# Toward Usable Access Control for End-users: A Case Study of Facebook Privacy Settings

**Maritza L. Johnson**

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

**COLUMBIA UNIVERSITY**

2012

# ABSTRACT

# Toward Usable Access Control for End-users: A Case Study of Facebook Privacy Settings

## Maritza L. Johnson

Many protection mechanisms in computer security are designed to enforce a configurable policy. The security policy captures high-level goals and intentions, and is managed by a policy author tasked with translating these goals into an implementable policy. In our work, we focus on access control policies where errors in the specified policy can result in the mechanism incorrectly denying a request to access a resource, or incorrectly allowing access to a resource that they should not have access to. Due to the need for correct policies, it is critical that organizations and individuals have usable tools to manage security policies.

Policy management encompasses several subtasks including specifying the initial security policy, modifying an existing policy, and comprehending the effective policy. The policy author must understand the configurable options well enough to accurately translate the desired policy into the implemented policy. Specifying correct security policies is known to be a difficult task, and prior work has contributed policy authoring tools that are more usable than the prior art and other work has also shown the importance of the policy author being able to quickly understand the effective policy. Specifying a correct policy is difficult enough for technical users, and now, increasingly, end-users are being asked to make access control decisions in regard to who can access their personal data. We focus on the need for an access control mechanism that is usable for end-users.

We investigated end-users who are already managing an access control policy, namely social network site (SNS) users. We first looked at how they manage the access control policy that defines who can access their shared content. We accomplish this by empirically evaluating how Facebook users utilize the available privacy controls to implement an access control policy for their shared content and found that many users have policies are incon-

sistent with their sharing intentions. Upon discovering that many participants claim they will not take corrective action in response to inconsistencies in their existing settings, we collected quantitative and qualitative data to measure whether SNS users are concerned with the accessibility of their shared content. After confirming that users do in fact care about who accesses their content, we hypothesize that we can increase the correctness of users' SNS privacy settings by introducing contextual information and specific guidance based on the their preferences.

We found that the combination of viewership feedback, a sequence of direct questions to audit the user's sharing preferences, and specific guidance motivates some users to modify their privacy settings to more closely approximate their desired settings. Our results demonstrate the weaknesses of ACL-based access control mechanisms, and also provide support that it is possible to improve the usability of such mechanisms. We conclude by outlining the implications of our results for the design of a usable access control mechanism for end-users.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgments

It would have been impossible for me to complete this work without the support of the wonderful people in my life.

I will be forever grateful to my adviser, Steve Bellovin. He patiently guided me through the transition from student to researcher, and happily answered all of my questions along the way. Angelos Keromytis, Sal Stolfo, John Karat, and Lorrie Cranor provided valuable feedback that helped shape this dissertation.

I was fortunate to have the opportunity to work with folks outside Columbia, each of them served as a mentor, a role model, and a friend. Stuart Schechter helped keep grad school fun by not only encouraging practical jokes but also taking it in stride when he was the target. Clare-Marie Karat graciously helped me identify and fill important gaps in my background while I was an intern at IBM. Janet Kayfetz went above and beyond, selflessly sharing her time and wisdom with me as a teacher and a friend.

My professors at USD encouraged me to undertake far more than I ever planned: Beth Simon, Christine Alvarado, and Lukasz Pruski incited my interest in graduate school and prepared me for the task. Early on, I had several excellent teachers who paved the way. From Mar Vista, Cathy Stutzman, Ricardo Gomez, Mark James, and David Holden.

I am thankful to have truly awesome friends: Curtis Chambers, Valerie Hajdik, Heather Burrus, Jennifer Hobdell, and Brittney Kuhn.

My husband David was an unyielding source of support and motivation. He listened to my stories of the highs and lows, and was always ready to encourage me to "just finish already."

I owe the most to my family. My extended family for their love and support, and finally my parents for giving me their best throughout my life.

For my parents, Wayne and Lupe

# Chapter 1

# Introduction

Policy-based systems are important in computing because they allow aspects of the system to be dynamically customized without the need for another cycle of requirements engineering and development. In this research, we consider policies as rules that define choices of system behavior. Policy-based systems and applications can provide valuable flexibility for organizations and users but, in practice, the difficulties of producing correct policies prevents policy-based systems from being used to their full potential. For our purposes, a correct policy is one that captures the intent of the user. Correct policies are critical because the system will enforce the policy as specified and problems will arise if the enforced policy does not match the policy author's intent. As noted by Cheswick et al., "The single most important factor of your firewall's security is how you configure it" [Cheswick *et al.*, 2003]. This concept is applicable to all policy-based systems.

Correctness is particularly critical to computer security policies where an incorrect access control policy can allow wrongful access to a protected resource or deny an individual's rightful access to a resource. Policy authoring tools must be designed using human-centered approaches, since a person will fill the role of the policy author and must be able to effectively use the tools. Policies are written by a person, or team of people, tasked with translating high-level management goals into implementable policies [Brodie *et al.*, 2005]. The aforementioned factors make policy-management an important topic for the field of human-computer interaction and computer security.

The importance of usability to access control mechanisms has been recognized since the early days of systems design [Saltzer, 1974]. Yet incorrect access control policies and inadequate policy management tools are still the norm. In 2004, Wool evaluated a set of firewall policies against known best practices and found that all of the rule sets had least one configuration error [Wool, 2004]. Even more recently, Das et al. identified several high level security issues in an organizational file server that were attributable to misconfigurations in the policy [Das *et al.*, 2010]. The policies examined in these examples were written by system administrators or, at the least, trained technical users. The usability requirement is much higher for a mechanism that will be used by non-technical users.

Access control systems are used in a wide range of domains and the policies are primarily managed by someone tasked with the role of the policy author. It is widely agreed that access control policies can be difficult to manage correctly even for an experienced system administrator, not to mention the difficulties experienced by knowledge workers or untrained end-users. As outlined in Chapter 3, quite a bit of research has focused on the study and design of usable policy management tools. In this thesis, we contribute to the growing body of work focused on end-users and usable access control policy management tools. Increasingly, end-users are asked to make access control decisions: who can access content they share on the Internet, which third-parties services can access the information either through an application or otherwise, which applications to install on a mobile device and which permissions to allow.

One place where non-technical end-users encounter access control policies is social network sites (SNS). SNS providers tend to present their access control mechanisms as “privacy settings” or “privacy controls.” It is through the manipulation of these features that an SNS user creates their access control policy. The user specifies who can access their shared content. In many social network sites (SNSs) the default policy is completely open — anyone on the Internet can view all of the user’s shared content. Some SNSs provide controls that allow a semi-restrictive policy — any of the user’s SNS connections can view her shared content. And some SNSs provide fine-grained controls that can be defined per person or per shared item.

In our work, we chose to study Facebook users and their ability to manage the accessibility of their profile content to other users. This provided an opportunity to observe the existing access control policies of end-users, one that for many users protects a large amount of personal information. Most of the prior work has studied people in a laboratory setting, or studied people using an experimental system that they only use for the purpose of the study. Additionally, it is common for a study design to require participants to protect synthetic data and complete tasks that they may never encounter in practice. Although some prior work has attempted to elicit users' desired policies in regard to an access control domain [Bauer *et al.*, 2008; Benisch *et al.*, 2011], there is a dearth of research on deployed access control policies. A goal of our research is to contribute to the literature an understanding of how end-users cope with managing an access control policy that determines who can access their personal shared content. For this reason, we studied how people use Facebook's privacy controls to protect their actual profile *in vivo*.

For the millions of users who share personal information like photos and status updates on Facebook, the ability to control who can view their content with confidence and ease is critical. Users find many benefits in socializing and sharing online including connecting with family members and old friends, but this experience can be dampened if users worry about the privacy of their shared data and might be concerned that they details of their personal life are being viewed or used by unintended audiences. Facebook provides privacy settings that allow users to specify which groups can view parts of their profile at varying levels of granularity. The granular controls allow the user to limit access to their profile as a whole, and the fine grained control allows the user to specify rules for individual content items. For example, a user could select a group of friends who can view an photo album of their vacation pictures, and preclude the rest of their friend network from viewing the pictures.

If the management of SNS privacy settings is analogous to managing an access control policy, then based on the prior work presented in Chapter 3 and the fact that even system administrators have trouble correctly configuring access control policies, we expect that SNS users have trouble managing their privacy settings correctly.

## 1.1 Thesis Statement

In this dissertation we focus on satisfying the need for a usable access control mechanism for end-users. We approach the problem by researching how end-users cope with managing an access control policy that they already use in their online activities — the access control settings for their Facebook content. In our preliminary work, we first assessed the correctness of Facebook users' current privacy settings compared to their sharing intentions. In the process of evaluating users' current settings, we discovered that presenting the user with examples of inconsistencies in their privacy settings is not sufficient to motivate the user to take corrective action. This revelation led us to conduct additional preliminary research to refute the claim that users are unwilling to take corrective action because they do not care about who accesses their shared content. From our preliminary work, we found reason to believe that the usability of an access control mechanism for SNS privacy settings depends on more than just a user interface that enables the user to correctly specify a specific access control rule or make changes to an existing rule set.

*Thesis Statement*: We found that SNS users do not use the available ACL-based privacy controls to protect their shared information in a way that is consistent with their sharing intentions. Based on this finding, we hypothesize that the correctness of a users' access control policy can be improved by supplementing the existing mechanism to include features such as (a) relevant contextual information and (b) specific instructions on how to use the SNS user can manipulate the privacy controls so that the new settings accurately reflect users' sharing intentions.

## 1.2 Contributions

Our research toward a usable access control mechanism for end-users contributes the following insights:

1. We present the first empirical research that uses end-users' sharing intentions to measure the correctness of their own deployed access control policies. Our results show that the privacy settings of many users do not match their sharing intentions. One example that illustrates this mismatch is exemplified by the group of users that stated

a desire to hide profile information from strangers when in fact they are sharing that content with 'Everyone' on Facebook.

2. We discovered that in some cases, even when end-users are made aware that an error exists in their privacy settings, simply presenting the information does not guarantee that the end-user will be motivated to take corrective action and make adjustments to fix the error.

3. Our research demonstrates the need to provide end-users with usable fine-grained access control mechanisms for social network sites. We contribute quantitative and qualitative data that show the wide variety of interpersonal privacy concerns SNS users experience and the ad-hoc methods they rely on to compensate for the shortcomings in the existing controls.

4. We contribute a novel approach for measuring SNS users' interpersonal privacy concerns by using a random sampling of individual social network friends and individual shared posts. We feel that this additional context increases the reliability of users' responses by grounding the survey questions in reality.

5. We increased the correctness of SNS users' privacy settings by introducing additional features to the existing mechanism to include viewership feedback and direct questioning, modifications do not require redesigning the existing access control mechanism.

6. We move beyond task-based laboratory user studies and measure the privacy settings that people are using in their actual lives. It is vital that we pursue this research direction as we work toward a usable access control mechanism that people will actually be motivated to use to protect their shared information. One aspect of this is, for example, understanding how the privacy settings fit into users' overall experience using Facebook.

## 1.3 Organization

Before we present prior work on the design and evaluation of usable access control management tools (Chapter 3), we first give an account of terms relevant to our research followed by an overview of pertinent Facebook features and privacy controls (Chapter 2).

In Chapter 4 we present an empirical evaluation of the correctness of Facebook users' privacy settings. We measure correctness by comparing their privacy settings against their self-reported sharing intentions. In this study, we found that not only did every single participant confirm at least one instance of over-sharing or under-sharing, but the majority of the participants claimed they would not change their privacy settings to fix the problem. We discuss several explanations for this result and evaluate the plausibility of the explanations in the two chapters that follow.

One explanation for Facebook users' unwillingness to change their privacy settings is that they do not care who sees their shared content. In Chapter 5, we present the results of a survey where we collected data from 260 active Facebook users to measure (1) are Facebook users concerned with audience outside their friend network viewing their shared content, (2) are Facebook users concerned with sharing content within their friend network, and (3) how do users with privacy concerns reconcile their concerns with the desire to share content on Facebook? We found overwhelming evidence that Facebook users do in fact care about who views their shared content. We also found that not only do they report to care, but most users employ privacy-preserving behaviors to control who can see their shared content, the content that is shared, and the membership of their friend network. The widespread use of mitigation strategies contradict the explanation that users would not fix their privacy settings because they do not care.

In Chapter 6, we investigate the possibility that users would be unwilling to correct their privacy settings due to a lack of situational awareness, a misunderstanding of which privacy control to manipulate, or a misunderstanding of how to use the existing privacy controls. We recruited 107 active Facebook users and first tested the effect of viewership feedback on the use of privacy preserving behaviors. We found that, in general, viewership feedback did not directly motivate participants to increase the use of privacy preserving behaviors.

We then tested whether users would address errors in their privacy settings if the mechanism alerted them to the fact that their privacy settings contradicted their stated sharing intentions, and (1) the mechanism identified the privacy control that would address the error and (2) gave instructions for how to use the control. We conducted this test on sharing preferences for photo albums and added it to the study of viewership feedback. To start, approximately 60% of our sample had at least one photo album with incorrect privacy settings. We found that participants who previously saw viewership feedback were significantly more likely to correct their inconsistencies, while the participants who did not see feedback made no changes to their privacy settings. This result suggests that the combination of viewership feedback and direct intervention does bring users closer to approximating their desired privacy settings.

Although the addition of feedback and direct intervention enabled some participants to correct inconsistencies in their privacy settings, many participants chose not to modify their settings. In Chapter 7, we discuss our findings with a focus on how they can influence the design of a better access control mechanism for SNS privacy settings, and the implications they suggest for progressing toward the goal of usable access control mechanisms for end-users.

# Chapter 2

# Background

## 2.1 Definitions

In this chapter we give an account of terms relevant to our research, followed by an overview of some key Facebook features and a brief timeline of Facebook's privacy controls.

Our research is concerned with increasing the correctness of access control policies — the set of rules that each attempted access to a system is compared against to determine whether the access is authorized. This process requires that "every access to a system and its resources be controlled and that all and only authorized accesses can take place" [Samarati and Vimercati, 2001]. A basic access control rule is composed of a subject, an object, and an action: the rule declares that the subject is allowed to enact the action on the object [Samarati and Vimercati, 2001]. An access control policy is a collective set of access control rules.

We consider access control policies to be the output of a larger policy authoring or policy management process. A process which encompasses several related tasks: the initial policy specification, the ability to "read" the existing policy and understand the effective policy, and the ability to make modifications and adjustments to the policy after the initial specification. Effective policy is "the function that results from a set of rules after the application of defaults and the resolution of conflicts." [Reeder, 2008]. A key aspect of policy authoring is the translation from an organization or individual's high-level goals into a set of policy rules in the allowable notation for the specific policy mechanism that satisfy the original high-level goals expressed [Karat *et al.*, 2009].

In our work we focus on the policy author, the person or group of people tasked with the role of policy author. The role of policy author is sometimes called policy implementer [Bauer *et al.*, 2009]. In some contexts an access control policy is created by two groups of people in two distinct roles: first the policy maker designs the policy by specifying high-level goals, then the policy author translates the high-level goals into a set of rules that capture the intent of the high-level policy and implements the rules using a policy management tool [Brodie *et al.*, 2005; Bauer *et al.*, 2009].

In our work the policy makers and implementors are Facebook users, and they manage their Facebook privacy settings to create their access control policy. The subjects in their policy are other Facebook users, and the objects in their policy are the individual pieces of content they enter on Facebook. As of late, the objects also include pieces of content that other users contribute on behalf of the policy author that become associated with their profile. The basic actions are to share content with the subject, or hide content from the subject. We also use the term *privacy controls* to refer to Facebook's privacy settings, these are the manipulatable user interface widgets (controls) in Facebook that allow a user to limit access to their shared content. The various ways that the subjects and objects can be grouped are discussed in Section 2.2.1.

## 2.2 Facebook Features and Terms

Facebook is the leading online social network service. In March 2012, the website reported the service had more than 900 million monthly active users [Facebook, 2012]. After signing up for an account, you are given a profile where you can submit personal information to share with other users and you can begin to build a friend network by establishing symmetric connections with other Facebook users. People in the friend network are referred to simply as "Friends" on Facebook. The average user has approximately 130 people in their friend network [Facebook, 2012].

In Facebook's early days, when the site was targeted at college students, a profile was composed of relatively static personal information. A Facebook profile included a person's name, birthday, gender, address, email address, phone number, schools attended, relationship status, political affiliation, religious beliefs, interests, favorite music, favorite books, favorite movies, and job history. Most of these fields are static in nature. Perhaps a user would modify the information at times, but one would expect the frequency of changes to be low. In addition to the information supplied by the user, there was a profile section called the *Wall* that acted like a bulletin board where people could leave timestamped messages to the user that were viewable by the profile owner and other users. Originally, it was required that each user be a member of a university *Network*.

Beginning in 2006, Facebook has introduced many new sharing opportunities including photo albums, status updates, notes, etc., giving people additional ways to share personal information, and creating a fertile ground for researchers [Boyd and Ellison, 2007].[1] The introduction of the *News feed* feature caused a shift in the way that Facebook users interacted with the service by providing users with a aggregate view of recent profile changes and recently contributed content. While it's true that the *News feed* feature did not change the availability of information, it undeniably changed the accessibility of shared information and users' activity.

Alongside *News feed*, Facebook introduced a feature to encourage users to share up to the minute information about what they were doing. The new feature, *Status updates*, consisted of an empty text box that prompted users to submit a brief message to complete the sentence, "*User's name* is ..." with their current mood, location, or activity. *Status updates* appeared on the primary user's profile and in other users' News feed.

### 2.2.1 Privacy Controls

Here we include a brief overview of Facebook privacy controls and describe how the features map to key aspects of an access control mechanism. We also include a timeline of significant changes to the privacy controls and default settings (see [Boyd and Hargittai, 2010] for additional information).

In the current Facebook privacy controls, the following groups are presented as options

[1] A comprehensive list of Facebook features accompanied by a description and the approximate date of introduction can be found on Wikipedia. `http://en.wikipedia.org/wiki/Facebook_features`

for the *Subject* field in access control rules:

- *Friends* - another Facebook user that the primary user established a symmetric connection with.
- *Friends of Friends* - a Facebook friend of one of the primary user's Facebook friends
- *Networks* - all Facebook users who are a member of a Facebook network that the primary use is also a member of.
- *Everyone* - any Facebook user.
- *Custom* - can be a combination of individual Facebook friends and groups of Facebook friends that have been previously organized into a custom friend lists.

When we began our research, the privacy controls used the type of information the *Object* field in the access control rules. For example, a Facebook user could specify from the possible list of *Subjects* who could see photos, videos, links, notes, and other types of information that could be shared on Facebook. One major change to this model was the introduction of per-post privacy settings. This feature introduced the ability to specify different privacy settings for an individual photo, instead of having the same privacy settings for all photos.

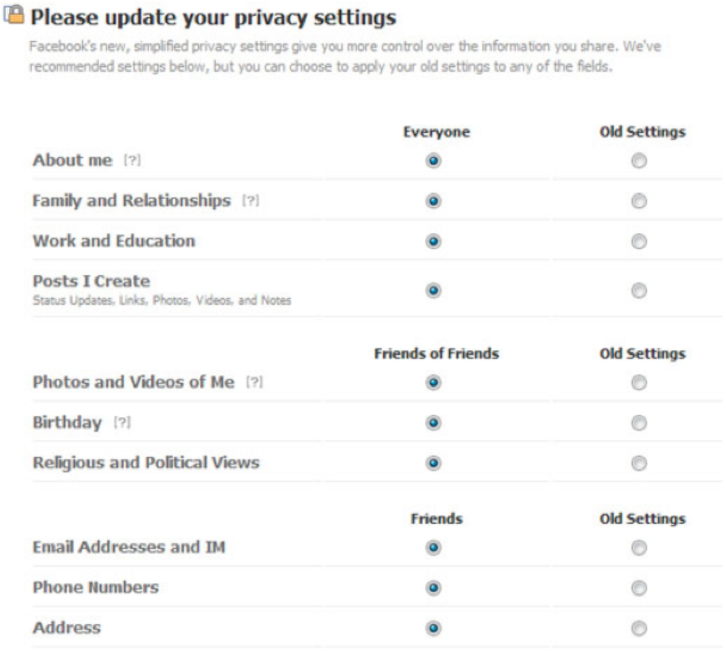
Since August 2011, the privacy controls are no longer based on data type. Now there is one privacy control with the options Everyone, Friends, and Custom that is applied to all posted content unless the user changes the privacy settings of a single content item when they post the item — in-line privacy controls.

#### 2.2.1.1 Timeline of Significant Changes

**2004** Facebook profiles are available to *Networks* by default.

**May 2006** The Networks feature is expanded to include other types of real world networks such as high schools, geographic regions, and workplaces.

**September 2006** Facebook changes the registration requirements and allows anyone to join Facebook.

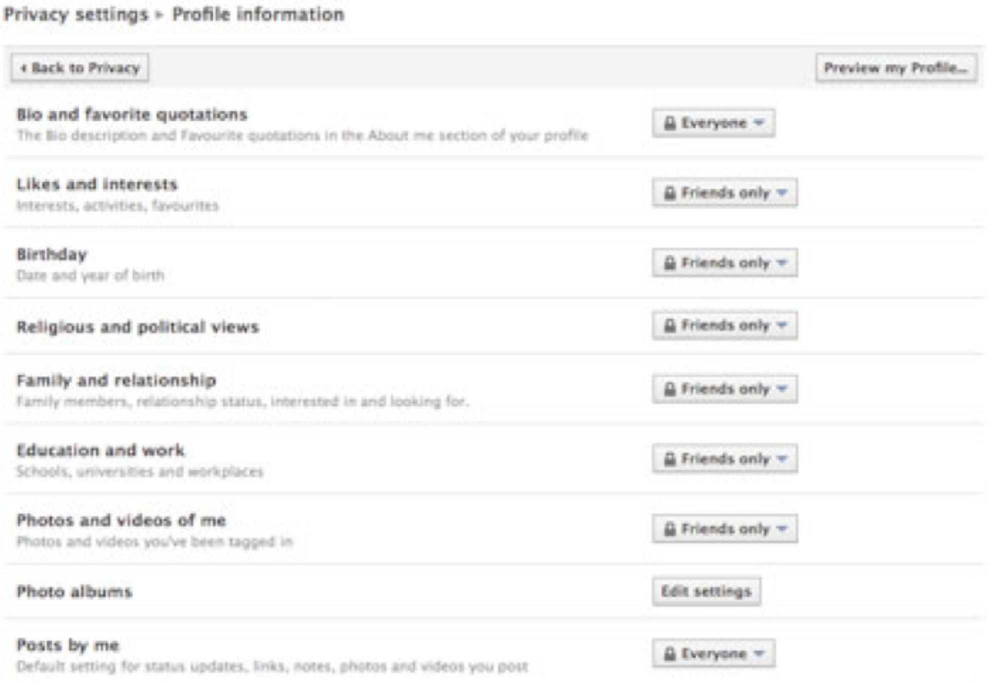


**Please update your privacy settings**

Facebook's new, simplified privacy settings give you more control over the information you share. We've recommended settings below, but you can choose to apply your old settings to any of the fields.

| | Everyone | Old Settings |
|---|---|---|
| About me [?] | ● | ○ |
| Family and Relationships [?] | ● | ○ |
| Work and Education | ● | ○ |
| Posts I Create<br>Status Updates, Links, Photos, Videos, and Notes | ● | ○ |
| | **Friends of Friends** | **Old Settings** |
| Photos and Videos of Me [?] | ● | ○ |
| Birthday [?] | ● | ○ |
| Religious and Political Views | ● | ○ |
| | **Friends** | **Old Settings** |
| Email Addresses and IM | ● | ○ |
| Phone Numbers | ● | ○ |
| Address | ● | ○ |

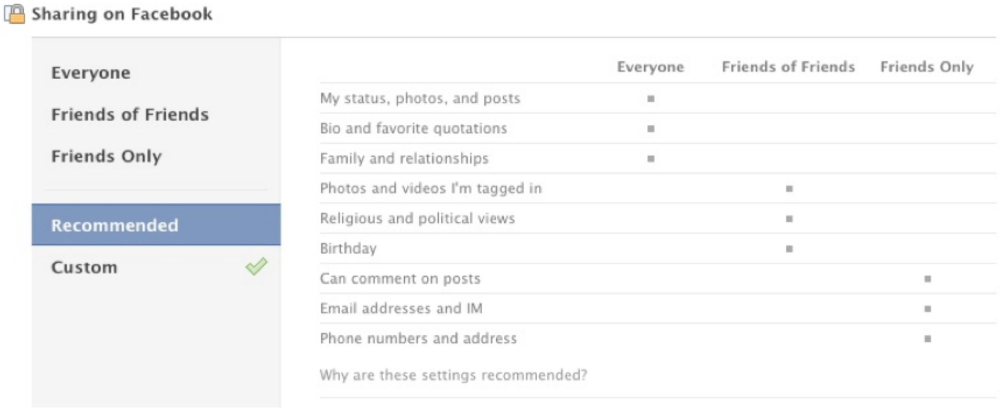
Figure 2.1: The dialog Facebook presented to users when they logged in December 2009. Users were forced to interact with the dialog before continuing to the website.



**Privacy settings ▸ Profile information**

| | |
|---|---|
| Bio and favorite quotations<br>The Bio description and Favourite quotations in the About me section of your profile | Everyone |
| Likes and interests<br>Interests, activities, favourites | Friends only |
| Birthday<br>Date and year of birth | Friends only |
| Religious and political views | Friends only |
| Family and relationship<br>Family members, relationship status, interested in and looking for. | Friends only |
| Education and work<br>Schools, universities and workplaces | Friends only |
| Photos and videos of me<br>Photos and videos you've been tagged in | Friends only |
| Photo albums | Edit settings |
| Posts by me<br>Default setting for status updates, links, notes, photos and videos you post | Everyone |

Figure 2.2: Facebook privacy settings before the redesign in 2010.



Sharing on Facebook

Everyone

Friends of Friends

Friends Only

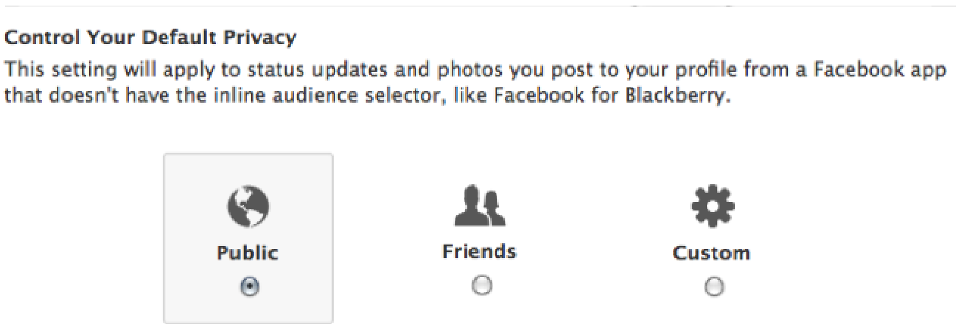
Recommended

Custom ✓

| | Everyone | Friends of Friends | Friends Only |
|---|---|---|---|
| My status, photos, and posts | ■ | | |
| Bio and favorite quotations | ■ | | |
| Family and relationships | ■ | | |
| Photos and videos I'm tagged in | | ■ | |
| Religious and political views | | ■ | |
| Birthday | | ■ | |
| Can comment on posts | | | ■ |
| Email addresses and IM | | | ■ |
| Phone numbers and address | | | ■ |

Why are these settings recommended?

Figure 2.3: Facebook privacy settings after the redesign in May 2010.



Control Your Default Privacy

This setting will apply to status updates and photos you post to your profile from a Facebook app that doesn't have the inline audience selector, like Facebook for Blackberry.

Public ●

Friends ○

Custom ○

Figure 2.4: Facebook privacy settings after the redesign in August 2011.

**2007** The default privacy settings share the user's name, profile picture, and networks to *Everyone.*

**November 2009** The list of people in the user's friend network is available to *Everyone* by default.

**December 2009** After making additional changes to the privacy settings, Facebook presented all users with a privacy controls dialog asking them to choose between keeping their *Current Setting* or sharing the content in question with *Everyone* (see Figure 2.1). Facebook also introduced per-post privacy controls, a feature that enabled the user to specify privacy settings for individual posts.

**May 2010** The main privacy controls user interface is changed from a list of settings (see Figure 2.2) to a tabular view (see Figure 2.3), though customizing the settings requires using the previous version of the user interface. *Recommended settings* are added to the tabular view that are quite open and also serve as the default settings.

**August 2011** Facebook debuts in-line privacy controls allowing the user to change the privacy settings of an individual post before the content is submitted (see Figure 2.4). The main user interface for privacy controls is redesigned to present one privacy control that is the default setting for all posted content. The choices are Everyone, Friends, and custom.

# Chapter 3

# Prior Work

Security and privacy policy authoring is an important topic for the usable security community and has received a fair amount of research attention. The majority of policy-based systems continue to rely on a human policy author, thus there is a strong need for usability and a user-centered design process. Much research has been conducted with the goal of improving the usability of access control mechanisms. Here we begin by presenting efforts to improve policy authoring in a specific domain, then present more general approaches to policy authoring tools.

## 3.1 Domain Specific Usable Policy Authoring

### 3.1.1 System Administrators and Access Control

To the best of our knowledge, Adage was the first instance of a concerted attempt to apply human-centered research principles to the design of an access control policy management mechanism [Zurko *et al.*, 1999]. Adage was an authorization service built on RBAC where the policy author can modify a policy using a textual or graphical user interface (GUI). Role-based access control was introduced in 1996 as an improved model for expressing authorization policies that control access to system resources: the crux of the new model was the introduction of a layer of indirection — rather than specify a policy such that permissions are assigned directly to a user, the policy author could create roles based on users' roles in the system, in this way permissions were assigned to roles, and users assigned to roles [Sandhu *et al.*, 1996].

In the GUI mode, the policy author is presented with a visual representation of the permissions, users, and roles in the system, and can use direct manipulation to author policies. The design was influenced by contextual interviews with security administrators where key features of their job and tasks were uncovered such as constant task switching, cooperating with colleagues to troubleshoot issues, and the prevalence of formulating the authorization policy while implementing it. A user study was conducted with security administrators where specific requirements were highlighted such as the ability to quickly understand an existing policy, and the importance of communicating states of conflict in the rule set. The authors concede that the features favored by the security administrators are unlikely to work for other users and that a different approach may be needed.

Firmato is a firewall policy management tool that was designed to minimize the errors in firewall policies [Bartal *et al.*, 2004]. Using an approach similar to Adage, Firmato contributed the ability to manipulate the policy using abstractions that hid the underlying mechanisms and network topology. Firmato also allowed policy authors to use named objects to represent groups of IP addresses and port numbers, and the GUI included features for specifying the policy, visualizing the topology, and visualizing the permitted and denied connections. We do not know how the new features impacted the correctness of firewall policies because no empirical tests were conducted.

Around the time that Firmato was designed, Wool published the results of a quantitative evaluation of 37 firewall policies [Wool, 2004]. Each policy was evaluated based on its compliance with twelve widely accepted practices for protecting a network. The practices were chosen to be an objective as possible, since the firewall policies were not paired with the administrator's goals or intentions. The conclusion was the most of the policies violated most of the established best practices. In other words, the policies were believed to be weak or inaccurate to some extent. This paper is notable in that the analysis was conducted on real firewall policies.

Bauer et al. contributed a set of requirements for designers of access control systems that are intended to be used by system administrators [Bauer *et al.*, 2009]. In contrast to the interviews associated with the research on Adage, the results of the interviews suggest that a policy is produced by people in at least two separate rules: the policy maker and the policy implementor. Recall that in the design of Adage, the administrators often created the policy as the were implementing it. A similarity to the Adage interviews is the finding that policy authoring is a collaborative task that is handled by multiple people. The interviews also revealed the difficulties system administrators experience when they find it is not possible to implement the desired policy due to limitations in the access control mechanism.

### 3.1.2 File Access Control

In response to the growing popularity of users sharing content on the web, ESCAPE was designed to allow users to configure access permissions implicitly through their actions [Balfanz, 2003]. To accomplish this, ESCAPE has a feature to allow the content owner to announce the existence of new material to intended recipients who receive an email notification with a customized link to the content. The solution relies on certificates for authentication and was not empirically evaluated.

Salmon was a file access control user interface designed to address the shortcomings of the Windows XP file permissions dialog [Maxion and Reeder, 2005]. A task-based laboratory study of the Windows dialog revealed that users were able to correctly complete tasks only about 25% of the time. Analysis of the results suggested that many errors were linked to the user interface's lack of an accurate overview of the policy state, or, as it is termed in the paper, a lack of "external representation of task-relevant information." To address this problem, Salmon's user interface presented the options to manipulate the policy alongside an overview of the effective policy, the addition of this information greatly improved participants' ability to correctly complete permission configuration tasks. IAM (intentional access management) took the focus on effective policy a step further by allowing users to specify the effective policy instead of entering a set of individual access control rules [Cao and Iverson, 2006]. An empirical evaluation of IAM demonstrated that with IAM users were able to author WebDAV policies more quickly and accurately than they did when they used an ACL-based tool to compose the individual rules.

In a study of file access control management in a corporate setting, Smetters and Good analyzed employees' file access control policies over the preceding ten years to measure how often policies were configured and to gain an understanding of the types of policies that were configured [Smetters and Good, 2009]. Toward the goal of understanding the requirements for fine-grained file access control, they found that in general users rarely set access control policies for individual files or folders but the policies are quite complex when they do.

Although much work in this area is focused at the user interface level, Reeder et al. empirically evaluated the effect of the conflict resolution policy on end-users' ability to author correct access control policies [Reeder *et al.*, 2011]. The results demonstrate that a specificity-based conflict resolution policy is more usable compared to a method where deny rules are given precedence.

### 3.1.3 P3P Privacy Policies

In the effort to design a privacy agent for P3P policies, Cranor et al. discussed the problems related to designing an interface for users to express their privacy preferences [Cranor *et al.*, 2006]. Most problems were related to finding the appropriate vocabulary to express the complex language found in privacy policies and structuring the interface to group common elements.

In 2008, Reeder et al. conducted a user study to measure participants' comprehension of typical website privacy policies expressed in natural language compared to a P3P policy presented with the Expandable Grid [Reeder *et al.*, 2008b]. The study measured comprehension with a set of survey questions, time to answer each question, and satisfaction with the presentation method. The results show that participants performed similarly with the two formats, and did not achieve better comprehension with the Expandable Grid nor were they able to answer questions faster.

### 3.1.4 Access to Physical Resources

Grey is an access control system that allows end-users to manage access to physical resources, a sample use case for Grey is managing the access policy for who can enter rooms and laboratory space in a university setting [Bauer *et al.*, 2008]. Research with Grey has explored the implications of allowing the user to express a policy that is more fine-grained than is possible with physical access control mechanisms (like a lock and key). In a limited deployment of Grey, the researchers found that resource owners were able to specify more accurate and more restrictive policies using Grey's fine-grained options compared to the access policies they enacted with keys.

This paper made several contributions to the study of usable access control systems. First, it introduced a new metric for measuring the effectiveness of an access control system by quantifying the gap between the resource owner's ideal (desired) policy — the one they would implement in the absence of limitations inherent to a system, and the implemented policy. In this particular study, the participants described their ideal policy and the policy they had in place with physical keys. Second, it was the first paper to report on how users manage an actual access control policy they use in their daily life.

### 3.1.5 Location Sharing

PeopleFinder is an example of a location sharing application, the application allows other people to request the location of the user and the user can specify privacy preferences to restrict who can access their location and when [Sadeh *et al.*, 2009]. In this context, it is critical to the adoption of the technology that users are able to correctly manage their privacy preferences. Sadeh et al. researched the amount of expressiveness users require for the privacy preferences of a location sharing application. Their results show that it is difficult for users to accurately specify their sharing preferences, and suggest that additional research is needed to understand users' needs. The results also suggest that perhaps a priori settings are not well-suited to the dynamic nature of social applications.

In a related study, Benisch et al. collected detailed data on end-users' preferences for sharing their location data with different groups of people over the course of a three week period [Benisch *et al.*, 2011]. Participants were given a mobile phone that collected their location throughout the day and each day the participant was asked to review their location logs and state for each observation who they would be willing to share the location with and under what conditions. Similar to prior interview-based research on location-sharing preferences, they discovered that the identity of the requester is an important factor in the decision to share [Lederer *et al.*, 2004; Consolvo *et al.*, 2005], but they found that other factors also affected users' decisions such as the time that the request was made and the location. The research looked at several options for specifying location sharing preferences and tested which would best satisfy the sharing preferences collected. This step took place offline, based on metrics predefined by the researchers, without participant feedback. They found that the most fine-grained option would offer the best match for a users' preferences but would also incur the highest burden on the users' time and attention.

Tsai et al. conducted a user study to measure the effect of feedback on users of a location sharing system [Tsai *et al.*, 2009]. They found that the introduction of location request feedback increased users' comfort in sharing their location, and actually reduced users' privacy concerns. Overall, users were encouraged to configure their privacy settings to allow more requests when feedback was present in the system.

Tang et al. found evidence that users' sharing preferences and attitudes toward location sharing change depending on their reason for using the location sharing application [Tang *et al.*, 2010]. The survey measured users' sharing decisions for social-driven applications and compared the responses against decisions for purpose-driven applications. The results demonstrate that users' attitudes differ by purpose and suggest that designers of privacy controls for these applications may benefit from being aware of such differences.

### 3.1.6 Social Networking Sites

Audience View is a policy authoring tool that gives the user immediate visual feedback as they make policy modifications [Lipford *et al.*, 2008]. With Audience View, the policy author configures privacy settings for a profile on a social networking website and the user is able to view their profile as different groups of users would see it. Rather than manage a list of rules, the user clicks on the profile to show and hide information. An empirical study shows the visual feedback contributes to better usability than the list-of-rules interface on social networking websites.

Egelman et al. conducted a laboratory study to evaluate how Facebook users react to limitations in the privacy controls [Egelman *et al.*, 2011]. The existing privacy controls were compared against an experimental interface that introduced ambiguity detection and provided actionable guidance. The results demonstrate that using the existing privacy controls many participants failed to complete the task, and that only ambiguity detection with actionable guidance improved participants' ability to complete the tasks. Many participants complained that they needed more options for fine-grained control.

Many SNS services present privacy controls as decisions about which users can view types of content (e.g., photos, updates, or posts), Klemperer et al. took a different approach and evaluated the usability of privacy controls for photos based on content tags[Klemperer *et al.*, 2012]. The results show that users understood the tag-based system and felt they could correctly control access to their photos with such a system.

## 3.2 General Usable Policy Authoring

Although the majority of research on usable access control mechanisms has focused on identifying and addressing usability issues for a specific domain, there has been some research on more general approaches to usable policy authoring.

SPARCLE began as a policy management workbench designed for people who author privacy policies for organizations [Brodie *et al.*, 2005; Karat *et al.*, 2005]. The research initially focused on applying a human-centered design process to meet the needs of privacy policy authors while providing strong ties across previously disjoint tasks: policy authoring, implementation, and compliance. Later work demonstrated the extensibility of the SPARCLE workbench to other domains [Brodie *et al.*, 2008]. In SPARCLE, the policy author could author policies using their choice of guided natural language or structured entry and switch between the formats as necessary. The interface for guided natural language displayed a syntax guide above the text area where the policy author typed the policy, which increased users ability to write correctly structured policy statements. After the policy was written, a natural language grammar was used to extract the policy elements from the natural language statements. To author a policy using structured entry, the policy author selected values for policy elements from predefined lists. Empirical studies showed the use of either authoring method produced higher-quality policies compared to unguided natural language [Karat *et al.*, 2005].

After analyzing the errors observed in a laboratory study on policy authoring, Reeder et al. recommended the following guidelines for the design of usable security and privacy policy authoring tools: support object grouping, enforce consistent terminology, make default rules clear, communicate and enforce rule structure, and prevent rule conflicts [Reeder *et al.*, 2007]. The research concluded that the application of these guidelines to the design of the policy authoring interface would have prevented the mistakes that were observed.

Expandable Grids is an interactive matrix for policy visualization and authoring that was designed as an alternative to the list-of-rules interface. The primary motivation behind the design was the realization that users need a reliable representation of the effective policy — the end result of the list of rules. Expandable Grids communicates the effective policy to the policy author throughout the policy authoring process, this feature increased participants' ability to author correct access control policies compared to their inability to use the Windows XP file permissions dialog correctly [Reeder *et al.*, 2008a]. The mechanism uses a table like view to present a policy element on each axis where the intersection is the policy decision for the pair of values. For example, if one axis represents files and the other axis represents users the intersection of the two represents what actions the user can perform on the file (if any). Expandable Grids was later demonstrated as applied to the presentation of P3P policies [Cranor *et al.*, ]. The results of the empirical study suggest the usability of Expandable Grids for the task is at least as high as the readability of natural language P3P policies [Reeder *et al.*, 2008b].

Lipford et al. conducted an empirical evaluation to compare the usability of Audience View and Expandable Grids and found that even though participants performed relatively well with both interfaces they identified disadvantages and advantages inherent to each design [Lipford *et al.*, 2010]. Although the researchers predicted performance differences would occur on specific tasks because of the different input and display styles, none of the predictions were realized. Participants even suggested an interface that combined the methods might increase usability, this observation is similar to SPARCLE's feature that allows multiple entry methods for the same policy.

## 3.3 Human Computer Interaction and Computer Security

In 1975, Saltzer and Schroeder listed psychological acceptability alongside other well-known security principles in their seminal paper on information protection in computer systems:

> It is essential that the human interface be designed for ease of use, so that users routinely and automatically apply the protection mechanisms correctly. Also, to the extent that the user's mental image of his protection goals matches the mechanisms he must use, mistakes will be minimized. If he must translate his image of his protection needs into a radically different specification language, he will make errors. [Saltzer and Schroeder, 1975]

The importance of the principle of psychological acceptability has grown because of the shift from all computer users being technically trained individuals to the present day where all types of people use computing devices in their jobs and in their personal lives. Whitten and Tygar outlined several distinct features of security systems for designers to keep in mind [Whitten and Tygar, 1999]. One important difference is that security is typically a secondary goal for users: typically the user's primary focus is on completing some task on their computer, and the user might expect security, or might think of security only as an afterthought.

# Chapter 4

# Measuring the Correctness of Facebook Users' Privacy Settings

Recently, there have been several news stories that highlight unwelcome uses of Facebook profile information: a man found himself discredited in his divorce proceedings after a photo from his profile of him drinking and partying was presented in court [Chen, 2010], a teacher was fired for a negative comment she made about the school on Facebook [Heussner, 2010], and a man's Facebook posts were used to question his injury in a workman's compensation claim [Little, 2012]. Why would someone publicly share information that could negatively impact their offline life?

One possibility is that the user deliberately and intentionally shares this information publicly. The other possibility is that the user unintentionally shares this information publicly because of problems inherent to the usability of the privacy controls.

In fact, each of the scenarios previously mentioned could have been avoided through the diligent use of the available privacy settings. Enabling SNS users to control who can access their shared information is critical to empowering the user to avoid undesirable sharing situations. Aside from anecdotal evidence, there is a paucity of quantitative data to support the perception that users are unable to accurately specify the intended audience for their shared information. In this chapter, we present a study where we empirically evaluate the question — do Facebook users' privacy settings accurately reflect their sharing intentions? Are their *actual* privacy settings correctly configured?

## 4.1 The Study

We empirically evaluated the preferences and behavior of Facebook users to determine if SNS users' privacy settings match their sharing intentions. We chose Facebook because of its overwhelming popularity: the company itself claims over 900 million monthly users [Facebook, 2012]. In this chapter we describe an empirical study with three parts: a survey to measure privacy attitudes, a questionnaire to collect sharing intentions, and a results phase where participants indicate whether potential violations represent an inconsistency between their sharing intentions and privacy settings. Privacy attitudes have previously been measured in various settings and laboratory studies have identified usability issues with Facebook's privacy settings and features. Nonetheless, our empirical study is the first to measure correctness by comparing sharing intentions against users' *actual* privacy settings in a real SNS.

Our results show that overwhelmingly, privacy settings do not match sharing intentions. That is, SNS users are sharing and hiding information incorrectly as judged by their beliefs. Furthermore, a majority of participants indicated that they could not or would not fix the problems. The prevalence of such errors — every participant had at least one incorrect setting — suggests the current approach to privacy controls is fundamentally flawed and a different approach is needed.

### 4.1.1 Background and Related Work

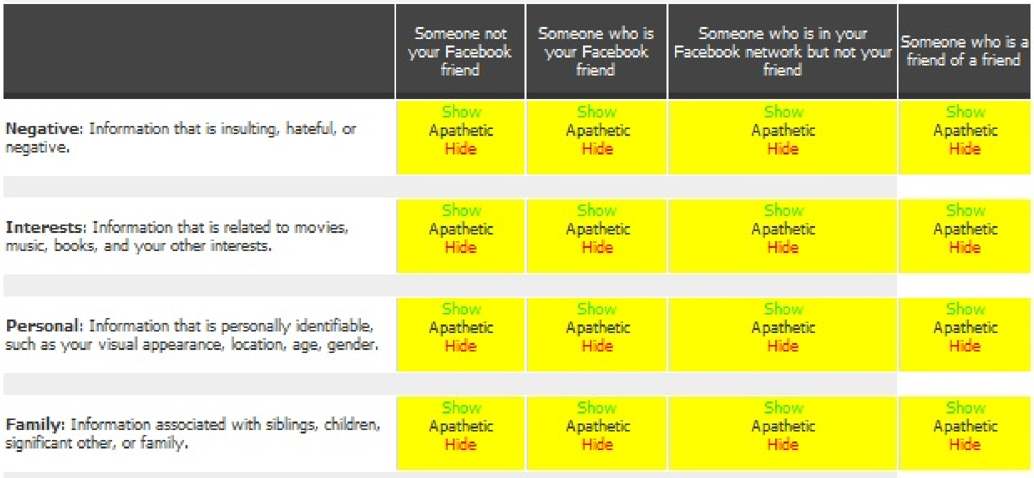
The study presented in this chapter draws upon many themes including users' motivations for sharing and communicating using an SNS, studies measuring the use of privacy settings, and the research on the usability of access control mechanisms discussed in Chapter 3.

In 2006, Acquisti and Gross measured the accuracy of Facebook users' perception of their level of disclosure on the site by surveying users on the visibility of their profile then comparing their answers against the amount of data that was available to all members of the participants' university network. A low number of participants (8% of 209 participants) were sharing more that they thought they were and some (11%) were sharing less than they thought, but in general most (77%) participants had an accurate perception of the amount of available information [Acquisti and Gross, 2006]. This study is similar in nature to our study, except it only measured users' awareness of the publicness of their profile, it did not measure users' sharing intentions. The study was a follow-up to an earlier study that passively measured information disclosure on Facebook [Gross and Acquisti, 2005]. In 2005, Gross and Acquisti analyzed 4,540 Facebook profiles to measure the information that was available and found that the majority of users shared a large amount of personal information on their profile, yet fewer users chose to limit access to just their friend network(0.06%).

Facebook has made significant changes to the website since 2006 (see Chapter 2 for details); it is now open to anyone, and many new features and privacy options have been introduced. The proportion of users who utilize the available privacy settings is also much different since at least 2008. Krishnamurthy and Wills measured the number of public profiles in 20 regional networks and found 53-84% of profiles were public [Krishnamurthy and Wills, 2008]. For some regional networks, the number of public profiles is much smaller than the 99.9% that were public in 2006.

Lewis et al. measured the behavior of a separate university Facebook network to determine the prevalence of privacy settings use and identify factors that contribute to a 'taste for privacy' [Lewis *et al.*, 2008]. The study measured 33.2% of the 1,710 participants had a profile that was private to the university network. This number represents the number of people who had manipulated their privacy settings in some way since the default setting at the time was that a profile was accessible to network members.

An investigation of privacy settings is incomplete without understanding how users want to share and their goals for using an SNS. Along these lines, prior research has found that many SNS users primarily interact with people they know offline. In a study of motivations for using Facebook, Joinson found that most users utilize Facebook for "keeping in touch" with people with whom they have an offline relationship with, this includes looking up information about friends and communicating with friends [Joinson, 2008]. Lampe et al. also researched how users interact with Facebook and reported similar motivations and uses [Lampe *et al.*, 2006]. Joinson found that users' privacy settings varied based on their



| | Someone not your Facebook friend | Someone who is your Facebook friend | Someone who is in your Facebook network but not your friend | Someone who is a friend of a friend |
|---|---|---|---|---|
| **Negative:** Information that is insulting, hateful, or negative. | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide |
| **Interests:** Information that is related to movies, music, books, and your other interests. | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide |
| **Personal:** Information that is personally identifiable, such as your visual appearance, location, age, gender. | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide |
| **Family:** Information associated with siblings, children, significant other, or family. | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide | Show<br>Apathetic<br>Hide |

Figure 4.1: The table used to collect participant's sharing intentions in Stage 2.

motivation for using Facebook. This point is critical to our evaluation — users find different uses for SNSs and their privacy settings should vary accordingly.

## 4.2 Method

In our study we investigated whether users' privacy settings match their sharing intentions. We implemented the study as a Facebook application which allowed us to conduct the study remotely. Each participant completed the study in two sessions. Prior to installing the study application, the participant read a consent form that explained the study and they reviewed the requested privileges in the application installation dialog.[1]

### 4.2.1 Stage 1: Survey

The study began with a survey to measure the user's privacy priorities, confidence in existing settings, Facebook usage, history of privacy violations, and exposure to privacy-related media coverage. We present the questions alongside the results in Section 4.3.1.

[1] Columbia University Protocol IRB-AAAF1543

| **Category** | **Description** | **Sample Keywords** |
|---|---|---|
| Religious | Information related to religion. | 'god', 'priest', 'torah', 'mosque' |
| Political | Information related to politics. | 'obama', 'republican', 'climate' |
| Alcohol | Information related to alcohol. | 'drunk', 'beer', 'keg' |
| Drugs | Information related to illegal drugs. | 'weed', 'smoked', 'toke' |
| Sexual | Information related to sexual relations. | 'sex','porn','hooker' |
| Explicit | Information with explicit words. | 'sh*t', 'f*ck' |
| Academic | Information related to academics. | 'homework','professor', 'lecture' |
| Work | Information related to work. | 'boss', 'internship', 'interview' |
| Negative | Information related to a negative opinion.. | 'hate', 'sucks', 'ugly' |
| Interests | Information related to interests, such as movies and T.V. shows. | 'band', 'movie', 'book' |
| Personal | Information that is personally identifiable, such as your visual appearance, location, age, gender. | 'birthday', 'new york' |
| Family | Information associated with siblings, children, significant other, or family. | 'father', 'sister', 'mom' |

Table 4.1: The list of information categories for collecting sharing intentions.

### 4.2.2 Stage 2: Collection of Intentions

We asked the participant to report their sharing intentions using a table where the columns displayed profile groups and the rows displayed information categories. In each cell, the participant indicated their attitude toward sharing the information category with the group. The choices were show, apathetic, and hide. Most privacy interfaces provide two mutually exclusive options of sharing or hiding information. These options, however, can not fully capture user intent. There may be information where a user does not have a strong opinion. For this reason, we included apathetic as an option when recording sharing intentions.

Our study focused on the default groups that are currently used in Facebook privacy settings: friends, friends of friends, network members, or everyone. Privacy settings can also be configured using custom friend lists though we did not measure this. In Facebook, a *friend* is another Facebook user that the user has confirmed or initiated a connection. A *friend of a friend* is a Facebook user who is the friend of one of the user's friends. In Facebook, people can be members of 'networks' that represent offline entities like geographical regions, schools, or companies. A *network member* is another Facebook user that is associated with one of the same 'networks.' We renamed the default group 'everyone' to *stranger* to indicate the absence of a relationship (i.e. not a network member, friend of a friend, or friend). In reality these groups may overlap; however, the study focused on profiles that fit in exactly one group.

We collected sharing intentions based on information categories instead of data types (e.g., photos, notes, links, events, and status updates). When the study was designed, Facebook's privacy settings presented configuration options by data type. A new feature, at the time, also allowed users to configure settings on a per post basis, which we did not study. The information categories were based on textual content, rather than data type, and spanned all data types. We collected sharing intentions to assist in the identification of potential violations. For this reason, we chose categories that users were likely to have a strong opinion about (the information categories are listed and described in Table 4.1).

| Intent | Result Based on Privacy Settings | Potential Violation |
|---|---|---|
| Hide | At least one object matched the category and was accessible to the profile group. | No |
| Hide | All objects that matched the category were hidden from the profile group. | Yes |
| Show | At least one object matched category and was hidden to the profile group. | No |
| Show | All objects that matched the category were accessible to the profile group. | Yes |

Table 4.2: The possible results of our violation identification process.

### 4.2.3 Stage 3: Identification of Potential Violations

The application identified potential sharing violations by comparing the participant's sharing intentions with their privacy settings. First, the application compiled a list of the information categories where the participant indicated a show or hide intention (apathetic intentions were ignored since they cannot produce a violation). Then the application classified the participant's profile data using our information categories. Next, the application iterated over the classified items and checked the privacy settings for the four profiles groups. The application recorded the identifier and type of violation when there was an inconsistency between the participant's intention and privacy settings. Stage 3 produced two lists: a list of the posts where the participant intended the category to be shown but the post was hidden, and another that included the posts where the participant intended the category to be hidden but the post was visible.

To classify the participant's posts using our categories, the application inspected all textual data associated with the participant's profile and activity. To execute this, the participant needed to grant the application permission to access their profile data including all posts that the participant had shared on their own profile, the posts the participant had made on their friends posts, and the posts the participant's friends contributed to their profile. The application classified the posts using sets of keywords. We created the sets of keywords manually, prior to recruiting, by collecting unique words that were common to each category. We did this by consulting sources such as existing Facebook data, terminology lists, and tags on related online content.

To determine the privacy settings for each post, we created four profiles to represent the default profile groups. We created the profiles such that they were mutually exclusive. The *friend* profile had a single friend which was the profile used to check the privacy settings for *friend of a friend*, we sent a friend request from the *friend* profile to the participant before the study began. *Stranger* did not have any friends and was not a member of any networks. The *network member* was a member of the Columbia University network and did not have any friends. Only *network member* was a member of the Columbia University network. At the time, the Facebook API did not allow direct access to the privacy settings for individual items.[2]

We define a *hide violation* to be the case where the participant's intent is to hide the information category from the profile group, but one or more object in the category is accessible. We define a *show violation* to be the case where the participant's intent is to show the information category to profile group, but one or more object in the category is not accessible (see Table 4.2).

To compile the set of potential violations, we compared the participant's sharing intentions against the privacy settings for each object that matched a category with a show or hide intention. The study application completed this task by attempting to access each object using the four study profiles and recording the actual privacy settings. The application considered all of the data associated with the participant's profile that was accessible through the API (the history of all Facebook interactions) and matched one of the following categories: *disclosed*, *entrusted*, and *incidental data* [Schneier, 2010]. The following lists demonstrate the shared objects that were evaluated by the study application; the lists are not exhaustive.

**Disclosed data (posted by the participant)** status updates, comments on status up-

[2] After our data collection was completed the privacy field became accessible for photo albums. The date that this change was implemented is not documented in the API.

dates, photo captions, comments on uploaded photos, links, comments on links, album captions, album comments, video captions, comments on videos, notes, comments on notes, and comments on wall posts, basic profile information and page memberships (e.g., About Me, and interests).

**Entrusted data (posted by the participant on another user's profile)** comments on status updates the participant was tagged in, comments on photos the participant was tagged in, comments on photo albums the participant was tagged in, comments on videos the participant was tagged in, comments on notes the participant was tagged in, public event RSVPs (title, tagline, type), and public group memberships (title, type, subtype).

**Incidental data (posted by a Facebook friend on the participant's disclosed data)** comments on status updates, comments on status updates the participant was tagged in, comments on the participant's photos and tagged photos, comments on participant's links, comments on participant's albums, comments on participant's videos and tagged videos, comments on participant's notes and tagged notes, and comments on participant's wall.

Facebook does not provide privacy controls that allow the user to control access to entrusted data. Despite the lack of privacy controls, in almost all cases the user has the option to delete entrusted data. For this study we modify the definition of *control* to mean a lack of control via the privacy settings. Thus, the user can delete the object but cannot *hide* the object [Schneier, 2010].

### 4.2.4 Stage 4: Confirmation of Violations

In the final stage, we asked the participant to review the potential violations and confirm the actual violations. We also asked the participant whether they would correct the confirmed violations. In this stage, the participant proceeded through twelve screens: one screen per information category that was divided into four sections, one section per profile group. In the case the application had identified a potential violation for the profile group and information category, the application presented the potential violation to the participant and asked the participant whether it contradicted their sharing intentions.

Our algorithm for identifying potential violations was designed to liberally assign categories to increase the chance of identifying actual violations. For potential violations, the application retrieved the object in question and displayed it to the participant. The justification (i.e. matching keywords) for the potential violation was shown in boldface to provide the participant with context. Within each section the potential violations were grouped based on the source (whether the data was posted by the participant or a friend) and on the data type (photo comment, group, event, status update, etc.). We asked the participant to confirm the potential violations. This is a key step that is novel in our study design, previous studies have only guessed at potential violations; it is not possible to distinguish an actual violation from a potential violation without knowing the user's sharing intentions.

### 4.2.5 Participants

Recruitment methods were targeted at the Columbia University community and included flyers, broadcasts to Facebook groups, broadcasts on mailing lists, and a paid advertisement on a campus blog. The final sample was a convenience sample of students who responded to the advertisements. A total of 65 people completed the study (38% male). The average age was 21.3 years ($S.D. = 1.90$). We compensated the participant $10 for their time.

We focused on checking the settings for the four default profile groups to measure the user's privacy settings; the Facebook API does not allow applications to directly query for access control settings. Thus, we created four new profiles, one to match each profile group, to conduct the study. In theory, the only profile that would be difficult to create is a network member. However, in our case this was trivial given our affiliation with Columbia University. The need to have a profile in the same network, to test the network member settings, restricted our recruiting to Columbia University students.

A perfect random sampling of Facebook users requires knowledge of the demographics and usage habits of Facebook's user base which is information that only Facebook has. Given the unequivocal nature of many of the results reported in this paper, the sample

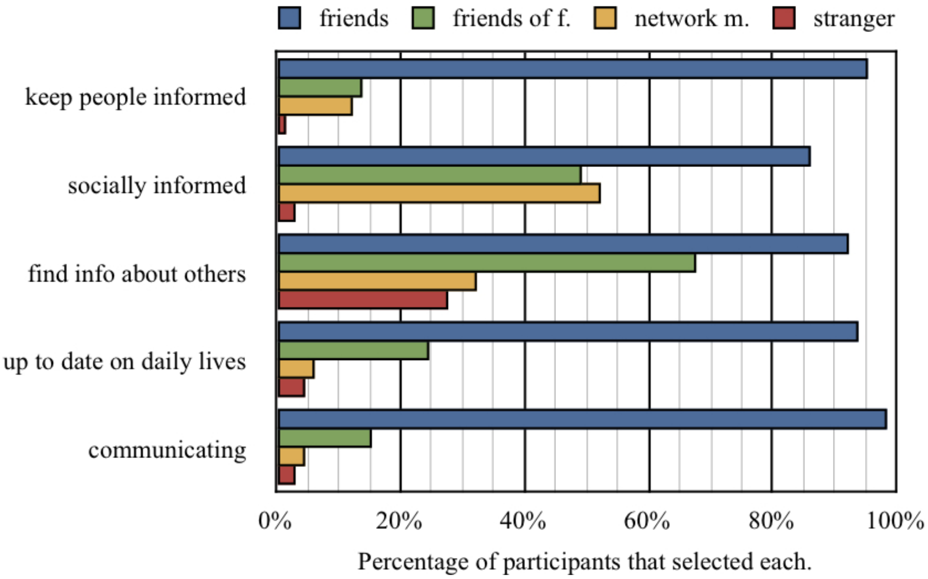


Figure 4.2: Participant responses to, "Why do you use Facebook?"

size is sufficiently large to provide insight to the trouble users experience with the existing privacy settings.

## 4.3 Results

In this section we present the results from each stage of the user study. We collected data October and November 2010. The data confirm that users are concerned with the privacy of their SNS profile data, however their privacy settings are not aligned with their sharing intentions. Moreover, many participants reported they do not intend to take corrective actions to the inconsistencies they confirmed.

### 4.3.1 Survey of Privacy Attitudes

Here we present the survey questions alongside the results (see Appendix A for a complete list of the survey questions).

First we asked, "What is the most important reason for online privacy?" Half (49%) the participants selected reputation security — to hide information to protect social reputation The next most poplar answer was economic security (39%) — to prevent identity

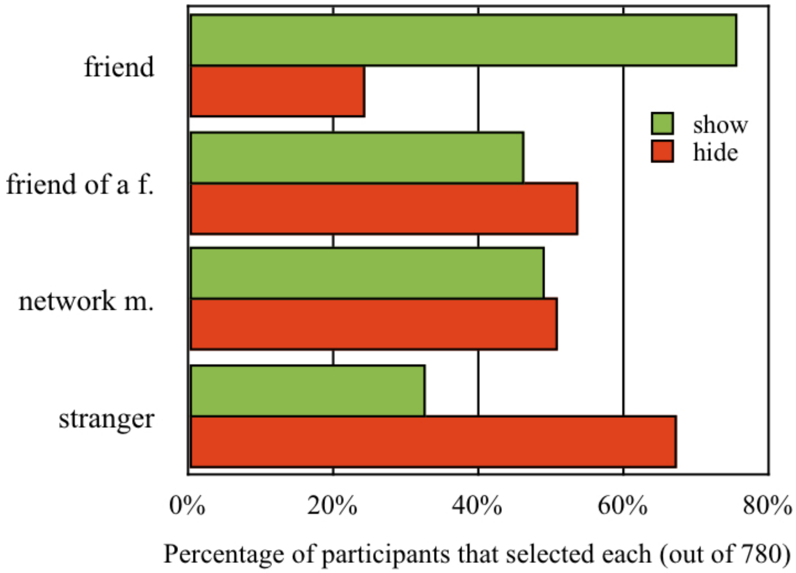


Figure 4.3: The participants' sharing intentions for each profile group. Each participant reported a total of 48 sharing intentions.

theft and protect browsing habits from advertisers. The least important reason (12%) was physical security — to ensure physical safety, by hiding your face, location, and/or contact information from strangers.

Then, we asked the participant to report their level of concern with economic, reputation, and physical security. The answer options ranged from "why would I be concerned?" to "I'm very concerned", with "I'm a little concerned" as the midpoint. The responses were converted to a numerical score from 1 to 5 (very concerned). In regard to economic security 63% ranked their concern as a 4 or 5, 71% ranked their concern with reputation security as a 4 or 5, and 42% ranked their concern as a 4 or 5 for physical security.

We asked how often they untag photos and described a few scenarios when a user might untag a photo. Most participants (94%) had untagged a photo because "I didn't like the photo of me (it was unattractive or unflattering)" and most (94%) had untagged a photo because "the photo displayed behavior I did not want to be associated with (something that could be embarrassing if others saw it)."

We asked whether they engage in five activities with the four default groups (presented as a table of 20 checkboxes): "keep people informed about my life," "finding information about people ," "finding information on people's daily lives (e.g. newsfeed)," "personal communication (e.g. messages, walls)," and "being socially informed (e.g. events, groups)." Participants reported to interact with 'friends' the most and 'strangers' the least (see Figure 4.2).

We asked, "Do you feel your Facebook settings reflect your attitude related to privacy?" Nearly every participant (95%) responded affirmatively ($CI_{.05} = 5.3$). We asked, "Have you ever had an accidental leak of information on Facebook that had a negative impact?" Most participants (91%) responded that they had "never had an accidental leak of information on Facebook."

We asked, "Have you heard anything regarding Facebook and privacy lately in the news lately?" Most participants (85%) had heard something from a general news source. We also asked participants, "Has the media coverage affected your behavior on Facebook?" Some (29%) replied the media had not affected their behavior at all. Those who answered yes (n = 46) could select more than one of the options listed: nearly all of them (83% of the 46) "became more selective about the information I post on Facebook," some (22%) deleted a Facebook friend, and most (91%) claimed to have modified their privacy settings to be more private.

### 4.3.2 Sharing Intentions

We asked the participant to state their sharing intentions across twelve data categories for four groups, then, for analysis, we combined show and apathetic intentions (Figure 4.3). Participants were willing to share most categories with a 'friend' (76%). Less than one-third of the categories were selected to be shared with a 'stranger' (see Figure 4.4). A few categories drew a large number of hide intentions for all groups like sexual, negative, drug, and alcohol.

Female participants selected more categories to share with friends and less to share with strangers. We computed a contingency table chi-square test on the frequency of show intentions for male and female participants. The difference in the number of sharing intentions between male and female participants is significant ($\chi^2(7) = 51.2$, $p$ <.0005).

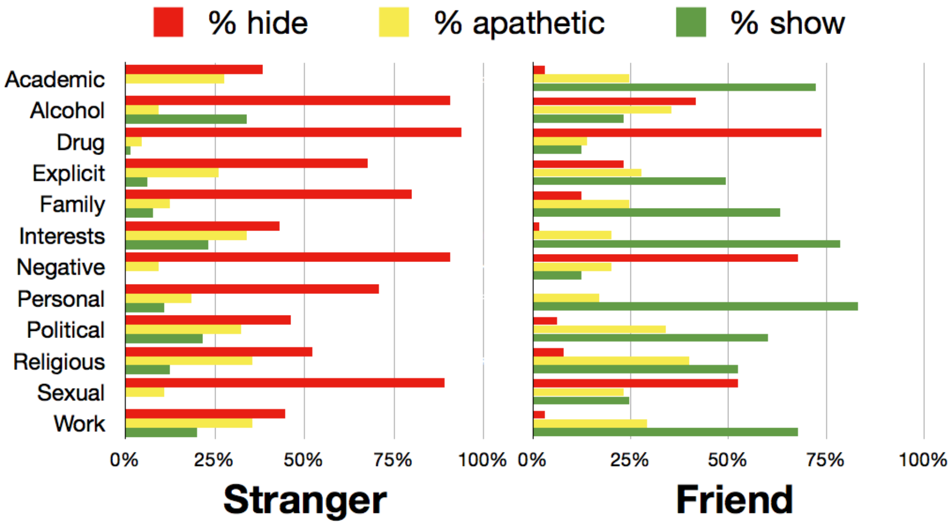


Figure 4.4: The collected sharing intentions by category for the *Stranger* and *Friend* group.

### 4.3.3 Confirmed Violations

Every single participant confirmed at least one sharing violation ($CI_{.05} = 2.42$): 94% of participants confirmed a hide violation — they were sharing something they intended to be hidden ($CI_{.05} = 5.77$), and 85% of participants had at least one show violation — they were hiding something they intended to be shared ($CI_{.05} = 8.68$). We recorded a total of 1191 confirmed violations across the sample ($M = 18$ per participant, $S.D. = 10.5$). More than half of the violations we recorded were hide violations (778 total, $M = 12$ per participant, $S.D. = 9.0$). Show violations represented 35% of the confirmed violations (413 total, $M = 6$ per participant, $S.D. = 5.7$).

For each confirmed violation, we asked the participant whether they would take action based on the violation, then estimated the perceived severity of the violation using their response. Even though every participant confirmed at least one sharing violation, only 58% of participants reported they would take action in response to at least one. Nearly all participants (97%) had at least one confirmed violation that they did not plan to address.

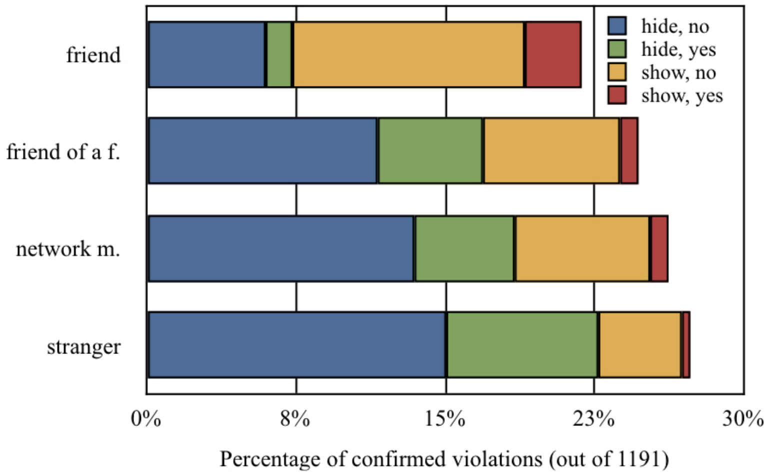


Figure 4.5: The percentage of confirmed violations presented by group. Each bar is divided into hide and show violations, then further divided to show the proportion of violations that elicited action.

In Figure 4.5, we present the confirmed hide and show violations per profile group, each bar is further divided based on the reaction to the violation. Overall, the distribution of violations across the four profile groups is nearly balanced, however, the composition of the violations differs by group. For example, 'friend' had the most show violations and 'stranger' had the most hide violations. Hide violations were more likely to motivate action (30% of 778 hide violations), especially for the non-friend groups (stranger = 12% of 778 hide violations, network member = 8%, and friend of friend = 8%). In general, the participants are not motivated to correct show violations (85% of 413 show violations), though show violations that involve the friend group are slightly more likely to motivate action (8% of 413 show violations). While some violations motivated changes, the most frequent response was 'no action' (76% of 1191 confirmed violations).

In Figure 4.6, we present the confirmed violations by information category. The hide violations most likely to motivate action were categorized as alcohol, sexual, explicit, and religious. The show violations most likely to motivate action were categorized as family, personal, and religious were most likely to motivate action. The high number of violations for academic (14% of 1191) may have been an artifact of our sample of students. Similarly,

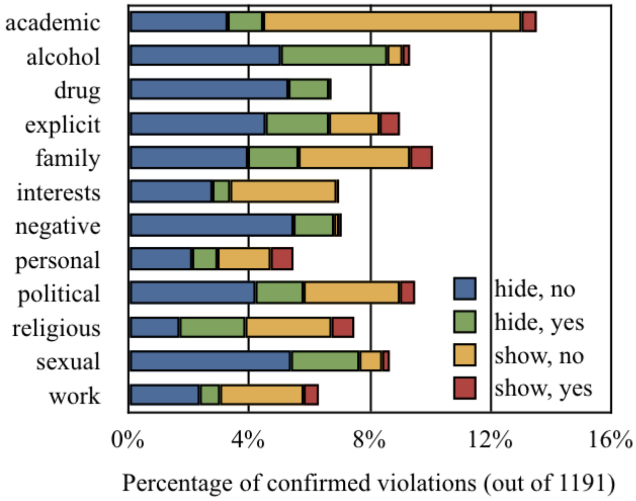


Figure 4.6: The percentage of confirmed violations divided by information category. Each bar is divided into hide and show violations, then further divided to show the proportion of violations that elicited action.

the high number of hide violations for alcohol (9% of 1191) may have been due to the fact that many of our participants were under the legal drinking age.

## 4.4 Discussion

We measured the accuracy of users' Facebook privacy settings by comparing their sharing intentions with their actual privacy settings. We found that every person in our sample had at least one confirmed inconsistency between their sharing intentions and privacy settings. Even though it's unclear why participants reported they would not correct many of the violations, the existence of these violations presents a clear message: not only are Facebook's existing privacy settings flawed but improvements must be made to minimize risk to users.

A subsequent study of the correctness of Facebook users' album privacy settings reports results that confirm our findings. In May 2011, Liu et al. asked Facebook users to report their ideal audience for ten of their photos and measured the correctness of their privacy settings using the actual privacy settings for the photos [Liu *et al.*, 2011]. The results reveal that more than half (63%) of the photos participants were questioned about had incorrect privacy settings. Both our study and this study evaluated the correctness of users' privacy settings though we took complementary approaches: we considered all text-based shared data and Liu et al. considered photos.

We suspect that the basic access control mechanism used by Facebook is irreparably flawed. Previous studies on the usability of access control mechanisms (e.g., [Reeder and Maxion, 2005; Egelman *et al.*, 2011]) have shown that this style — a list of items, and a set of permissions for various users which must be set manually by the owner — is difficult to use. A drawback of past studies is that they use contrived scenarios, and synthetic data which users may not feel motivated to protect. A benefit of studying Facebook is that the data is personal, and users are, presumably, motivated to protect it. Our results, however, show that even with personal data our participants were not able to protect it successfully — an unfortunate result given that our survey data indicate they are concerned with privacy, take steps like untagging or deleting content to protect their privacy, and believe their privacy settings are correct. Furthermore, the results of a related study suggest that users do not understand the limitations of the current Facebook mechanism [Egelman *et al.*, 2011]. We believe it is reasonable to conclude the problem is inherent in the basic design.

One explanation for users' unwillingness to correct confirmed violations is that users do not care about privacy. This explanation is unlikely given that almost every participant supplements the existing privacy settings by untagging and deleting content. Such privacy preserving behaviors have been observed in other research as well, such as a survey conducted by Pew Internet in 2010 that reports in the 18-29 age group about half of surveyed SNS users had deleted unwanted comments (47%) [Madden and Smith, 2010].

We speculate that one of the largest culprits for privacy flaws is Facebook's reliance on data types (e.g., photos, events, and status updates) for defining privacy settings. These data types are misrepresentative of the real world that Facebook attempts to model. Outside of a social network, an individual does not determine visibility of personal data by format but instead by the context of the content. A key improvement would be to automatically categorize information with a predicted context, and define privacy settings per context that reflect the user's intent. Our data demonstrate that users are opinionated about showing or hiding specific categories of information. Prior work has explored the possibility of using content-based access control for blog posts; further investigation is necessary to determine if a similar approach would be beneficial to an SNS environment [Hart *et al.*, 2008].

Incorporating privacy settings that rely on information categories may also improve the process of introducing new sharing features. Currently, new features, such as Facebook Places, are implemented by introducing entirely new privacy settings and have historically been introduced with default settings that require the user to actively manage the new settings to reach a level of privacy that is consistent with their previous level. Our suggested information tags feature makes it feasible to infer settings for unanticipated features. For example, in the case of Facebook Places, if a user wishes to hide alcohol related information from everyone, it is reasonable to conclude that all location check-ins at a bar should be hidden from everyone. This method of inference provides a reasonable compromise between user privacy intentions and SNS providers' bias toward data visibility.

The recommended privacy settings contradict how people interact with other Facebook users. The responses to our question about how users interact (Figure 4.2) and the overall sentiment expressed in the sharing intentions (Figure 4.3) indicate that SNS users have little to no use for 'strangers.' The ideal recommended settings would reflect the needs of a majority of users; we estimate that users who wish to share their profile publicly are in the minority.

We wrote the survey in May 2010, and collected data October through November 2010. The popular media coverage of Facebook and privacy was quite abundant after Facebook changed the privacy settings user interface and we were concerned that users' privacy settings would reflect the coverage. We have no way to verify whether the people who claimed to modify their privacy settings in response to the media coverage of Facebook actually did. Our results suggest that even if changes were made additional changes are needed

We purposefully avoid a thorough treatment of the data categories used in the study to avoid unduly emphasizing their importance. The data categories used in the study were chosen based on the assumption that they would help identify sharing violations, particularly ones where users would have an opinion. Future work could select more representative categories while investigating a more sophisticated classifier.

Determining the root cause of violations is one possible follow-up study; this is better suited to an in-person interview (as opposed to the remote study reported here). An in-person study would provide a format where the study coordinators could adjust the questions to identify the source of the violation. Participants who have violations may not understand the privacy settings well enough to identify the reason behind a violation.

### 4.4.1 Limitations of Study Design

Typically, a sample of students is a weakness but for our study it may be an advantage. Most of our participants were tech-savvy and experienced Facebook users. Also, students will almost certainly be on the job market in the near future, which means the correct use of privacy settings is critical, and this demographic is likely to be sensitive to that. The size of our sample is defensible given the extreme nature of our results, i.e. that every participant had at least one violation.

The Facebook API does not allow applications to directly query the privacy settings of a user. To work around this issue we chose to focus on checking the privacy settings for the four default profile groups since it is simple to create a new profile and ask participants to accept a friend request, create a second profile as a friend of the first, and create a third profile to represent a stranger. In theory, the only profile that would be difficult to create is a network member. However, in our case this was trivial given our association with Columbia University. The necessity of having a profile in the same network restricted our recruiting to Columbia University students. In hindsight, this restriction was unnecessary, particularly given our finding that users have similar attitudes toward *Network Members* and *Friends of Friends*, and the deprecation of the feature.

Our statistics on confirmed sharing violations are a lower bound. We hypothesize that, in practice, each participant has more violations than were counted, which is an artifact of our identification algorithm and the study design. Across the 65 participants, the study instrument identified a total of 70,402 potential violations ($M = 1083$, $S.D. = 1056$). Rather than present each violation to the participant individually, the application grouped potential violations by data category then by profile group and asked the participant to answer each question based on correctly classified data and true violations. Furthermore, our algorithm only classified the textual posts, a future study might identify additional violations if photo content and videos were also considered.

In regard to the participant sample, the prerequisite of installing a Facebook application and granting full offline access to the study application may have biased the sample. During a pilot study we received several comments from potential participants who opted-out because of this requirement. We did not, however, receive similar feedback when recruiting on campus.

We are unable to analyze the nature of the observed violations beyond the analysis presented in the Section 4.3.3 because our application did not store the content that was a potential violation or confirmed violation. To respect the privacy of our participants, we chose to minimize the amount of data we collected and collected only the data necessary to determine the correctness of their privacy settings. When the study application identified potential violations we only stored to the database the content item's Facebook identifier, the keywords that were used to categorize the item, and the set of profile groups that could access the item. This was the information needed for the confirmation phase of the study.

## 4.5 Significance

The data we collected on users' sharing intentions indicate clearly that users have preferences for who they want to share their content with. Yet 100% of our participants confirmed that in at least one instance they were either sharing their personal information with people they want it hidden from or they were hiding information from people they want to share it with. This discrepancy clearly indicates that our participants' Facebook privacy settings are incorrect, and that their settings do not accurately express their sharing intentions. Furthermore, even though our participants confirm that their settings are incorrect, they are nonetheless unwilling to correct their privacy settings.

We know that the SNS participants in our study have sharing intentions; we also know that their privacy settings do not reflect these intentions and that the settings are incorrect; and we know, sadly, that they are not willing to adjust their settings so that the they reflect their deliberate intentions. How could this be? One possible interpretation is that our users are lazy or even exhibitionist. But the more realistic interpretation is that our users simply do not know how to correctly manage their privacy settings because the mechanism is too opaque. It could be that the actual privacy controls are too difficult to interpret or that the choices the user wants to implement are not even available as options. Either of these explanations would lead to the problems that we observed — that our users' privacy settings do not reflect their sharing intentions.

We are thus able to conclude that Facebook's privacy controls have serious flaws that must be fixed. The seriousness of the matter is demonstrated by the severity of the possible ramifications — people are experiencing consequences from SNS use that affect them offline, like losing a job. Of course, it would be a gross exaggeration to state that everyone will experience such a negative consequence as a direct result of SNS use. Nevertheless, many people are sharing personal content via SNS services and it is critical that they have usable controls to accurately limit the audience for their shared content.

# Chapter 5

# Facebook Users' Privacy Concerns and Mitigation Strategies

In Chapter 4 we provide quantitative evidence that Facebook users do not correctly specify their desired access control policy using the available privacy controls. A new mechanism is clearly needed, or at a minimum significant changes must be made to the existing mechanism. We suggest the application of human-centered design processes will produce a more usable design and initiate the process by posing the question — what are users' privacy concerns?

Prior work has investigated motivations for using social network sites and has also investigated how users interact with other SNSs users. Notably missing from the body of work is an understanding of users' privacy concerns. In order to design a new mechanism we must first understand what information users want to protect and from whom they want to protect it from in order to design the necessary privacy controls. With this in mind, we approach the question of users' privacy concerns by investigating the people and content that cause concern, and the strategies users employ to mitigate their concerns. We consider this process of understanding SNS users' privacy concerns akin to building a threat model for SNS users, an important first step when addressing any security problem. For the purpose of this Chapter, we limit our scope to privacy concerns related to sharing with other SNS users — interpersonal privacy.

## 5.1 Our Study

We constructed an interactive Facebook application to survey 260 Facebook users about specific pieces of content that they had posted to their profiles, as well as their levels of comfort sharing content with randomly selected people from their friend networks. We observe that many participants deny access to their profile to people outside their friend network, thus effectively mitigating the chances of experiencing a privacy mishap with a stranger. We also found that users increasingly experience sharing concerns that involve people who are members of their friend network, and out of necessity users have developed a range of coping strategies to address these concerns.

We begin by discussing prior research on SNS usage, and prior studies of SNS users' privacy concerns and mitigation strategies. Then, we present our methodology, including a description of the Facebook application we implemented to execute the study. Next, we present data on the privacy concerns users experience and the techniques they employ to mitigate their concerns. We conclude by discussing how our data demonstrate a shift in privacy concerns from situations that involve outsiders to situations that involve people within the friend network. Based on our results, we suggest a new focus for future research and highlight aspects of our approach that could be adjusted to ensure meaningful progress toward the goal of usable privacy controls for SNS users.

## 5.2 Background

Our research is on SNS users' ability to manage their interpersonal privacy while sharing and interacting with other users. We subscribe to Altman's definition of privacy and equate interpersonal privacy to " an interpersonal boundary regulation process used by people to regulate their interactions with others" [Altman, 1975]. We also adopt his terminology for desired privacy — a user's ideal level of privacy, and actual privacy — the achieved level of privacy, or the outcome from the boundary regulation process. Palen and Dourish's discussion of digital privacy is also relevant, specifically the point that privacy is dynamic and requires users to satisfy constraints that vary across contexts [Palen and Dourish, 2003]. We focus our discussion of related work on the topics of Facebook privacy and the available privacy controls, the difficulties users face in their attempts to manage interpersonal privacy, qualitative studies of SNS users' privacy concerns, and users' strategies for mitigating their concerns.

The aggregate of media reports and the results of prior work create a perplexing view of SNS users and privacy. Despite the multitude of privacy concerns that accompany SNS use [Gross and Acquisti, 2005], the number of people who are active SNS users continues to grow and users feel there are real benefits to interacting with others via an SNS and are motivated to share personal information online [Joinson, 2008]. Privacy concerns related to the use of Facebook have grown with the addition of new features and an expanding user base. Originally, Facebook membership was limited to university students, and the default privacy settings were configured to allow 'network members' access to user content. In 2005, Gross and Acquisti found that only 0.06% of a university network — three people — had changed the default settings [Gross and Acquisti, 2005].

Since 2006, Facebook has introduced many new sharing opportunities including photo albums, status updates, notes, etc., giving people more ways to share more personal information, and creating a fertile ground for researchers [Boyd and Ellison, 2007] (see [Boyd and Hargittai, 2010] for an overview of the evolution of Facebook's features and privacy controls). In light of the increase in the amount of content shared and the increase in the number of users, recent research results indicating that users' actual privacy settings do not match their sharing intentions are particularly troubling [Madejski *et al.*, 2012; Liu *et al.*, 2011]. These results confirm those of earlier work. In 2006, Acquisti et al. found that a significant minority of users were aware of the privacy settings available [Acquisti and Gross, 2006]. A later study by Egelman et al. indicated that users have difficulty configuring Facebook privacy settings to satisfy task requirements in a laboratory setting [Egelman *et al.*, 2011]. Some of the difficulties that participants experienced were related to a failure to understand the limitations of the privacy settings.

### 5.2.1 SNS Users' Privacy Concerns

To identify distinct categories of SNS users' privacy concerns, Krasnova et al. held focus groups with university students in Berlin about their privacy concerns related to the use of Facebook [Krasnova *et al.*, 2009]. The most frequent theme was concern over unwanted audiences viewing shared content, where the list of audiences mentioned included future employers, supervisors, family members, peers, and subordinates. Participants also frequently mentioned "organizational threats" related to the collection and use of their data by the SNS provider and third parties. Concerns about social threats were another common theme for concerns including people purposefully posting content to harm the individual, and general concern over a lack of control over the actions of other users.

Tufekci investigated the relationship between users' privacy concerns and their level of disclosure on an SNS, and found no relationship [Tufekci, 2008]. Even users who expressed many privacy concerns divulged large amounts of personal information on their profiles. However, the study only asked about the relatively static fields of a profile like age, sex, gender, religion, political affiliation, interests, and favorite books, rather than concerns over dynamic content (e.g., status updates, comments, etc.).

In a three year longitudinal study of university students, Lampe et al. found that users' imagined audiences for their profiles were changing over time [Lampe *et al.*, 2008]. For example, in 2008, significantly more users expected family members had viewed their profiles compared to 2006 (an increase from 49% to 70%). Similarly, more students thought a total stranger might have viewed their profile (24% in 2008, 14% in 2006). Some of the changes in attitude can be attributed to the evolution of Facebook sharing features and privacy settings.

The shift from Facebook as a social network for universities to a social network for everyone forced users to adapt to a new model of sharing: suddenly users' friend networks included coworkers, family members, and friends from other life stages, in addition to classmates. Interested in understanding the tensions that arise from a heterogeneous friend network, Lampinen et al. conducted semi-structured interviews with twenty people about their friend networks and their methods for managing group co-presence. They reported that many users fear that a boss or acquaintance might see something embarrassing that was not intended for them, and that users attempt to avoid these situations through self-censorship and using context to carefully selecting a suitable communication medium.

Skeels and Grudin also studied the dynamics of group co-presence, but focused on SNS usage in the workplace, and found that users have trouble coping with the co-presence of coworkers and other contacts in an SNS friend network [Skeels and Grudin, 2009]. Many people noted the burden associated with constantly maintaining an awareness that the two groups are present in their audience. Participants also noted the need to limit access to selected content based on relationship.

Wang et al. conducted a survey of SNS users' privacy concerns to compare the attitudes of American, Chinese, and Indian users [Wang *et al.*, 2011]. The survey covered topics such as demographics, usage habits, attitude toward sharing various types of content, desire to restrict access, and the use of fabricated profile data. The results show that, in general, the American participants were the most privacy concerned, followed by the Chinese participants. The Indian participants were the least privacy concerned.

These studies provide a strong foundation for the observation that protecting content from unwanted audiences is more than simply a matter of preventing strangers from accessing profiles. However, we are unaware of any large-scale studies that have attempted to quantify the extent to which users are sharing content inappropriately with members of their friend networks through the use of users' previously posted content or questions about specific friends. We also build upon previous work by recruiting a more generalizable sample, rather than members of a particular institution.

### 5.2.2 Strategies for Mitigating Privacy Concerns

In terms of users' strategies for mitigating their privacy concerns, SNS users regulate their interactions with others using many techniques and not all are based on the official privacy controls. Young and Quan-Haase identified boundary regulation mechanisms that include deleting tags, and using direct messages to limit audiences [Young and Quan-Haase, 2009]. Stutzman and Kramer-Duffield found that users who employed supplemental privacy preserving behaviors, like curating the posts on their wall and collaboratively adjusting SNS behavior among friends, were more likely to have a "friends only" profile [Stutzman and Kramer-Duffield, 2010].

Several papers have reported that users cope with conflicting social spheres by maintaining separate profiles, limiting access to subsets of the friend network, carefully selecting a communication medium, or using separate SNSs for different audiences [Stutzman and Hartzog, 2009; Skeels and Grudin, 2009]. PEW Internet reports that in 2011, 63% of Facebook users had removed someone from their friend network [Madden, 2012], an increase compared to the 56% of users who reported to have "unfriended" someone in 2009. The same survey found deleting and untagging posts to be common among all user demographics.

Some users resort to changing their offline behavior to mitigate their privacy concerns. In a study of sharing photos in an SNS, Besmer and Lipford found that users adjusted their offline and online behavior to mitigate their privacy concerns: participants reported avoiding having their pictures taken in the first place, untagging photos, and asking friends to remove photos rather than adjusting privacy settings [Besmer and Richter Lipford, 2010].

### 5.2.3 Summary

Usable privacy controls are critical to SNS users' boundary regulation process. While it may be possible for some users to achieve their desired privacy without the help of technical mechanisms, it is unlikely that this is the case for all users considering the wide range of privacy concerns and the overhead involved with using ad hoc techniques. Usable privacy controls are needed, but first a thorough understanding of users' privacy concerns is necessary such that the design can optimize the number of concerns addressed and the number of users who benefit.

Prior work leaves an important question unanswered — which privacy concerns are rampant enough that they ought to be designed for in the controls, and which mitigating behaviors are prevalent enough to motivate the design of new privacy controls? Prior work has demonstrated the wide range of users' privacy concerns and that users manage their concerns through a number of techniques other than the use of the access controls. This suggests that the existing access controls can be improved. However, it is unreasonable to expect that an access control mechanism will prevent all users from ever sharing content inappropriately. Therefore, we need metrics to determine how often problems currently occur and what would be an acceptable failure rate [Egelman and Johnson, 2012]. In our study, we attempted to answer the former.

## 5.3 Method

We designed our study to collect data on SNS users' interpersonal privacy concerns and situate their answers in context with the composition of their friend networks and the sensitivity of content posted to their profiles.[1] We also collected data on users' strategies for mitigating privacy concerns. We chose to focus on Facebook based on the functionality of the API and because of the large user base. We instrumented the survey as a Facebook application; this enabled us to pose questions using real profile data. We were specifically interested in how users manage their friend network, their use of Facebook's privacy features, and whether users had privacy concerns with content they had already posted to their profiles (e.g., photos and comments).

### 5.3.1 Survey Content

The survey had three sections. In the first section, we asked participants general questions about their Facebook usage so that we could compute correlations with real and perceived privacy risks. In the second section, we asked participants to report their level of concern with general scenarios describing situations with common unwanted audiences. Finally, in the third section we used the API to ask questions about individual Facebook friends and shared posts.

#### 5.3.1.1 General Usage

We asked about participants' SNS habits to measure the activities users engage in most often, the amount of time spent on each activity, the relationship between the user and the people in their friend networks, which privacy features are used, and whether other means of controlling access to information are employed.

#### 5.3.1.2 Concerns with Unwanted Audiences

SNS users run a risk of unwanted audiences accessing their information. Previous work asked users to report the perceived likelihood of specific audiences viewing their profiles

[1] Columbia University IRB Protocol AAAI1077.

(e.g., employers, law enforcement, thieves, political parties, or sexual predators) [Young and Quan-Haase, 2009]. We reused many of the scenarios that were used by Young and Quan-Haase, but instead of asking participants to guess the likelihood that the scenario was already occurring we asked participants to rate their level of concern — unconcerned to concerned on a 5-point Likert scale — that "each scenario could happen by using Facebook." We asked these questions to examine participants' levels of concern in the general sense before asking similar questions about specific posts selected from their shared content.

#### 5.3.1.3 Incorporating Profile Data

In the final section of the survey, we used the Facebook API to select content from participants' profiles and ask questions that incorporated the content. The questions were designed to ascertain the composition of participants' friend networks and the perceived level of sensitivity of content that they and others had posted to their profiles. The decision to incorporate profile data into the survey questions was driven by the assumption that it would contextually ground the questions, a technique that improved users' recall when making location-sharing decisions[Venkatanathan *et al.*, 2011].

For instance, if a participant indicated that she would not want coworkers to see a specific photo, and her friends network included coworkers, this may indicate a situation where fine-grained control is needed to manage access to content within her friend network, particularly if access to the photo was not restricted to anything more granular than friends. In this manner, we attempted to quantify the frequency with which privacy violations may be occurring. That is, previous work has focused on qualitatively describing privacy concerns, whereas we were interested in measuring the likelihood with which these concerns come to fruition.

We asked questions about nine randomly selected friends for each participant to gain an understanding of how SNS users know the members of their friend networks, as well as to measure how much they trust their friends with access to their profile information. For each of these friends, we asked the following questions:

1. What is your relationship to *FRIEND-NAME*?

| | General | Friends | Posts |
|---|---|---|---|
| Member of your immediate family | ✓ | ✓ | ✓ |
| Member of your extended family | ✓ | ✓ | ✓ |
| Coworker | ✓ | ✓ | ✓ |
| Someone you know from high school, college, or grad school | ✓ | ✓ | ✓ |
| Friend of a friend | ✓ | ✓ | ✓ |
| Someone you have not met in person | ✓ | ✓ | ✓ |
| Someone you socialize with in person | | ✓ | ✓ |
| Not sure | | ✓ | |
| Stranger | | | ✓ |

Table 5.1: We asked about common Facebook audiences throughout the survey. 'General' shows the groups used for a question about general friend network composition. 'Friends' shows the groups presented for the classification of individual friends. 'Posts' shows the groups used in questions about individual posts.

2. How do you feel about *FRIEND-NAME* viewing all the information you have uploaded to Facebook?

To determine whether participants were being diligent in describing their friend network, we also asked these questions about a fictitious friend whose profile picture we took from a free stock photo archive. Thus, we asked these questions for ten friends, nine of whom were actually members of their friend networks, the tenth was a randomly assigned male or female fictitious person.[2]

To understand the type of content a user might be uncomfortable sharing and why, ten posts were randomly selected from the participant's profile. We posed eight questions to measure their level of comfort with sharing the post. The audiences used for this set of questions are listed in the last column of Table 5.1.

[2] The fake profile picture appeared as the fifth in the sequence of ten. We observed that 83.1% of participants correctly answered that they did not know this person (95% CI [78.0, 87.2]), though found no correlations with demographics or Facebook usage and thus did not analyze this further.

| | **Sample** | **Facebook** |
|---|---|---|
| Age | | |
| 18-24 | 16% | 25% |
| 25-34 | 48% | 25% |
| 35-54 | 31% | 30% |
| 55+ | 5% | 11% |
| | | |
| Gender | | |
| female | 75% | 55% |
| male | 25% | 43% |

Table 5.2: Comparison of our sample's demographics to the demographics reported by istrategylab.com. The table shows only the age groups that are present in our sample, and istrategylab's numbers for gender total 98% (2% of users were recorded as unknown).

### 5.3.2 Participants

We recruited participants via ResearchMatch, a website that pairs researchers with potential participants.[3] The recruiting email did not mention privacy, it requested "Facebook users to take a twenty minute survey on their Facebook usage habits." As compensation, participants were entered in a drawing for one of five $100 gift cards. We received completed surveys from 260 respondents, ages ranged from 18-62 ($\mu = 33.8$, $\sigma = 10.6$).

Facebook does not publish detailed demographic data and so we rely on the statistics reported by istrategylabs [istrategylabs, 2010]. Based on their most recent report of user demographics, released in June 2010, our sample closely resembles the larger Facebook population of users, though the 18-24 group is underrepresented, and the 25-34 age groups is overrepresented 5.2. Our underrepresentation of younger users is in contrast to the related work focused on youths or undergraduates (e.g. [Acquisti and Gross, 2006; Boyd and Hargittai, 2010; Egelman *et al.*, 2011; Madejski *et al.*, 2012; Stutzman and Kramer-Duffield, 2010]). We restricted participation to users in the United States, and our sample represents users from several states including New York (25% of 260), Alabama (13%),

[3] www.researchmatch.org/about/

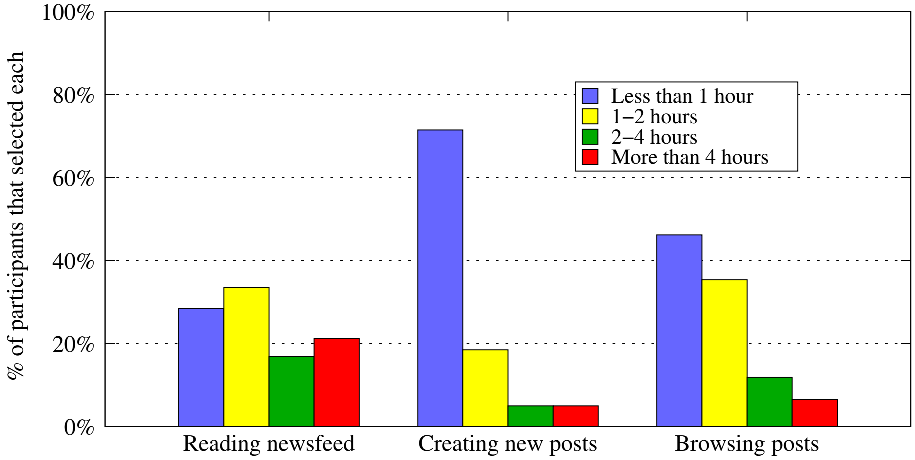


Figure 5.1: Responses to, "About how much time do you spend . . . each week?"

Minnesota (10%), and California (9%). Approximately 50% of our sample had completed at least some college.

Many of the participants are active users and report using Facebook several times a day (68.8% of 260), while very few participants said they log in less than once per week (5% of 260). Most of the participants have had their Facebook account for more than two years (77.3% of 260). We also asked about the amount of time spent on specific activities: time spent reading the newsfeed, creating new posts, or browsing friend's profiles (see Figure 5.1). In general, participants spend more time consuming content than they do creating content (see Figure 5.2).

### 5.3.3 Limitations

It is possible that our method introduced bias; two people refused to participate because of the use of a Facebook application. We cannot estimate how many others chose not to participate for similar reasons. It would seem our sample is biased toward users who are unconcerned with privacy. Though our sample might not include the users most concerned with privacy, the number of privacy concerns and mitigation strategies recorded suggest that the assumption is inaccurate. For the purpose of this study, the ability to collect detailed data about users' sharing concerns in relation to their actual friend network and

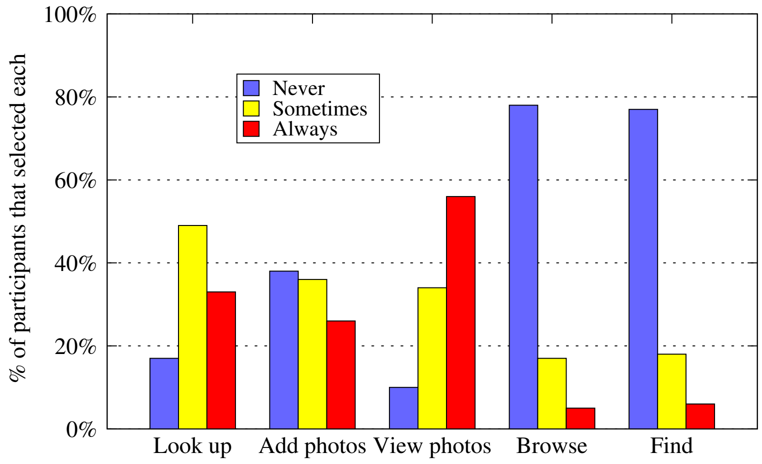


Figure 5.2: Responses to, "How frequently do you use Facebook to ...?" Answers correspond to: look up information about a friend, upload photos, view friends' photos, browse profiles of people you don't know, and find new friends.

shared posts outweighs the amount of bias that may have been introduced by our method.

## 5.4 Results

We collected data from February to April 2011, to examine Facebook users' privacy concerns and privacy-preserving strategies. Overall, we observed that the vast majority of users successfully navigate the privacy settings interface in order to prevent strangers from viewing their posted content. However, users are less prepared to deal with threats arising from the intermingling of "friends" who are associated with differing contexts (e.g., work, family, etc.) and currently use a variety of ad-hoc approaches that are unlikely to completely address their concerns. In this section we first present our results in terms of participants' strategies to prevent strangers from viewing their content, concerns about disclosures to strangers, and specific examples of content that participants would not want strangers to view. Next, we examine what we call "the insider threat," which involves inappropriately sharing content with members of the friend network. In this context, we present the compositions of participants' friend networks, their concerns with regard to sharing content with specific subgroups of their friend networks, and the strategies they employ to address their concerns.

## 5.4.1 Stranger Danger

### 5.4.1.1 Strategies to Block Strangers

We used the survey application to check the amount of profile information that was viewable to all Facebook users (i.e., the number of users not using any access control settings). The application checked the visibility of each participant's wall (e.g., status updates and comments), the people in her friend network, and her photos. We found that across the sample:

- 14.2% had a public wall (e.g., status updates and comments).
- 6.5% had at least one public photo album.
- 53.8% had the list of people in their friend network public.

Less than half of our participants (45.4% of 260) had no information accessible to the general Facebook network. The application was not able to measure whether access was further restricted beyond complete strangers (e.g., whether certain subsets of friends were also prohibited from viewing certain content). Additionally, since Facebook's default privacy settings have changed over time, we cannot definitively say whether participants' had actively blocked strangers from accessing this content or if it could be partially attributed to changes in default settings. However, these numbers do indicate that the vast majority of participants (94.6% of 260) have either photos or other content blocked from strangers.

### 5.4.1.2 Concerns with Broad Scenarios

We measured general privacy concerns using ten scenarios about unwanted audiences and asked participants to indicate on a 5-point Likert scale their level of concern that each could happen as a result of using Facebook (the markers were "unconcerned," "neutral," and "concerned"). We measured participants' level of concern that each *could* happen, as opposed to prior work that used similar scenarios to measure users' belief that the scenarios were already taking place [Young and Quan-Haase, 2009]. The set of concerns that involved profile access by strangers (i.e., people who are not members of the friend networks) are depicted in Table 5.3. The table also presents the percentage of participants who reported being concerned with each scenario, as well as the median ranking from the Likert scale. Twenty-eight participants (10.8% of 260) reported being unconcerned with any of the scenarios involving strangers (85.7% of those participants had a private profile — neither photos nor walls were accessible to strangers). However, we did not observe a statistically significant correlation between public profile access and concerns over strangers; we attribute this to the fact that so few profiles were accessible to strangers. On average, participants were concerned with 4.4 scenarios ($\sigma = 2.84$).

#### 5.4.1.3 Specific Concerns

Beyond their stated concern levels for the aforementioned scenarios, we showed participants ten random pieces of content that they had previously posted to their profiles. This may have included comments, photos, and status updates. For each piece of content, we asked participants to rate how concerned they would be if a stranger were to view it. As before, answers were reported on a 5-point Likert scale that ranged from "concerned" to "unconcerned," with "indifferent" as the neutral option.

Upon performing a Pearson correlation between participants' mean levels of concern averaged over the ten posts and whether they had a public wall, photos, or both, we observed a statistically significant positive correlation ($r = 0.15$, $p < 0.016$). One interpretation of this is that participants who knew that strangers could not access their profiles were therefore less concerned about the likelihood of it happening. The corollary to this is that participants who did not use privacy controls were more concerned about the threat of strangers.

We cannot say whether the forty participants' profiles (15.4% of 260) that were accessible to strangers were configured that way intentionally, and therefore these participants were significantly more wary of stranger dangers ($U = 3,357.5$, $p < 0.017$); or whether it was because these participants were more concerned because they were aware that they have never personally modified their privacy settings.

| **Scenario** | **Concerned** | **M** |
|---|---:|---:|
| 1. Thieves using Facebook to track, monitor, locate, and identify you as a potential victim. | 68.8% | 4 |
| 2. Your employer seeing an inappropriate photo or comment on your profile. | 62.7% | 4 |
| 3. Your employer using your profile to assess your suitability for the company. | 55.0% | 4 |
| 4. Sexual predators using Facebook to track, monitor, locate, and identify you as a potential victim. | 51.9% | 4 |
| 5. Your employer using Facebook to monitor your conduct while you're at work. | 46.2% | 3 |
| 6. Your employer using Facebook to monitor your conduct while you're away from work. | 44.6% | 3 |
| 7. A stranger will see an inappropriate photo or comment on your profile. | 40.8% | 3 |
| 8. Political parties using Facebook to target you through the use of ads and data mining. | 30.4% | 3 |
| 9. Your university using Facebook to identify you as a university code violator. | 20.0% | 1 |
| 10. Law enforcement using Facebook to track drug use and other illegal activities. | 17.3% | 1 |

Table 5.3: We asked participants to rate their level of concern that each scenario could occur on a 5-point Likert scale from "unconcerned" to "concerned," with "indifferent" as the neutral option. The second column reports the percentage of the sample that was concerned about each scenario, while the third represents the median rating for each question.

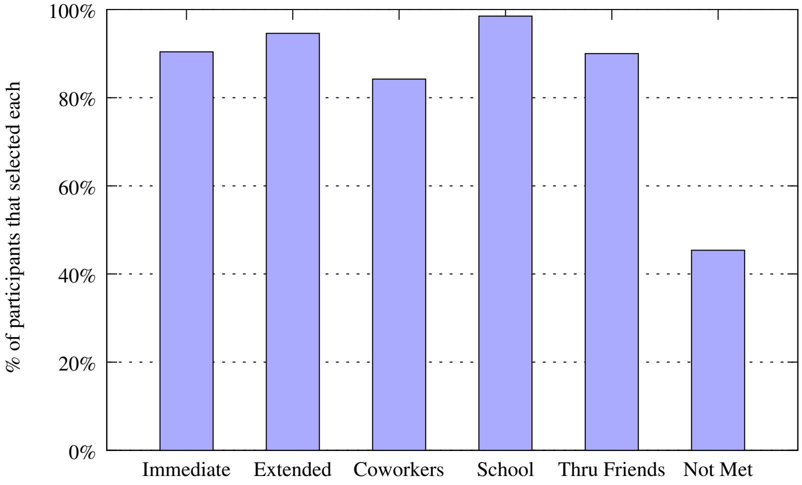


Figure 5.3: Responses to, "Which of these groups are you friends with on Facebook?"

However, our results do suggest that the vast majority of participants are aware that their profiles are relatively well protected from strangers.

### 5.4.2 The Insider Threat

#### 5.4.2.1 Friend Network Composition

Our sample's average friend network size was 357 friends (median = 291, range = [22, 3280], $\sigma = 319.5$). This is much larger than the statistic Facebook reports: the average user has 130 friends [Facebook, 2012]. It is likely that this discrepancy is indicative of a very long tail; Facebook reports the average network size for all users — including users who add a handful of friends and never access their accounts again — whereas we limited our sample to only active users. Our numbers are consistent with other academic studies of active Facebook users. For instance, Kelley et al. reported a median of 222 friends [Kelley *et al.*, 2011], Young and Quan-Haase reported an average of 401 friends, and Stutzman and Kramer-Duffield reported an average of over 400 friends [Stutzman and Kramer-Duffield, 2010].

To get a general sense of friend network composition, we asked participants, "Which of these groups are you friends with on Facebook?" and instructed them to select all options that applied to their friend networks. The choices for this question are presented in the first column of Table 5.1. As shown in Figure 5.3, a minority of participants selected people that they had not met in person, whereas the other five groups were each selected by over 80% of our participants.

For a broad overview, we asked each participant, "What percentage of your friend network do you trust with access to your profile and shared information?" The choices were presented in increments of ten from 0-100%. On average, participants claimed to trust 75.4% ($\sigma = 26.3$, median = 90%) of their friend networks. Of the 55 people who answered 50% or less, only 35 (63.6% of 55) of them claimed to have modified their privacy settings so that some of their friends have limited access to their profiles.

We validated participants' perceptions about their friend network composition by asking them to categorize a random sampling of their friends. The survey randomly selected nine people from each participant's friend network and asked questions about each selected friend. For example, if *Alice Smith* was selected, the participant was shown Alice's profile picture and was asked, "What is your relationship to *Alice Smith*?" We asked each participant to select one category from those listed in the 'Friends' column of Table 5.1.

A plurality of the friends were reported to be known from school (42.6% of 2,340). The remaining groups were also chosen at least once, though immediate family members and *not sure* were chosen least frequently.[4] Based on the categorization of the nine friends, we can estimate the average composition of participants' friend networks based on the frequency that each group was selected (see the 'Frequency' column of Table 5.4). We hypothesized that while labeling individual friends participants might discover their friend networks contained more groups than they remembered. However, we found that this was not the case; none of the nine randomly selected friends was a member of a group that a participant had not already selected in the first part of the survey. Thus, participants are by and large aware of the composition of their friend networks.

For each of the nine friends we asked participants to categorize, we also asked how they felt about that friend viewing all the content they had uploaded to Facebook. Participants

[4] This says more about the *not sure* category, as immediate family members were likely selected infrequently because each participant was likely to have only a limited number of immediate family members in real life, relatively speaking.

| | Frequency | Participants | Comfort |
|---|---:|---:|---:|
| School | 42.6% | 88% | 97.0% |
| Socialize with | 15.4% | 57.3% | 98.9% |
| Friend of a friend | 12.4% | 62.7% | 97.0% |
| Coworker | 11.1% | 45% | 96.9% |
| Extended family | 9.4% | 48.5% | 95.4% |
| Have not met | 5.3% | 20% | 95.2% |
| Immediate family | 2.1% | 14.2% | 98.0% |
| Not sure | 1.7% | 13% | 75.0% |

Table 5.4: Summary of responses about individual friends. 'Frequency' shows the number of times each group was selected across the sampling of 2,340 friends (i.e., nine friends for each of 260 participants). 'Participants' shows the percentage of participants represented in the frequency column. 'Comfort' shows the percentage of selected friends that the participant is unopposed to sharing with in that group.

responded using a 5-point Likert scale that ranged from "uncomfortable" to "comfortable," with "indifferent" as the neutral option. We provided participants with a text box to optionally explain their responses when they selected uncomfortable or slightly uncomfortable. We define "unopposed to sharing" throughout the rest of this paper to mean participants who answered with either "comfortable," "slightly comfortable," or "indifferent;" we consider those who answered with either "uncomfortable" or "slightly uncomfortable" as being opposed to sharing. The majority of participants were unopposed to sharing with all nine of the selected friends (79.2% of 260). Participants indicated that they would be opposed to sharing at least some of their profile with 3.3% of the 2,340 selected friends (this number corresponds to 54 unique participants). When a participant indicated they were opposed to sharing with a specific friend, we asked a follow-up question to prompt an explanation. We present and discuss this data in Section 5.5.

We asked participants to rate their levels of concern that two additional scenarios may happen, similar to those presented in Section 5.4.1.2. While the first ten scenarios were centered around strangers — people unlikely to appear in participants' friend networks — the additional scenarios focused on concerns with sharing inappropriate content with known

| **Scenario** | **Concerned** | **M** |
|---|---|---|
| 1. A coworker seeing an inappropriate photo or comment on your profile. | 55.0% | 4 |
| 2. A family member will see an inappropriate photo or comment on my profile. | 46.5% | 3 |

Table 5.5: We asked participants to rate their level of concern that each scenario could occur on a 5-point Likert scale from “unconcerned” to “concerned,” with “indifferent” as the neutral option. The second column reports the percentage of the sample that was concerned about each scenario, while the third represents the median rating for each question.

recipients: family members and coworkers. In Section 5.4.1.2, we asked about “employers,” whereas here we discuss coworkers. We intended for the distinction between coworkers and employers to be that the latter are in management positions (i.e., have the ability to hire and fire). We cannot say with certainty that this distinction was apparent to all participants. However, McNemar's test between participants' concern levels between when a coworker and an employer see “an inappropriate photo or comment on your profile” yielded statistically significant differences ($\chi^2 = 9.50$, $p < 0.002$). This indicates that participants viewed these two groups differently.

We observed that participants claimed to be significantly more concerned with the prospect of coworkers viewing content than family members ($\chi^2 = 5.80$, $p < 0.016$). We performed Phi correlations to examine whether participants' concerns over sharing content with these two groups were correlated with having members of these groups included in their friend networks, but found no significant correlations when examining both coworkers and family members. If there is a correlation, it is too small for us to observe among our relatively large sample. In either case, since roughly half of the participants indicated they would be concerned by these scenarios (Table 5.5), they choose to mitigate them in ways beyond preventing family members and coworkers from being included in friend networks.

#### 5.4.2.2 Concerns over Specific Content

Access control decisions are typically phrased in terms of who can access a resource, and what the resource is. For this reason, we also asked questions about sharing preferences based on specific posts. As we explained in Section 5.4.1.3, we randomly selected ten posts from each participant's profile and asked questions to measure how they perceived the sensitivity of the content. For each post, we asked eight questions of the form, "How would you feel if *group* saw this?" The groups are listed in the last column of Table 5.1. Again, participants responded using a 5-point Likert scale from "uncomfortable" to "comfortable," with "indifferent" as the neutral option. We asked participants for an optional explanation when they selected uncomfortable or slightly uncomfortable.

The groups with whom participants were least comfortable were strangers, people they had not met in person, coworkers, and those who were a friend of a friend, respectively. Table 5.6 depicts participants' levels of comfort sharing their ten posts with the various groups. The first column of the table depicts the total percentage of posts (of 2,600) that participants reported that they were unopposed to sharing with each group. However, these numbers by themselves do not give an accurate representation of the number of participants who have posted sensitive content. Thus, the second column shows this information on a per-participant basis. Specifically, this shows the percentage of participants (of 260) who were unopposed to sharing all ten randomly selected posts with each of the eight groups.

We compared the number of participants who were opposed vs. unopposed to sharing with each of the eight groups using McNemar's test across each pair of groups, and then applied the Holm-Bonferroni correction. Based on the results of these tests, we were able to partition the eight groups based on significant differences with regard to participants' comfort sharing all ten posts. This hierarchy can be seen in Figure 5.4. As expected, participants were significantly less comfortable sharing with complete strangers than any other group. Because we found no significant differences among our sample between sharing with immediate family, extended family, people from school, and people with whom participants socialize, as well as no significant differences between a coworker and a friend of a friend, we were able to consolidate our hierarchy into four discrete levels of trust.

We took this analysis a step further in order to detect the frequency with which partici-

| | Posts | Participants |
|---|---|---|
| Immediate family | 99.0% | 91.2% |
| Socialize with | 98.9% | 96.5% |
| Extended family | 98.9% | 91.2% |
| School | 98.8% | 95.0% |
| Friend of a friend | 97.2% | 90.0% |
| Coworker | 97.0% | 83.9% |
| Have not met | 91.6% | 72.3% |
| Stranger | 84.4% | 55.4% |

Table 5.6: Responses to questions about individual posts. Columns depict percentage of posts (of 2,600) participants were unopposed to sharing with the given groups, as well as the percentage of participants (of 260) who were unopposed to sharing all ten posts with the given groups.

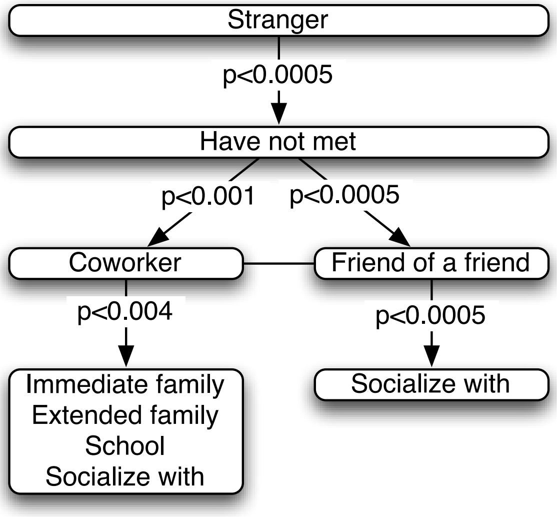


Figure 5.4: The hierarchy of participants' comfort for sharing all ten randomly selected posts.

pants' concerns were being realized: we examined whether participants who were concerned about posts being viewed by specific groups also included members of those groups in their friend networks. For instance, if a participant indicated that it would be inappropriate to share some of the ten aforementioned posts with coworkers, we also examined whether that participant had also categorized one of the nine aforementioned friends as being coworkers. We performed this analysis for coworkers, immediate family, and extended family. We found that:

- 14 of 117 (12.0%) participants whose nine friends included one *coworker* were uncomfortable sharing at least one of the ten content items with coworkers.
- 1 of 37 (2.7%) participants whose nine friends included one *immediate family member* were uncomfortable sharing at least one of the ten content items with immediate family members.
- 9 of 125 (7.2%) participants whose nine friends included one *extended family member* were uncomfortable sharing at least one of the ten content items with extended family members.

However, we observed no statistically significant correlations between the inclusion of specific groups in the friend list and discomfort with sharing specific content with members of those groups. Thus, while our data document that users are sharing content inappropriately, this occurs relatively rarely. At the same time, we have quantitatively showed that the friend network consists of varying subgroups with whom participants are not always comfortable sharing all of their their content. Thus, based on users' inability to solely rely on the friend network for boundary regulation, while at the same time realizing relatively few instances of inappropriately shared content, it is likely that they are minimizing their exposure through other means.

### 5.4.3 Strategies for Mitigating Concerns

We know from the results of qualitative surveys and anecdotal experience that SNS users employ a multitude of techniques for managing their interpersonal privacy in addition to using the actual privacy settings: custom lists, culling their friend network, deleting posts, untagging posts, asking friends to delete posts, and maintaining more than one account. We can build a better understanding of SNS users' privacy management needs by collecting data on the instances where these privacy preserving behaviors are used. In our survey, we asked participants whether they had employed each technique, and if so, to provide an explanation.

#### 5.4.3.1 Custom Lists

Custom lists allow users to subdivide their friend networks, and can be used to configure fine-grained privacy settings — to grant additional access or to block access. To examine how lists are used, we accessed each participant's set of custom lists using the API, and for each list asked, "Why did you create the custom list *LIST-NAME*, and what do you use it for?" Our sample yielded 555 custom lists: most users had created at least one (52.3%, 95% CI [46.3, 58.3]). From the explanations provided, three themes emerged: 100 lists were for privacy (18% of 555 lists), 372 were for use with other features (67%), and 83 were for created for reasons that participants can no longer remember (15%).

Approximately a quarter of the participants had created a custom list for privacy reasons (23.8% of 260). We further categorized the privacy lists based on whether the list was created to include or exclude specific friends. Exclusive lists were created to prevent access by that group (75% of the 100 privacy lists). Inclusive lists were created to give additional access to that group(17% of the 100 privacy lists). In some cases the intended use of the list was unclear (8% of the 100 privacy lists). Most of the lists were created to separate groups of friends or to differentiate contacts by closeness. A few users created custom lists for family members (24 lists) and coworkers (14 lists).

The majority of the custom lists were created for use with other Facebook features (67% of 555 lists, created by 136 participants). The descriptions provided were generic and did not explain how the lists were actually used (66.1% of 372 lists). Although these lists might not be used for privacy reasons, the names and descriptions of the lists provide insight to friend network composition and the user's many social identities like shared interests, activities, and location. Several participants created a custom list to group friends that play the same Facebook game, filter their newsfeed, and manage their chat list (14.7%, 6.7%, and 6.5% of the 372 lists, respectively).

#### 5.4.3.2 Curating the Friend Network

In an SNS like Facebook, where a friend relationship is reciprocal and friends are granted additional access to content, curating the friend network can be a privacy preserving behavior. The options for curating the friend network are to deny a friend request or delete (unfriend) a person. Nearly every participant had turned down a friend request in the past (96.2%), and so we asked them to select their reasons. The most common answer was 'didn't know the person' (selected 211 times), followed by 'knew the person but did not want them to have access to my profile' (129 times).

Most participants had unfriended at least one person (69.6% of 260). A commonly selected reason was 'because we were no longer friends in real life.' Forty-five participants had deleted a friend because they were unsure whether they knew them. Of the forty-six explanations that were provided, twenty-five participants noted they wished to stop seeing updates from the person ("because they kept posting negative or critical things without relent", "their political views were exactly opposite of mine and I did not agree with any of their posts").

#### 5.4.3.3 Control via Deletion

SNS users can also manage access to the data associated with her profile by deleting or 'untagging' posts. We posed a set of questions related to untagging and deleting posts to measure how often they are used in a privacy preserving manner. Each question was posed in the format: *have you ever. . . ? If yes, why?* We then asked participants to check all the reasons that applied from the options: you didn't want anyone to see it, you didn't want a specific person to see it, you didn't like it, or other. We also provided a text box for the participant to describe the circumstances.

We asked participants, "Have you ever untagged yourself in a photo that was posted by a friend?" Over half (58.5%of 260) the participants answered yes, and the participants provided sixty-two descriptions. Most of the descriptions related to reputation or image management: the picture looked bad (e.g., "I was making a very unattractive face"), they didn't want people to see them partying/drinking/etc., or the photo wasn't actually them (e.g., "I've been tagged in spam before").

More than half the participants responded affirmatively when we asked if they had ever deleted a photo they had uploaded to Facebook (60.8% of 260). The explanations of the circumstances varied. In some cases a photo was deleted to satisfy the request of a friend who was also in the photo (e.g., "a relative requested I remove it"). In other cases, it was to preserve privacy (e.g., "I have taken most of the pictures of my kids off because when I think about it's weird to me that random people I don't know well are looking at my kids").

About one-fifth of the participants (22.3% of 260) said that they had asked a Facebook friend to delete a photo that they were in. The most popular reason was because they 'didn't like it' (e.g., "Me getting drunk at a party, not appropriate.").

We asked participants if they had ever posted a status update or comment and later deleted it. More than half answered yes (65.4% of 260). The most popular reason was that they 'didn't like it' (e.g., "I decided it was stupid."), followed by 'didn't want anyone to see it.' Some of those who selected 'other' explained "I've written posts that later seem too personal," and "I changed the way I felt."

#### 5.4.3.4 Control via Per-Post Privacy Settings

In addition to providing global access control settings, Facebook also allows users to customize access control settings on a per-post basis through a drop-down menu at the time the post is made. We asked participants, "Have you ever changed the privacy settings for a single status update?" Then asked the 92 participants who answered yes to describe an instance where they had used the feature. From the explanations it was clear that only 47 participants had actually used the feature (18.1% of 260), the other 45 conflated the global privacy settings with the per-post feature.

We coded the situations described by the 47 participants who correctly used the feature based on whether they desired to exclude specific people, include specific people, or did not specify. Thirty-one answers were for the purpose of exclusivity, for example:

- "I have blocked family members and conservative friends from status updates they might view as inappropriate."

- "When I was talking about my roommates, I didn't want one of them to see it."
- "It was hidden from a person to announce a surprise party about them."

Fourteen answers were situations that specifically included some subset of friends (e.g., "posting a personal link regarding a vacation, I made it viewable only to a specific group"). The remaining descriptions were either unspecified or too ambiguous to categorize.

#### 5.4.3.5 Multiple Accounts

Although the terms of service mandate that each person have at most one account, twenty-two participants (8.5% of 260) had two or more accounts. Among these, sixteen cited reasons related to managing social spheres, either for dividing friends and family, friends and game friends, or their professional and social lives.

## 5.5 Implications

Our results contribute to an understanding of Facebook users' privacy concerns and the strategies they employ to mitigate their concerns. We found most users are concerned about strangers viewing their profiles, and that many users are also concerned with the *insider threat*—inappropriately sharing content with members of the friend network. Many users have private profiles, either by default or manually adjusting the global privacy settings, which means that strangers are unable to view posted content (e.g., status updates, photos, and comments). However, these settings do not adequately address the insider threat, and therefore these concerns likely go unmitigated.

### 5.5.1 Users are Concerned with Strangers and Many Effectively Mitigate Their Concern

We found evidence that many users are concerned with the possibility of an outsider accessing their shared content. In fact, the broad scenarios that elicited the highest concern were those that involved audiences that were not represented in most participants' friend networks (see Table 5.3 and Figure 5.3), we also found that participants were least comfortable sharing individual posts with a 'stranger' or someone they had 'not met in person' (see Table 5.4).[5]

We found that 89.2% of our participants were concerned with the outsider threat (232 of 260 indicated concern for at least one of the scenarios involving strangers). This figure represents the participants who selected 'concerned' for at least one of the general scenarios in Table 5.3. Of this set of concerned participants, we observed that 84.5% had a private profile. Which means that overall 15.5% of our participants were highly concerned with the outsider threat but they were not managing it through the available privacy controls. Put another way, 86.2% of participants (224 of 260; 95% CI [81.4, 91.1]) were either not very concerned with the stranger threat or they were concerned and were able to address their concerns.

Facebook users are increasingly opting for a 'friends only' profile. In 2010, Stutzman and Kramer-Duffield reported more than half of their participants had a 'friends only' profile [Stutzman and Kramer-Duffield, 2010]. This is a significant increase in adoption compared to Gross and Acquisti's 2005 observation that 0.06% of their sample had a 'friends only' profile. Thus, the existing privacy settings interface and the default settings may reasonably address users' privacy concerns regarding strangers.

### 5.5.2 The Insider Threat Prevails Unmitigated

We identified interpersonal sharing concerns in 37% of our sample (96 of 260; 95% CI [31.0, 43.1]), this figure represents the participants who either expressed concerns about sharing with a specific friend or expressed concerns about sharing a specific post. We have reason to believe that this figure is a lower bound since the sample represents such a small portion of most users' friend networks and shared data. Even so, we can positively say that this group

[5] The option 'not met in person' was used in two sections of the survey: reporting friend network composition and in the section on individual posts. We intended it to describe a person known online but not in the real world, however, users understood the intended meaning in the first section but did not in the posts section. Based on the explanations offered, many equated 'not met in person' with 'stranger' in the posts section.

of participants have interpersonal sharing concerns that have been realized. Furthermore, our data on the number of users who supplement the privacy controls with ad hoc strategies suggests that the privacy controls are unsuited for dealing with this threat. It is important to understand the characteristics of users' actual privacy concerns to determine which should be addressed with privacy controls and which would be better handled through other means like self-censorship or removal.

Based on our findings, we recommend additional research to understand the taxonomy of privacy concerns experienced by users. We found that users' conceptualization of the insider threat varies depending on the context of of a situation. For example, some participants described their concerns by their relationship with a person while others described them based on the content of posts. In situations where participants were uncomfortable with specific friends, some participants described their discomfort as a general distrust of the person, while others described specific types of content they would not want the person to see. Also, in some cases the participants described their concern in terms of their intended audience, while others described their concern by the people that should be excluded. Participants almost never described their concerns in terms of broad groups, though this could be an artifact of our question style. In Section 5.4.3.2, we provide sample participant responses to illustrate users' reasons for deleting content and using per-post privacy settings.

Facebook users could use custom friend lists to address the insider threat. However, based on our data and prior work, it seems this feature is largely a failure. Nearly all of our participants who had created a custom list also utilized additional privacy preserving behaviors, and many of the custom lists we recorded were not used for managing privacy concerns. Furthermore, according to prior work, although users are able to organize their friends into lists, the lists are effectively useless because the user-created lists fail to accurately capture their desired audiences for shared content [Jones and O'Neill, 2010; Kelley *et al.*, 2011]. As mentioned above, the threat model has changed such that now the problem consists of edge cases that are highly contextual. One reason custom lists do not address the problem is that they are created a priori, before the user thinks about the situations in which they will be used [Kelley *et al.*, 2011]. We note this problem is most likely present in other SNSs that rely on audience management, such as Google Plus.

Perhaps the user interface for the privacy controls could do more to support the notion that curating the friend list is an effective strategy for privacy management, assuming the user has a 'friends only' profile. Our participants' ability to manage the stranger threat suggests that a basic model of outsiders vs. insider is more effective than many of the existing privacy controls. Deleting shared content is also an effective method for mitigating privacy concerns, but it can only be done post hoc, that is, after the potential damage has been done. Based on our data, it seems that users are aware and sensitive to their privacy concerns, but they lack reliable mechanisms to address them.

### 5.5.3 Generalizability

We believe that our sample more accurately reflects the current demographics of Facebook users compared to most prior work on this topic. Despite significant changes in the demographics of Facebook users, most of the empirical research on Facebook users' privacy concerns has been limited to undergraduates or teenagers [Acquisti and Gross, 2006; Boyd and Hargittai, 2010; Egelman *et al.*, 2011; Krasnova *et al.*, 2009; Lewis *et al.*, 2008; Madejski *et al.*, 2012; Stutzman and Kramer-Duffield, 2010; Tufekci, 2008]. Notable exceptions include two studies on group co-presence [Lampinen *et al.*, 2009; Skeels and Grudin, 2009]. In the early days of Facebook, when membership was limited to university students, studying undergraduates made sense. However since 2009, Facebook has become popular with other demographics as well [Smith, 2009]. College-aged individuals now represent a minority of the Facebook population [istrategylabs, 2010].

### 5.5.4 Limitations

The numbers we present in this paper are likely not exact due to several confounding factors. First, we only asked participants about a limited number of their friends and posts. Thus, participants' concerns for these likely represent lower bounds. Additionally, we chose to only ask participants about their levels of concern, rather than their perceived likelihood of negative outcomes or whether they regretted sharing specific content. Thus, while participants may have concerns, it is unclear under what circumstances these concerns may rise to the level of altering behavior.

It is also likely that our method inherently introduced bias: we have direct evidence that at least two people refused to participate because of the use of a Facebook application. We cannot estimate how many others chose not to participate for similar reasons. Thus, it would at first seem that our sample is biased toward users who are unconcerned with privacy. While our sample might not include the users most concerned with privacy, the number of privacy concerns and mitigation strategies recorded indicate that our participants had very clear privacy concerns. It is possible that without this bias, our sample would reflect even stronger privacy concerns.

### 5.5.5 Future Work

The existing privacy controls fail to empower users to adequately manage their concerns because they mostly focus on strangers. Our study identifies several privacy preserving behaviors that users rely on to manage their interpersonal privacy. We suggest that familiarity with the strategies used and when they are employed will help researchers identify specific weaknesses to address in future privacy controls. We do not mean to imply that privacy controls must replace all privacy preserving behaviors or that the privacy controls are the only way to manage interpersonal privacy. Rather, we believe SNS users would benefit from improved controls that match their needs.

By measuring Facebook users' interpersonal privacy concerns we learned that the insider threat is a significant concern for many users, but not a concern for all. As a result, we collected detailed data about the insider threat from only a subset of our sample. Our data suggest that users' conceptualization of the insider threat is individualized and dependent on context, but they do not reveal generalizable themes. A follow-up study might take a similar approach to measuring the threat and choose to focus on participants who are concerned with the insider threat.

As we mentioned previously, we believe our estimate for the number of people who are concerned with the insider threat is a lower bound. A follow-up study could test this by using a similar methodology and asking questions about a larger sample of the participants' friend networks or a larger sample of the participants' shared content. In most cases, our sampling of the friend networks represented less than 3% of a user's friends. For the purpose of understanding the insider threat, it may be useful to devise a way to select friends of interest or shared items that are most likely to be problematic.

We found that the threat model has evolved, which suggests that our approach to the problem needs to evolve as well. In the past, we measured the usability of access control interfaces in strict terms: How many unintended parties can view the content? How many intended audiences cannot view the content? These metrics may have served us well in the past, but applying it as we move forward would be a mistake. One complication is that with the insider threat, the threat is dynamic, unlike the more static stranger threat; the appropriateness of the audience is highly contextual. Moreover, because of the complexity of this problem, we will never have an interface that provides perfect coverage in all situations, for all users. Instead, we should focus on designing fine-grained controls that work for *most* users, *most* of the time. Ideally a mechanism that achieves this goal would also effectively communicate its limitations and promote alternative privacy preserving behaviors that mitigate users' residual concerns.

## 5.6 Summary

Our survey of Facebook users' privacy concerns contributes to human-centered efforts to correct the inadequacies of existing SNS privacy settings. As we saw in Chapter 4, the inconsistencies in users' privacy settings suggest that usable access control mechanisms are needed. The number of inconsistencies we identified in the previous chapter demonstrate that the existing mechanism is unusable and a better mechanism is needed. The purpose of the survey presented here is twofold: to collect data on users' interpersonal privacy concerns to show that indifference is not the explanation for users' unwillingness to correct inconsistencies in their settings, and to collect data on users' concerns and mitigation strategies to inform the design of a better mechanism based on the users' needs.

Our survey investigated users' attitudes toward various scenarios of inadvertent sharing, measured level of comfort sharing with individual friends and shared posts, and measured how users mitigate their concerns. We found that many users who are concerned with outsiders viewing their profile effectively manage this threat by enabling 'Friends Only' access to their shared content. We also found that many users have concerns about sharing with specific friends or sharing specific types of content. These situations are of higher concern to the users and are quite individualized. SNS users desire fine-grained control, and the frequency with which they rely on ad-hoc privacy preserving behaviors suggests that the existing controls fail to meet their needs.

# Chapter 6

# The Effect of Viewership Feedback on Privacy Preserving Behaviors

There appears to be a discrepancy between the results we present in Chapter 4 and 5: first we presented our finding that users' Facebook privacy settings are incorrect and that users claim they will not fix the errors, then we presented evidence that users have privacy concerns and employ strategies to manage their interpersonal privacy concerns. We might have interpreted users' unwillingness to fix their sharing violations as an indication of a lax attitude toward privacy, however, Chapter 5 largely dispels that conclusion. We interpret these findings as evidence that users want privacy controls that would allow them to address their privacy concerns, however, adequate controls do not currently exist. And so, users are forced to adopt ad-hoc mechanisms that allow them to mitigate concerns in the ways they understand.

In Chapter 4 we speculated that someone might be unwilling to correct a sharing violation because they are unmotivated to do so, or because they do not understand how to utilize the existing privacy controls; we also called for immediate improvements to the existing Facebook privacy controls. As a follow-up, we considered ways to increase the usability of the existing privacy controls without demanding a complete redesign. Toward this goal, we identified an SNS user's ignorance to viewership — who has viewed their profile and which posts have been viewed — as a potential barrier to the user's ability to correctly manage their privacy settings. In many access control domains the policy author evaluates the correctness of a policy using access logs, this is information that Facebook users are not privy to in the current design. Presently, users only receive viewership feedback when another user comments on a post, 'Likes' a post, or mentions the content in an offline context. It is understandably difficult to construct an accurate model of your audience from this limited feedback. To be clear, we are not suggesting that Facebook users should receive a detailed log of who accessed their profile content and when. The revelation of such information would introduce unwanted privacy issues for the viewers.

In this chapter we present a study where we tested whether the correctness of a user's actual privacy settings can be increased by making slight changes to the privacy controls. As a first step, we began by testing the effect of introducing viewership feedback on an SNS users utilization of privacy preserving behaviors. Then, we tested whether users would fix incorrect photo album settings when given notification of the problem to fix and instructions for how to fix it.

## 6.1 The Study

We were motivated to add viewership feedback by the workflow of system administrators, environmental psychology [Altman, 1975], and human-centered research on location sharing systems [Lederer *et al.*, 2004; Consolvo *et al.*, 2005; Tsai *et al.*, 2009].

System administrators rely heavily on log files to evaluate the status of the network and the effectiveness of deployed policies. It would be difficult, if not impossible, for a system administrator to effectively do their job without log files. Considering SNS privacy settings are essentially an access control policy, it is understandable that without feedback a user's privacy settings would be incorrect.

Prior to the widespread use of SNSs, Altman described the process by which people manage interpersonal privacy in the offline world and introduced the term *boundary regulation* [Altman, 1975]. Altman's definition aptly captures the nature of SNS interpersonal privacy in his definition of privacy "selective control of access to the self." Altman claims that "a person adjusts self boundaries based on past experience" [Altman, 1975]. Presently, there is a limited amount of feedback available to Facebook users to portray their actual audience or which of their posts are most frequently viewed, which denies users the ability to iteratively adjust their boundaries according to reality.

In human-centered research on location sharing systems, Lederer et al. and Consolvo et al. independently found that knowing *who* requested access to the user's location is the most important factor to the user's decision to allow or reject the request [Lederer *et al.*, 2004; Consolvo *et al.*, 2005]. Consolvo et al. found that the reason *why* for initiating a request was also an important factor. While there are differences between the location sharing systems and Facebook, both systems enable the primary user to share personal information with a network of contacts, and both allow the primary user to manage an access control policy. Tsai et al. found the addition of feedback to a location sharing system encouraged users to loosen their access control policies and increased users' comfort in using the system.

In our investigation of users' privacy concerns we found that users express their privacy concerns in terms of specific people or specific types of content. This led us to test two types of viewership feedback: descriptions of the audience for shared information and recently viewed content. We generated artificial but plausible feedback for the purpose of our study — the actual viewership data is not accessible to Facebook users, or to Facebook application developers. In addition to the viewership feedback, we chose to present suggestions for privacy preserving behaviors for the participants' consideration as potential responses to the feedback. This addition was motivated by our results from Chapter 4, where we speculated that Facebook users might be unaware of the options in the existing privacy controls or that the user might not know how to use them.

## 6.2 Method

We designed our methodology under the assumption that there is likely a gap between users' desired level of privacy and their actual level of privacy on Facebook, and that they may not be aware of their options for managing the privacy of their shared content. With this in mind, we tested whether the use of privacy preserving behaviors would increase after a participant was shown feedback about the (hypothetical) viewership of their shared data.

### 6.2.1 Design

We implemented the study as a Facebook application to collect data from the user's Facebook profile, to customize and deliver feedback to the participant based on her profile data and friend network, and to administer the survey questions.[1]

Out study design tested two types of feedback to represent the hypothetical viewership of shared data: high-level descriptions of Facebook users who recently viewed the participant's profile and the presentation of individual posts and photos from the user's shared data as posts that had recently been viewed. The viewership feedback was contrived for the purpose of the study and was selected to be plausible based on the users' privacy settings and friend network. It is important to stress that the viewership data was entirely contrived — Facebook does not release the data on which posts are viewed or which users view profiles or data items.

The study application randomly selected four posts and three photos to be shown as "recently viewed." When possible, the application selected older posts from the user's shared data. Since the introduction of the *News feed*, Facebook has positioned itself as a social network for communicating real-time updates even though the full history of a user's *wall* activity is accessible via her profile. Users' attitude toward Facebook as an archive of all this data will change as they become accustomed to Timeline and the ability to search someone's profile by year, but it is unclear how long it will take for users' mental models to shift.

The generic viewership descriptions were generated based on the participant's shared content and their privacy settings. For example, if a user had a public photo album they might see:

> Someone who is *not* your Facebook friend browsed your public photos.

If a user did not have a public wall or public albums the viewership descriptions were randomly generated to include a viewer subject and viewed profile data. The subjects were chosen based on gender and current city, and the profile data was chosen from the set: wall,

[1] The research protocol was approved by the Columbia University Institutional Review Board as protocol AAAJ2451.

photos, friend list, about me, and profile. A representative list of audience descriptions:

- Someone who is *not* your Facebook friend but is in *one of your networks* browsed some of your public photos.
- A female Facebook friend browsed your friends list.
- A Facebook friend who lives in Los Angeles, California viewed your profile.
- A male Facebook friend browsed your entire wall.
- One of your Facebook friends viewed your About Me.

The Facebook application presented the experimental groups with viewership feedback and collected data on participants' reactions to the feedback and their use of privacy preserving behaviors. For the purpose of the study, we define privacy preserving behaviors (PPBs) as:

1. Changing the visibility of the Wall
2. Managing the privacy settings of a photo album
3. Creating a custom friend list
4. Deleting a friend network member
5. Deleting a post

These privacy preserving behaviors were chosen to represent privacy management features that are currently available in Facebook (#1, #2, and #3), and other strategies that users employ to manage access to their content (#4 and #5) [Stutzman and Kramer-Duffield, 2010; Madden, 2012].

After measuring the participant's reaction to the feedback, the application displayed links to describe actions they might want to take in response to the feedback. This design decision was influenced by our assumption that users might not be aware of how to use the available privacy features. The links included advice on how to manage the privacy of a user's wall, managing album privacy settings, changing the privacy setting of a single post, creating a new friend list, removing a friend network member, and deleting a single post ( see Appendix C).

We randomly assigned participants to one of three conditions: control, viewership feedback about recent audiences of shared data, and viewership feedback about recently viewed posts and photos. For the remainder of the paper we refer to the three conditions as Control, Audience, and Posts, and use the term experimental groups to refer to Audience and Posts in aggregate. We assigned participants to a group randomly, with an attempt to balance the conditions by age, gender, and publicness of profile. The control group did not receive any viewership feedback.

### 6.2.2 Procedure

Our study was comprised of several stages that participants completed over approximately twenty-four days. First, all participants completed a pre-screening questionnaire to confirm age, country of residence, and level of activity on Facebook. After a condition was assigned, all participants completed the same pre-study questionnaire. At this point, the study procedure diverged depending on group assignment. The experimental groups received two treatments of viewership feedback during the study. The first treatment took place one week after the pre-study questionnaire. The second treatment took place one week after the first. The control group did not receive viewership feedback, though the application measured the use of PPBs for the control group at the same increments as the Audience and Posts groups. Finally, the participants completed a post-study questionnaire followed by a final stage devoted to photo album privacy settings.

#### 6.2.2.1 Pre-study Questionnaire

In the pre-study questionnaire, we primarily asked questions to provide context for the participants' response to the feedback. We began by collecting additional demographic data like employment status and race, in addition to a set of questions to gauge tech-savviness. We borrowed the Likert scale on tech-savviness from Boyd and Hargittai's work on Facebook privacy settings [Boyd and Hargittai, 2010].

We also asked participants to report their interpersonal privacy concerns and organizational privacy concerns using Likert scales borrowed from Krasnova et al. [Krasnova *et al.*, 2009]. Finally, we asked users about their use of Facebook privacy settings, their confidence in their ability to use the privacy settings correctly, and their confidence in their current privacy settings.

#### 6.2.2.2 Feedback Presentation and Survey

Each feedback period began was preceded by the survey application measuring PPB use for all participants. The participants in the experimental groups were notified by email to return to the study application for the next stage of the study. Upon the participant's return, the study application presented the feedback followed by three 7-point Likert items:

1. This information is what I expected.
2. I am a little surprised by this.
3. I am comfortable with this information.

On a second screen of questions, the study application gave the participant as chance to provide an open-form response with their reaction to the feedback. Both feedback periods concluded with a presentation of links to brief instructions for actions the participant might consider based on the feedback.

- I only want Friends to see my Wall
- I want to limit access to some of my Friends
- I want to change the privacy settings of a photo album
- I want to delete a post
- I want to remove a tag
- I want to remove some people from my Friend network

This stage of the study was repeated twice for participants in the experimental groups. During this time, the Control group did not interact with the study application. If a participant from the Control group visited the study application between the pre-study and post-study questionnaire, the study application displayed an overview of their progress in the study design and a reminder that they would be notified when it was time for the next stage in the study.

#### 6.2.2.3 Post-study Questionnaire

All participants completed a post-study questionnaire approximately seven days after the completion of the second round of treatment. The post-study questionnaire repeated a subset of questions from the pre-study questionnaire such as the Likert scales for measuring privacy concerns, and the questions about Facebook privacy settings. We added a set of questions to collect reactions to participating in the study.

#### 6.2.2.4 Album Privacy Settings

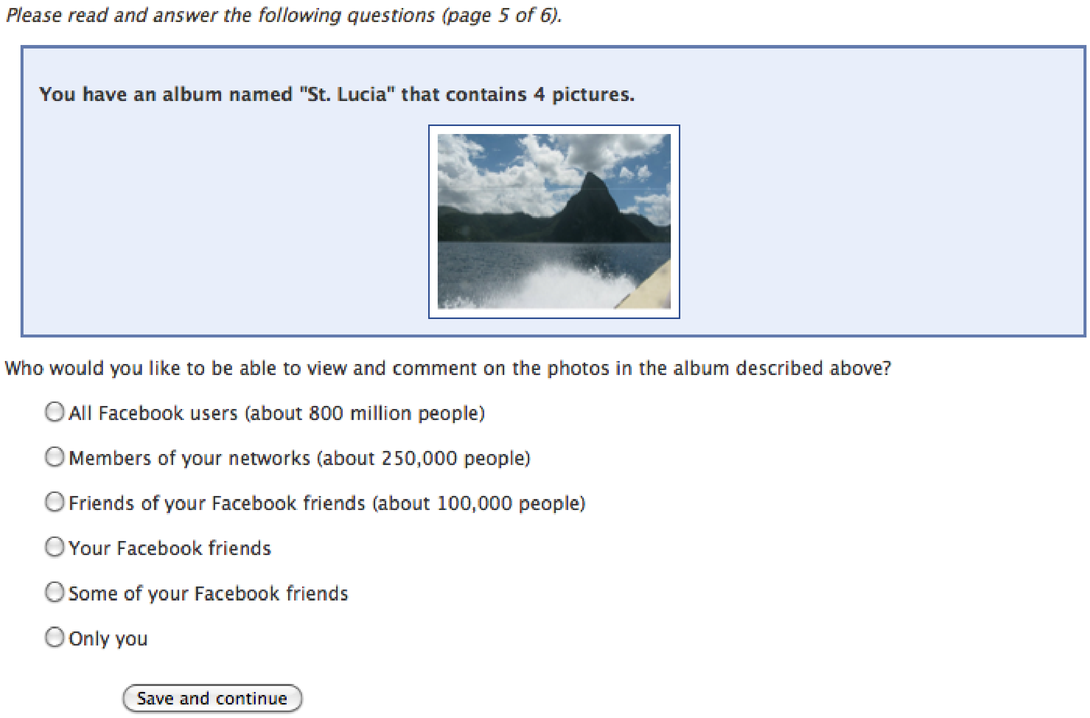
Participants who completed the study and still had photo albums accessible to people outside their friend network (i.e.,'Everyone', 'Network Members', or 'Friends of Friends') answered a final set of questions to confirm that their album privacy settings were consistent with their sharing intentions. The study application asked the participant, "Who would you like to be able to view and comment on the photos in the album described above?" for up to ten randomly selected photo albums accessible to an audience greater than their friend network.

Our study application presented metadata about each photo album and asked, "Who would you like to be able to view and comment on the photos in the album described above?" (Figure 6.1). We chose to display the album metadata to help the participant establish some context. The displayed metadata included: the cover photo of the album, the name of the album, and the number of photos in the album.

The participant selected one of the following answers to indicate their sharing intention for the photo album:

1. All Facebook users (about 900 million people)

2. Members of your networks (about 250,000 people)



*Please read and answer the following questions (page 5 of 6).*

**You have an album named "St. Lucia" that contains 4 pictures.**

Who would you like to be able to view and comment on the photos in the album described above?

○ All Facebook users (about 800 million people)

○ Members of your networks (about 250,000 people)

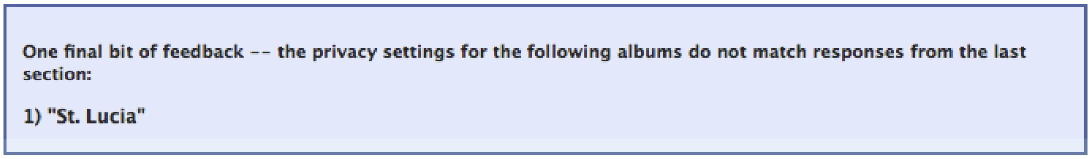
○ Friends of your Facebook friends (about 100,000 people)

○ Your Facebook friends

○ Some of your Facebook friends

○ Only you

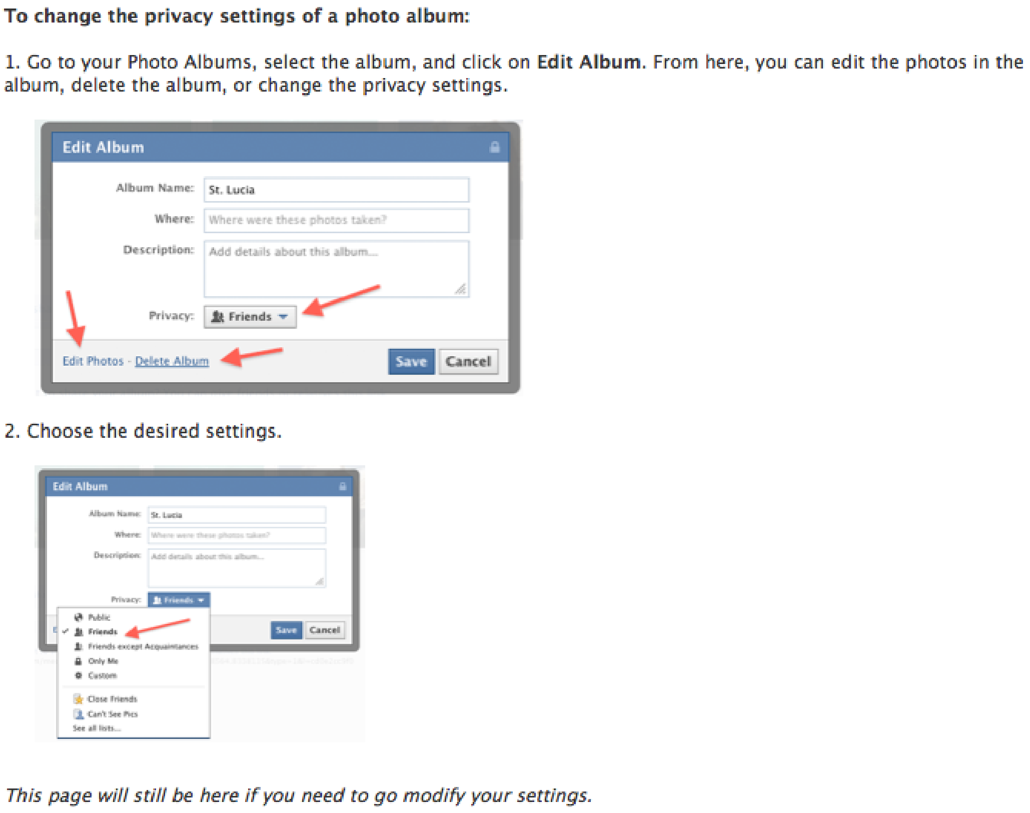
Figure 6.1: We asked each participant to report their sharing intentions for up to ten photo albums. We presented metadata for each album to provide context (e.g., album name, number of photos, and the cover photo). The choices listed are the choices provided by the privacy controls, supplemented with an estimate of audience size.



**One final bit of feedback -- the privacy settings for the following albums do not match responses from the last section:**

**1) "St. Lucia"**

Figure 6.2: If the participant had any albums with inconsistent privacy settings, we presented a list of album names and mentioned the inconsistent settings.



**To change the privacy settings of a photo album:**

1. Go to your Photo Albums, select the album, and click on **Edit Album**. From here, you can edit the photos in the album, delete the album, or change the privacy settings.

Edit Album

Album Name: St. Lucia

Where: Where were these photos taken?

Description: Add details about this album...

Privacy: Friends

2. Choose the desired settings.

*This page will still be here if you need to go modify your settings.*

Figure 6.3: On the same screen as the list of albums, we presented brief instructions for how to change the privacy settings of a photo album.

3. Friends of your Facebook friends (about 100,000 people)

4. Some of your Facebook friends

5. Only you

The sharing choices match those available in the Facebook privacy settings, and were also used by Liu et al. in their study of Facebook privacy settings [Liu *et al.*, 2011]. The additional descriptions that estimate audience size are similar to the method used by Caine et al. in their study of Facebook sharing preferences [Caine *et al.*, 2011].

After the participant this set of questions, the study application intervened with an additional stage of feedback when their privacy settings did not match their sharing intentions. In this case, the final screen of the survey would state, "One final bit of feedback — the privacy settings for the following albums do not match your responses from the last section." This message was followed by the album names and cover photos of the albums with a mismatch, and step-by-step instructions for changing the privacy settings of a photo album. The instructions were the same as the instructions for managing album privacy settings that were provided to the experimental groups after each round of feedback (see Figure 6.2 and Figure 6.3).

Data collection ended with one final observation of the participant's album privacy settings to measure whether this last interaction encouraged a participant to change their privacy settings.

### 6.2.3 Participants

We recruited participants using ResearchMatch[2], a service funded by National Institute of Health for the purpose of pairing researchers with potential participants. Our recruitment material stated we were, "looking for people to participate in a research study on how Facebook users balance their privacy concerns with the desire to share personal information online." We prescreened participants based on age, gender, and their level of activity on Facebook. Participants consented to participating in the research after reading a consent

[2] www.researchmatch.org/about/

form and a description of the research. The informed consent process took place before the participant installed the study's Facebook application. We offered participants a $10 gift certificate in exchange for their time. We also held a drawing for four $50 gift certificates among the people who completed the study to encourage participants to complete the entire study.

We advertised our study to 3,019 people via ResearchMatch and contacted 432 people with additional information about the study purpose and procedure. 190 people installed the Facebook application and provided pre-screening information. We selected 127 people to participate in the study and 107 people completed the study. Before we began recruiting we decided to only include participants who reported to use Facebook at least weekly, and have more than ten photos and more than 35 shared posts on their profile. The 63 people who were not selected as participants failed to meet one of both of these requirements.

### 6.2.4 Demographics

A total of 107 people completed the study: 33 people in the control condition, 39 in the audience condition, and 35 in the posts condition. We collected age, gender, highest level of education, employment status, and race to describe the demographics of our sample. The mean age for our sample is 34 years (min = 19, max = 62). The age distribution for the three groups is quite similar, with the most noticeable variation being the median age for Posts. Our sample is 55% female and 45% male. The proportion of female and male participants is maintained in the experimental conditions, but the control group has a slight overrepresentation of females with 63% (Figure 6.4 shows gender by condition). More than half of the sample self-reported their race as White (74.2%), approximately 15.0% selected Asian, Pacific Islander, and 6.5% selected African-American. In terms of the highest level of education completed, all of the participants were high school graduates, 46.7% were college graduates, and 38.3% had completed a post graduate degree. Approximately three-quarters of the participants were employed for wages at the time of participation (74.8%), and 15.9% of the participants were students.

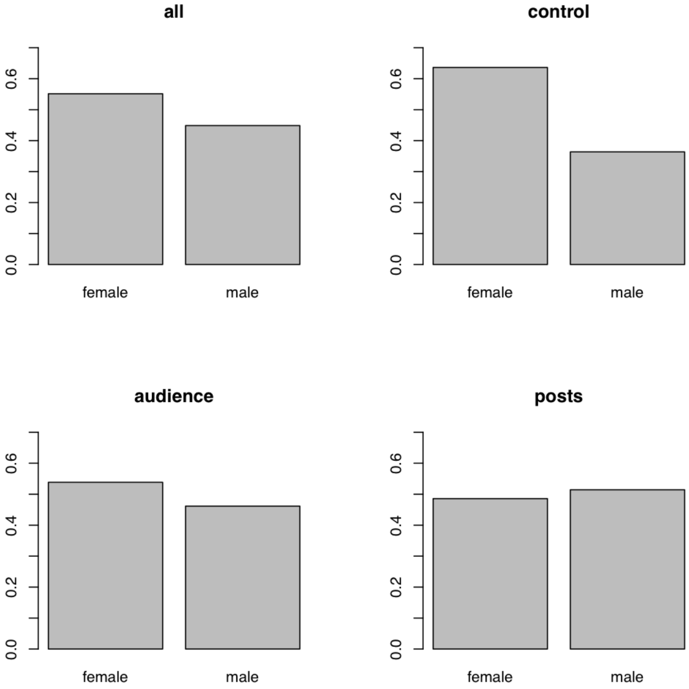


Figure 6.4: Percentage of male and female participants by condition.

### 6.2.5 Limitations

Our decision to conduct the research using a Facebook application had some pros and cons. The benefits of using a Facebook application include the ability to collect data on users' actual privacy settings, and to ask questions using real profile data. A drawback is the potential bias that it introduces. Two people who received the recruiting message contacted us to say they would not participate because of the requirement to install a Facebook application. Two others contacted us for clarification regarding the requested permissions. We suspect the presentation of the permissions in the Facebook installation dialog is a root cause of such inquiries (see King et al. on users' misunderstandings about Facebook applications and their capabilities [King *et al.*, 2011]).

Despite our efforts to balance the conditions by age, gender, and profile visibility, the final demographics for each condition are slightly imbalanced due to attrition. Our sample appears to overrepresent the White demographic for race, however, PEW Internet reported 78% of Facebook users were White in 2011 [Hampton *et al.*, 2011]. Our sample represents people with higher levels of education than the PEW sample (20% with a college degree, and 15% with a graduate degree).

## 6.3 Results

In this section we summarize the data collected and present our analysis of the data. We begin by presenting the data related to participants' use of the privacy preserving behaviors described in Section 6.2: wall visibility, photo album privacy settings, custom friend lists, deleting posts and tags, and removing people from the friend network. Then we present responses to the pre-study and post-study questionnaires to contextualize our analysis. We discuss the results and their implications in Section 6.4.

### 6.3.1 Privacy Preserving Behaviors

We measured the participants' use of the PPBs at three points in time. The first PPB measure represents the changes that took place during the seven days after the pre-study questionnaire and before the first feedback treatment. The second PPB measure repre-

| **Observation** | **Control** | **Audience** | **Posts** |
|---:|:---:|:---:|:---:|
| PPB-1 | 3% | 2.6% | 0% |
| PPB-2 | 0% | 2.6% | 0% |
| PPB-3 | 3% | 2.6% | 2.9% |

Table 6.1: The percentage of participants who made changes to the visibility of their wall during the period of observation indicated in the first column.

sents the changes that took place after the first feedback treatment and before the second. Similarly, the third PPB measure is after the second feedback treatment and before the post-study questionnaire. Throughout this section we will refer to these measures as PPB-1, PPB-2, and PPB-3.

#### 6.3.1.1 Wall Visibility

The study application measured wall visibility as public if any posts were accessible to 'Everyone.' Very few participants made changes to the visibility of their wall during the study (see Table 6.1). Several of the observed changes were participants who changed their wall from private to public (Control = 1, Audience = 2, and Posts = 1). Two of the participants changed their settings to make their wall private.

#### 6.3.1.2 Album Privacy Settings

Only a few participants in each condition changed their album privacy settings during the course of the study. We did not apply any statistical tests to this data because of the low number of participants who made changes (see Table 6.2).

#### 6.3.1.3 Custom Friend Lists

The API permits access to the user's set of custom friend lists, but it is not possible to know how a user uses the custom lists via the API. Previously, all custom friend lists were created by the user. In 2011, Facebook introduced "smart lists" that are auto-created and populated using profile information like employment, schools attended, and familial

| Observation | Control | Audience | Posts |
|---|---|---|---|
| PPB-1 | 3% | 2.6% | 5.7% |
| PPB-2 | 0% | 5.1% | 5.7% |
| PPB-3 | 6% | 5.1% | 0% |

Table 6.2: The percentage of participants who changed the privacy settings of at least one photo album during the period of observation indicated in the first column. The percentages reported represent unique participants for each observation period — no participant managed their album privacy settings more than once during the stages of the study presented in this table.

| Observation | Control | Audience | Posts |
|---|---|---|---|
| PPB-1 | 15.2% | 17.9% | 37.1% |
| PPB-2 | 27.3% | 17.9% | 22.9% |
| PPB-3 | 24.2% | 23.1% | 20.0% |

Table 6.3: The percentage of participants who deleted or untagged a post during the period of observation indicated in the first column.

relationships. The set of custom friend lists includes metadata that indicates the list type: work, school, user created. None of the participants in any of the conditions created a new custom friend list during the study. One participant in the Posts condition deleted a user-created custom friend list.

#### 6.3.1.4 Untagging and Deleting Posts

In each condition, approximately a quarter of the participants deleted a post or untagged themselves in a post that was previously on their wall (see Table 6.2). The number of participants that made changes after completing the pre-study questionnaire is roughly the same as the number of participants that made changes during the feedback periods.

| Observation | Control | Audience | Posts |
|---:|---|---|---|
| PPB-1 | 57.8% | 53.8% | 51.4% |
| PPB-2 | 57.8% | 59.0% | 34.2% |
| PPB-3 | 66.7% | 51.2% | 51.4% |

Table 6.4: The percentage of participants who removed a friend from their friend network during the period of observation indicated in the first column.

#### 6.3.1.5 Changes to Friend Network

Throughout the study the study application detected that many participants made changes to the membership of their friend network. It is possible that the numbers reported in Table 6.4 are misleading because of the symmetric nature of Facebook friendships. The study application was unable to distinguish between two actions: Alice deleting Bob from her friend network, or Bob deleting Alice from his friend network. If Alice is a participant in the study, both actions would be recorded as a friend deletion, but only one is initiated by the participant.

### 6.3.2 Self-Reported Reactions to Feedback

In Table 6.5 we summarize the participants' reactions to the feedback by group and by period. For analysis, we reduced the responses to dichotomous data for each question: unexpected and expected, not surprised and surprised, and uncomfortable and comfortable. We compared the responses between the experimental groups to compare participants' reactions to the different feedback styles.

For the question about expectations, we classified 'neutral' responses as 'expected.' We computed a two-tailed Fisher's exact test to compare the experimental groups' responses to the first round of feedback. The two-tailed p-value is 0.0143. Thus, the Posts group found the feedback more 'unexpected' in a statistically significant manner.

For the question about whether the feedback was surprising, we classified 'neutral' responses as 'unsurprised.' A two-tailed Fisher's exact test produced a p-value of 0.8181. Overall, the posts group more frequently answered 'surprised' but the difference between

| | **Likert Item** | **1st Qu.** | **Median** | **3rd Qu.** |
|---|---|---|---|---|
| Audience 1 | This information is what I expected. | 3 | 4 | 4 |
| | I am a little surprised by this. | 2 | 4 | 4 |
| | I am comfortable with this information. | 3 | 4 | 4 |
| Audience 2 | This information is what I expected. | 3 | 4 | 4 |
| | I am a little surprised by this. | 2 | 2 | 4 |
| | I am comfortable with this information. | 4 | 4 | 4 |
| Posts 1 | This information is what I expected. | 2 | 3 | 4 |
| | I am a little surprised by this. | 2 | 4 | 4 |
| | I am comfortable with this information. | 3 | 4 | 4 |
| Posts 2 | This information is what I expected. | 2 | 3 | 4 |
| | I am a little surprised by this. | 2 | 3 | 4 |
| | I am comfortable with this information. | 4 | 4 | 5 |

Table 6.5: Summary of participants' reactions to viewership feedback. The left most column describes the condition and the feedback period (period 1 or 2). Each Likert item was presented with a 5-point scale with the anchors strongly disagree, neutral, and strongly agree.

the two groups is not significant. For the question about whether the participant was comfortable with the feedback, we classified 'neutral' responses as 'comfortable' and found that in both conditions about 90% of participants were comfortable with the feedback. We did not analyze this question further.

### 6.3.3 Responses to Questions about Individual Photo Albums

In Table 6.6, we present the results from the final set of questions about the privacy settings of individual photo albums. We asked participants about their sharing intentions for up to ten photo albums that were still available to people outside their friend network. We found that approximately 66% of the 74 participants in the experimental groups had some publicly accessible albums even after exposure to viewership feedback. We needed to determine if this was intentional. We found that 96% of the participants in the experimental groups who had publicly available photo albums were sharing their albums more widely than intended.

From the data presented in Table 6.6, the viewership feedback increased the likelihood that a participant would make changes in this last stage of the study. We tested whether the experimental groups were more likely to respond by adjusting their privacy settings. The difference in the number of participants that made changes between the experimental groups and the control group was significant (two-tailed Fisher's exact test $p = 0.0267$).

We conducted additional tests to determine if there was a difference between the types of feedback. The Posts participants were significantly more likely to change their privacy settings than the Control participants (two-tailed Fisher's exact test $p = 0.0109$). The Audience participants made more changes than the Control participants but the difference is not significant (two-tailed Fisher's exact test $p = 0.1273$).

### 6.3.4 Facebook Privacy Settings

At the beginning and at the end of the study, we asked participants to remark on their confidence in their ability to use Facebook's privacy settings and their confidence in their current privacy settings. We report the results by group in Table 6.7. At the onset of the study most users were confident in their ability to use the privacy controls and confident that their privacy settings matched their sharing intentions. After participating in the

| | **Control** | **Audience** | **Posts** |
|---|---|---|---|
| Respondents (of sample) | 60.1% | 64.1% | 68.6% |
| Inconsistent (of sample) | 54.5% | 64.1% | 62.8% |
| Inconsistent (of respondents) | 90% | 100% | 91.67% |
| Corrected (of inconsistent) | 0% | 16% | 31.8% |
| Uncorrected (of inconsistent) | 100% | 92% | 95.4% |
| Additional (of inconsistent) | 0% | 0% | 13.6% |

Table 6.6: Results from the set of questions about individual photo albums. *Respondents* represents participants who answered questions about photo albums (only those who had a photo album that was accessible beyond their friend network). *Inconsistent* represents participants who had an inconsistency between their privacy settings and their responses. *Corrected* represents participants who responded by changing their privacy settings for at least one album.*Uncorrected* represents participants who had at least one uncorrected inconsistency.*Additional* represents participants who changed the privacy settings of a photo album that they were not directly questioned about (a maximum of ten albums were shown to each participant).

| | **Likert Item** | **1st Qu.** | **Median** | **3rd Qu.** | **N** | **Y** |
|---|---|---|---|---|---|---|
| Control pre | Confident | 4 | 4 | 5 | 9.1% | 90.9% |
| Control post | Confident | 4 | 4 | 5 | 15.2% | 84.8% |
| Audience pre | Confident | 4 | 4 | 4.5 | 15.4% | 84.6% |
| Audience post | Confident | 3 | 4 | 4 | 30.8% | 69.2% |
| Posts pre | Confident | 3.5 | 4 | 4 | 25.7% | 74.3% |
| Posts post | Confident | 3 | 4 | 4 | 34.3% | 65.7% |
| Control pre | Correct | 3 | 4 | 5 | 27.3% | 72.7% |
| Control post | Correct | 3 | 4 | 4 | 27.3% | 72.7% |
| Audience pre | Correct | 3 | 4 | 4.5 | 30.8% | 69.2% |
| Audience post | Correct | 3 | 4 | 4 | 30.8% | 69.2% |
| Posts pre | Correct | 2.5 | 4 | 4 | 37.1% | 62.9% |
| Posts post | Correct | 3 | 4 | 4 | 37.1% | 62.9% |

Table 6.7: The results from the pre-study and post-study questionnaires from Likert Items about Facebook privacy settings. *Confident* = I feel confident changing the privacy settings of my Facebook account. *Correct* = I feel confident that the privacy settings of my Facebook account accurately reflect my attitude toward sharing on Facebook. The final two columns show the responses they represent yes and no responses. Note: the percentage of yes and no answers for the Correct Likert item are correct and perhaps slightly misleading because in each condition a small number of participants changed their response.

study the participants were still confident in their current privacy settings, but many lost confidence in their ability to use the privacy controls.

We further investigated the relationship between the experimental groups before and after answers regarding confidence in their ability to manage their privacy settings. For the Audience group, about 15% less of the participants were confident in their ability, comparing the values with McNemar's test z = 3.125, p = 0.0771 shows the different is nearly significant. For the Posts groups, less participants were confident in their ability though the difference is not statistically significant (z =0.3077, p = 0.5791).

## 6.4 Discussion

Although the data we collected do not support our primary hypothesis on that viewership feedback would directly affect privacy preserving behaviors, we did find evidence that viewership feedback can positively influence users' attitudes and actions toward their desired privacy settings.

### 6.4.1 Viewership Feedback Did Not Directly Influence Use of Privacy Preserving Behaviors

Our data do not support our hypothesis that reviewing viewership feedback will increase Facebook users' use of privacy preserving behaviors. In fact PPB use was not just low for the experimental groups, we also found that the participants in the control group exhibited privacy preserving behaviors that were quite similar to those exhibited by the experimental groups, even throughout the observation periods after the experimental groups were shown viewership feedback.

One possible explanation for the similarities is that the privacy focus in the pre-study questionnaire affected all of the participants. Our observations for PPB use after the pre-study questionnaire and before the first period of feedback hints that it is possible that the questionnaire encouraged PPB use (see the Table rows labeled PPB-1 in Section 6.3.1). The participants who made changes before the feedback and advice, and the participants in the control group who made changes could be described as users that understand how to use the privacy controls but need a reminder to check their settings and make changes. Well-designed viewership feedback could help this set of users maintain their privacy settings in the future.

A second explanation for the periods when we observed PPBs is that we observed behaviors that the users would have employed anyway and the pre-study questionnaire had no effect. From the PPBs we measured, it is plausible that the users are deleting friends and posts at the rate we observed [Madden, 2012]. But the number of changes to album privacy settings we observed in the control group and prior to the feedback warrants additional investigation. To the best of our knowledge, prior work has not quantitatively measured the rate of change for photo album privacy settings and so we cannot determine post hoc whether the behavior could be considered normal. Our study design should have specified a stricter control group: the control group should have installed the same application but completed survey questions unrelated to privacy to control for any effect the privacy questions might introduce.

#### 6.4.1.1 Believability of Viewership Feedback

In our study design, we were limited to randomly generated feedback that we attempted to increase the authenticity of using profile information observable through the API. In some cases the feedback might have been too general to have an effect on behavior. For example, when choosing feedback for the Posts group, we could only randomly select photos and posts as "recently viewed." For the Audience group, we generated feedback based on the known accessibility of profile data, incorporating reference to public walls or public photo albums when possible.

In some cases, participants remarked that they thought the feedback was improbable. Participants in the Posts group were more likely to express surprise or disbelief that someone would view their old posts. For example, "Many of these posts are old. I'm surprised that someone was looking at them." And, "I find it disturbing that people are going through my old posts and/or looking at old pictures." For photos, this reaction is odd because it's trivial to browse old photos by looking at a user's photo albums. Previously, it required work to view old wall posts — a user would basically have to do a linear search backward in time. However, with the introduction of Timeline, which organizes posts by year and allows direct access to previous years, it requires little effort to access old posts. We purposefully selected old posts in the attempt to show items the user had forgotten about, and to highlight the fact that Facebook does not 'forget' a post unless the user deletes it.

Participants in the Audience group questioned the feedback more than participants in the Posts group. One participant saw feedback that said, "A Facebook user browsed all of your public photos." To which the participant commented, "If one of my fb friends browsed all my photos it would take HOURS - so I'm surprised about that." We think the believability of the viewership feedback might affect the user's response, but we did not evaluate the notion in this study.

### 6.4.2 Behavior and Attitudes Affected By Viewership Feedback

Although our hypothesis on the effect of viewership feedback is not supported by the data, we did identify aspects of the participants' behaviors and attitudes that appear to have been affected by the viewership feedback. The clearest demonstration of changed behavior is the difference in the number of experimental group participants who corrected their photo album settings in the last stage of the survey. We also observed a few small changes in participants' attitudes.

#### 6.4.2.1 Corrections to Photo Album Privacy Settings

In the last stage of the study, the study application presented the participant with albums that were accessible beyond their friend network and asked the participant to select one group that they would like to share the album with (see Figure 6.1). The study application then compared the responses against the albums' privacy settings and immediately produced (1) a list of albums with inconsistent privacy settings and (2) gave explicit instructions for changing the privacy settings for an album (see Figure 6.2 and Figure 6.3).

As discussed in Chapter 4, we speculated that a participant in that study might have claimed they would not correct a sharing violation because they were unsure of which privacy control to use or because they were unsure if a privacy control existed to address their needs. In other words, we thought there was a chance that the participant might not be able to identify the underlying problem based on the information they were presented with in that study, or perhaps they were not able to map the problem to a an existing control. We attempted to eliminate this problem in this stage of the study: the study application clearly identified the inconsistencies, presented the privacy control that would address the problem, and described how to use it. We expected that the majority of the participants would respond to this information by correcting at least one of the identified inconsistencies.

In reality, though, most of our participants chose to not correct the inconsistencies we identified in their album privacy settings. Most of the participants who did respond to the information they received in this stage of the study were people who received the feedback on “recently viewed” posts. For some reason exposure to this viewership feedback seems to have influenced people to take corrective action when other people did not. The Posts viewership feedback was particularly effective in motivating participants to make changes, while the Audience feedback was less effective. Interestingly, only participants in the Posts group took an extra step toward their desired settings and changed the privacy settings for photo albums that were not directly asked about. We further discuss the implications of these observations in Chapter 7.

Earlier, we posited that there was a chance that the pre-study questionnaire affected the behavior of all participants and might have encouraged the use of privacy preserving behaviors. Despite the fact that by the time participants reached this part of the study they had answered an additional set of questions with more specific questions about the privacy settings (i.e., “Did you learn anything about managing your privacy?”), none of the participants in the control group chose to correct an inconsistent setting. This observation weakens the argument that the pre-study questionnaire put all participants in a heightened state of privacy awareness.

#### 6.4.2.2 Feedback Lowered Participants’ Confidence in Using the Privacy Controls

We asked the participants in the pre-study and post-study questionnaire about their confidence in their ability to use Facebook’s privacy controls and about their confidence in their current privacy settings. We thought the experimental groups’ confidence in using the privacy controls might increase after reading the advice for using individual privacy controls, but it seems to have had the opposite effect. The experimental groups reported a lowered confidence in their ability to manage their privacy settings, though reported the same confidence in the correctness of their privacy settings. At the end of the study, some participants commented that participating in the study made them of aware of changes that they were previously unaware of.

### 6.4.3 Limitations

Our vantage point limited our ability to collect detailed data on the participants' use of privacy preserving behaviors. This is particularly true for the use of custom lists; it would be useful to know how often Facebook users configure privacy settings using custom lists as the subject. Our measurement of the number of friends deleted in each observation could be inflated because our count could include the participant being deleted by a member of their friend network. Also, photo albums are the only objects accessible via the API that allow direct access to the privacy settings. We could collect more detailed data on privacy behaviors if this information was also available for wall posts and other items.

We could have more restrictively controlled the conditions for the control group. This would eliminate the question we mention earlier in the text about whether the pre-study questionnaire influenced participants' behavior. Considering the data we collected, it does not seem that a stricter control would have yielded additional insights about the effectiveness of viewership feedback.

## 6.5 Implications and Future Work

We found that viewership feedback and direct solicitation of sharing intentions combined with follow up feedback and advice enabled some users to achieve more accurate privacy settings. At first glance the aggregate of these might appear to be a major change to the existing mechanism, however, we claim that these would be minor additions to the existing user interface.

SNS providers could design effective viewership feedback using information they already collect and could balance the need to protect the identity of the viewer while providing useful information to the primary user. For example, feedback on viewed content could show the user a list of recently viewed posts over the past week, or even month, or it could show a list of popular posts for the lifetime of the account. The audience feedback might be slightly more cumbersome to generate in a privacy preserving manner, though some version of $k$-anonymity could be used [Sweeney, 2002]. The feedback could be displayed in a non-obtrusive manner beside the *News feed* feature where the user could glance at it each time they log in to Facebook.

The direct solicitation of sharing intentions combined with follow-up feedback and advice is similar to an audit mechanism. In our study design we took a simple approach to including this auditing feature: we asked the participants about the privacy settings for a small set of their photo albums and only presented the options that represented default granular options. The feature could be extended to include questions about the privacy settings of data types other than photo albums, and the posts could be chosen by age, number of views, or by other features of the post like subject matter.

A topic for future research is the question of *when* users would be most receptive to interacting with an audit feature. In other areas of human factors and computer security research, like the design of warning dialogs, it is clear that timing is critical to whether a user responds to an attempt to interrupt their primary task [Egelman *et al.*, 2009]. Another important question to address is — how often should users engage with such a feature to ensure long term effectiveness?

Our data suggest that the feedback on recently viewed posts was more effective than the feedback describing recent profile viewers. One explanation is that the participants felt the audience feedback was not plausible. If this is the case, then presenting real feedback might be more effective. It's also possible that the audience feedback was too vague, whereas the posts feedback was very specific. Future research could test the effectiveness of various types of feedback in an attempt to determine the critical attributes of feedback.

## 6.6 Significance

In this chapter we empirically evaluated two methods of supplementing Facebook's existing privacy settings with the goal of motivating users to change their privacy settings to reduce the gap between their desired privacy settings and their actual privacy settings. We first tested the effect of adding viewership feedback to the user experience and did not find evidence to suggest that the feedback encouraged participants to change their privacy settings. Then we tested a more direct intervention: we asked the participant to select an audience for up to ten photo albums, presented them with a list of any inconsistencies, and then provided instructions for implementing the necessary changes using the existing mechanism. We found that the participants who previously saw feedback were more likely to make changes to their privacy settings when presented with the list of albums with inconsistent privacy settings. Even though some participants corrected the inconsistencies we presented them with, not all participants were motivated to make changes to their settings based on the information provided.

We feel it is an important finding that the existing settings can be supplemented to improve users' actual privacy settings. However, determining effective methods for assisting these remaining users remains an important topic for future research. From the data we collected, we are unable to distinguish users who ignored the advice completely to users who attempted to act on the advice but were unable to follow the instructions as presented. Similar to the study presented in Chapter 4, we refrained from asking participants *why*. We did not ask *why* they chose not to or were unable to follow the advice, or *why* their privacy settings. It's unclear whether such inquiries would yield useful data, particularly because our user studies are conducted remotely.

# Chapter 7

# Conclusion and Future Work

We began with the claim that correct access control policies are critical and the assumption that policy authors must have a usable access control mechanism in order to produce correct policies. In our research on Facebook privacy settings, we found that users' privacy settings do not match their sharing intentions and that providing additional information to the user helps some users converge toward their desired privacy settings. We improved the correctness of users' privacy settings by supplementing the existing mechanism with contextual information but without modifying the actual privacy controls. In this chapter we discuss the implications of our results in terms of lessons learned for future research toward a usable access control mechanism for social network site users.

## 7.1 Lessons Learned

### 7.1.1 Users Need More Contextual Information in Addition to Usable Privacy Controls

Prior work on usable access control focuses on users' ability to complete policy authoring tasks using a specific user interface. In our work, particularly in Chapter 6, it appears that users need additional information before they reach a stage where they have a clear idea of the policy to implement using the privacy controls — the provided user interface to manage the access control policy. The positive effect of viewership feedback suggests that this information motivated users to take action and correct the inconsistencies that the study application presented.

From the way that many people interact with Facebook, there is little reason for a user to revisit their privacy settings. Maybe a user changes their settings when they first create an account, when they read a mainstream news article about Facebook and privacy, or after a sharing mishap. The positive effect of the viewership feedback and audit on the photo albums suggests users might adjust their privacy settings more frequently if presented with this additional information, though this is a topic for future research.

### 7.1.2 Users Want Fine-Grained Control

In our study of Facebook users' privacy concerns, we learned the users have a wide variety of privacy concerns that are quite individualized. Furthermore, users described their concerns both in terms of specific types of content they are wary of sharing and specific friends they are wary of sharing with. The available fine-grained controls are per-post privacy settings and limiting the audience of content by explicitly including or excluding custom lists of friends. The existing privacy controls do not give users a viable option for restricting access based on content. Consistent use of per-post privacy settings requires foresight from the user and constant awareness of their desired sharing policy. Custom friend lists require the user to create practical lists of friends that match their desired sharing policy. It is unclear whether users are able to successfully create practicable friend lists [Jones and O'Neill, 2010; Kelley *et al.*, 2011].

### 7.1.3 Users Need Assistance Formulating Their Desired Policy

Much of the work on usable access control assumes that the user has a desired policy they would like to implement using the policy authoring mechanism. Evaluating the process by which users formulate their desired privacy settings is a topic for future research. Currently, many Facebook users manage their privacy settings when they create their Facebook account, at that point in time users have no notion of who will be in their friend network or what information they will share using the service. In the study presented in Chapter 6, several users expressed surprise at viewership feedback that featured old photos or posts. From the comments, it seemed that some participants forgot about content they had previously shared on their profile but was clearly still available for other users to access. The *News feed* feature might support a mental model that the information posted is ephemeral. Future work could also investigate whether the introduction of *Timeline* alters this mental model.

### 7.1.4 Better Metrics for Measuring the Effectiveness of an Access Control Mechanism

In our early work, we evaluated the correctness of privacy settings by the number of inconsistencies identified by comparing the user's sharing intentions against their actual privacy settings. After completing the two subsequent studies, we wonder if that is still a useful metric. On one hand, it is important to have a quantitative metric for determining the effectiveness of an interface. But on the other hand, a single realized mishap that involves over-sharing to the wrong audience can have disastrous effects. The metric might depend on the user and on their privacy concerns.

## 7.2 Directions for Future Work

We chose to experiment with additions to the existing user interface to increase the use of privacy preserving behaviors by altering the users' viewership awareness, knowledge of current privacy settings, and knowledge of their options for managing their privacy settings. Our additions to the Facebook user experience motivated people to make changes to their settings and ultimately improved their actual Facebook privacy settings. However, there were many participants who experienced no measurable benefit from the additions we introduced. Thus, a large opportunity remains to develop an improved understanding of users' needs that can inform the design of other enhancements to the privacy controls.

The study presented in Chapter 4 left us with a similarly perplexing result — that users did not intend to correct the privacy violations they confirmed. We did not explore this result further in that study and reasoned that a remote study is the wrong format for determining why a user would be unwilling, or uninterested in correcting their privacy settings when they were obviously inaccurate. In this study, despite knowing the limitations of our prior choices of methodology, we evaluated a similar question though in the follow up study we added instructions for how to use specific features. We found that the instructions helped some users and that they made changes to their privacy settings but others did not. An interesting question for future work would be to evaluate whether a study with similar conditions conduced in the lab results in more people responding to the feedback on inconsistencies. If users' make changes in a lab under otherwise the exact same conditions, what would we learn about the results? We would know that the mechanism is usable if the user is properly motivated, but we would not know the right way to motivate the user to take action.

Another direction for future work is determining whether Facebook's 900 million users could be meaningfully classified into distinct groups that would benefit from different types of access control mechanisms. With such a high number of users, one would expect that a single approach to the privacy settings would not suit most users. Similar to the features in SPARCLE and Adage that allowed policy authors to enter policy in two different format and toggle between modes, what mode would be useful to Facebook users? [Karat *et al.*, 2005; Zurko *et al.*, 1999].

## 7.3 Conclusion

In computer security, we rely on policy-based mechanisms like firewalls and intrusion detection systems to protect our computer systems. These protection mechanisms will blindly enforce the policy as specified, and for this reason there is a strict requirement for correctness. We are particularly interested in access control policies, where an incorrect policy can lead to unwanted consequences such as the wrong party being granted access to a resource, or the correct party being denied access to a resource. Presently, most security policies must be managed by a user. This fact, combined with the correctness requirement, implies that usability is critical to the practicability of an access control system. While prior work on usable access control mechanisms has focused on the design of user interfaces for policy management, in our work we focus on the usability of an existing access control mechanism — Facebook privacy settings, an access control domain where millions of end-users are asked to manage an ACL-based policy for their shared content.

In our research, we found evidence of inconsistencies between Facebook users' *desired* policy and the *actual* access control policy for their shared content, and that even when made aware of these inconsistencies many of our participants claimed that they would not fix them. In other words, simply presenting the user with information about errors in their privacy settings is not an effective method to motivate users to take corrective action.

In our study of users' privacy concerns, we found that users are concerned about who has access to their shared content and that their concerns are nuanced and individualized. Our findings dispel the notion that a user would be unmotivated to take corrective action when presented with inconsistencies due to a lack of concern. Whereas, previous studies of privacy concerns gather data using interview techniques, we measured users' interpersonal concerns by asking questions about people in their friend network and items they had shared on Facebook. We also found that users employ a multitude of privacy-preserving behaviors to mitigate their concerns, such as utilizing the coarse-grained privacy controls, they use the existing fine-grained controls much less and prefer to rely on ad-hoc strategies like deleting content or friends.

Our modifications to the existing mechanism enabled some of our participants to adjust their privacy settings so that they more accurately reflect their sharing intentions. Perhaps because the study was conducted remotely, we observed that a subset of our participants did not respond to our modifications by making the prescribed changes. Since we were able to assist some participants but not all, we conclude that additional research on SNS users' access control needs is necessary. In particular, research with a focus on determining the reasons why the existing mechanisms fail, and developing an understanding of the modifications that will motivate various types of users.

# Bibliography

[Acquisti and Gross, 2006] Alessandro Acquisti and Ralph Gross. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *6th Workshop on Privacy Enhancing Technologies*, pages 36–58. Springer Berlin, 2006.

[Altman, 1975] Irwin Altman. *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding.* Brooks/Cole Publishing Company, 1975.

[Balfanz, 2003] Dirk Balfanz. Usable access control for the World Wide Web. In *Proceedings of the 19th Annual Computer Security Applications Conference*, ACSAC '03, pages 406–415, 2003.

[Bartal *et al.*, 2004] Yair Bartal, Alain Mayer, Kobbi Nissim, and Avishai Wool. Firmato: A novel firewall management toolkit. *ACM Trans. Comput. Syst.*, 22(4):381–420, 2004.

[Bauer *et al.*, 2008] Lujo Bauer, Lorrie Faith Cranor, Robert W. Reeder, Michael K. Reiter, and Kami Vaniea. A user study of policy creation in a flexible access-control system. In *Proceedings of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 543–552, NY, NY, USA, 2008. ACM.

[Bauer *et al.*, 2009] Lujo Bauer, Lorrie Faith Cranor, Robert W. Reeder, Michael K. Reiter, and Kami Vaniea. Real life challenges in access-control management. In *Proceedings of the Twenty-Seventh Annual SIGCHI conference on Human Factors in Computing Systems*, CHI '09, pages 899–908. ACM, 2009.

[Benisch *et al.*, 2011] Michael Benisch, Patrick Gage Kelley, Norman Sadeh, and Lorrie Faith Cranor. Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. *Personal Ubiquitous Comput.*, 15(7):679–694, October 2011.

[Besmer and Richter Lipford, 2010] Andrew Besmer and Heather Richter Lipford. Moving beyond untagging: photo privacy in a tagged world. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, CHI '10, pages 1563–1572, New York, NY, USA, 2010. ACM.

[Boyd and Ellison, 2007] Danah M. Boyd and Nicole B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13:210–230, 2007.

[Boyd and Hargittai, 2010] Danah Boyd and Eszter Hargittai. Facebook privacy settings: Who cares? *First Monday*, 15(8), August 2010.

[Brodie *et al.*, 2005] Carolyn Brodie, Clare-Marie Karat, John Karat, and Jinjuan Feng. Usable security and privacy: A case study of developing privacy management tools. In *Proceedings of the First Symposium on Usable Privacy and Security*, SOUPS '05, pages 35 – 43, 2005.

[Brodie *et al.*, 2008] Carolyn Brodie, David George, Clare-Marie Karat, John Karat, Jorge Lobo, Mandis Beigi, Xiping Wang, Seraphin B. Calo, Dinesh C. Verma, Alberto E. Schaeffer Filho, Emil Lupu, and Morris Sloman. The coalition policy management portal for policy authoring, verification, and deployment. In *IEEE Workshop on Policies for Distributed Systems and Networks*, POLICY '08, pages 247–249, 2008.

[Caine *et al.*, 2011] Kelly Caine, Lorraine G. Kisselburgh, and Louise Lareau. Audience visualization influences disclosures in online social networks. In *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 1663–1668, New York, NY, USA, 2011. ACM.

[Cao and Iverson, 2006] Xiang Cao and Lee Iverson. Intentional access management: making access control usable for end-users. In *Proceedings of the 2nd Symposium on Usable Privacy and Security*, SOUPS '06, pages 20–31, NY, NY, USA, 2006. ACM.

[Chen, 2010] Stephanie Chen. Divorce attorneys catching cheaters on facebook. `http://articles.cnn.com/2010-06-01/tech/facebook.divorce.lawyers_1_privacy-settings-social-media-facebook`, June 2010.

[Cheswick *et al.*, 2003] William R. Cheswick, Steven M. Bellovin, and Aviel D. Rubin. *Firewalls and Internet security: repelling the wily hacker*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 2003.

[Consolvo *et al.*, 2005] Sunny Consolvo, Ian E. Smith, Tara Matthews, Anthony LaMarca, Jason Tabert, and Pauline Powledge. Location disclosure to social relations: why, when, & what people want to share. In *Proceedings of the SIGCHI International Conference on Human Factors in Computing Systems*, CHI '05, pages 81–90, New York, NY, USA, 2005. ACM.

[Cranor *et al.*, ] Lorrie Cranor, Marc Langheinrich, Massimo Marchiori, Martin Presler-Marshall, and Joseph Reagle. The Platform for Privacy Preferences 1.0 (P3P 1.0) Specification. W3C Recommendation, April 2002.

[Cranor *et al.*, 2006] Lorrie Faith Cranor, Praveen Guduru, and Manjula Arjula. User interfaces for privacy agents. *ACM Trans. Comput.-Hum. Interact.*, 13(2):135–178, 2006.

[Das *et al.*, 2010] Tathagata Das, Ranjita Bhagwan, and Prasad Naldurg. Baaz: a system for detecting access control misconfigurations. In *Proceedings of the 19th USENIX Security Symposium*, USENIX Security '10, pages 161–176. USENIX Association, 2010.

[Egelman and Johnson, 2012] Serge Egelman and Maritza Johnson. How good is good enough? the sisyphean struggle for optimal privacy settings. In *Proceedings of the CSCW Workshop Reconciling Privacy with Social Media Workshop*, 2012.

[Egelman *et al.*, 2009] Serge Egelman, Janice Tsai, Lorrie Faith Cranor, and Alessandro Acquisti. 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*, CHI '09, pages 319–328, New York, NY, USA, 2009. ACM.

[Egelman *et al.*, 2011] Serge Egelman, Andrew Oates, and Shriram Krishnamurthi. Oops, I did it again: Mitigating repeated access control errors on Facebook. In *Proceedings of the 29th SIGCHI International Conference on Human Factors in Computing Systems*, CHI '11, pages 2295–2304, New York, NY, USA, 2011. ACM.

[Facebook, 2012] Facebook. `http://newsroom.fb.com/content/default.aspx?NewsAreaId=22`, 2012.

[Gross and Acquisti, 2005] Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, WPES '05, pages 71–80, New York, NY, USA, 2005. ACM.

[Hampton *et al.*, 2011] Keith Hampton, Lauren Sessions Goulet, Lee Rainie, and Kristen Purcell. Social networking sites and our lives. `http://pewinternet.org/Reports/2011/Technology-and-social-networks.aspx`, June 2011.

[Hart *et al.*, 2008] M. Hart, R. Johnson, and A. Stent. More content-less control: Access control in the Web 2.0. In *Proceedings of the First Workshop on Online Social Networks*, WOSP '08, pages 43–48, 2008.

[Heussner, 2010] Ki Mae Heussner. Teacher loses job after commenting about students, parents on facebook. `http://abcnews.go.com/Technology/facebook-firing-teacher-loses-job-commenting-students-parents/story?id=11437248#.T1dZ_5hq5Ic`, August 2010.

[istrategylabs, 2010] istrategylabs. Facebook demographics and statistics report. http://www.istrategylabs.com/2010/06/facebook-demographics-and-statistics-report-june-2010—-privacy-concerns-dont-stop-growth/, June 2010.

[Joinson, 2008] Adam N. Joinson. Looking at, looking up or keeping up with people?: motives and use of Facebook. In *CHI '08: Proceedings of the SIGCHI International Conference on Human Factors in Computing Systems*, pages 1027–1036, New York, NY, USA, 2008. ACM.

[Jones and O'Neill, 2010] Simon Jones and Eamonn O'Neill. Feasibility of structural network clustering for group-based privacy control in social networks. In *Proceedings of the 6th Symposium on Usable Privacy and Security*, SOUPS '10, pages 9:1–9:13. ACM, 2010.

[Karat *et al.*, 2005] John Karat, Clare-Marie Karat, Carolyn Brodie, and Jinjuan Feng. Privacy in information technology: Designing to enable privacy policy management in organizations. In *International Journal of Human-Computer Studies*, volume 63, pages 153–174. Elsevier, 2005.

[Karat *et al.*, 2009] John Karat, Clare-Marie Karat, Elisa Bertino, Ninghui Li, Qun Ni, Carolyn Brodie, Jorge Lobo, Seraphin Calo, Lorrie Cranor, Ponnurangam Kumaraguru, and Robert W. Reeder. A policy framework for security and privacy management. In *IBM Journal Research & Development*, volume 53, pages 4:1–4:14, 2009.

[Kelley *et al.*, 2011] Patrick Gage Kelley, Robin Brewer, Yael Mayer, Lorrie Faith Cranor, and Norman Sadeh. An investigation into Facebook friend grouping. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction*, INTERACT '11, pages 216–233, Berlin, Heidelberg, 2011. Springer-Verlag.

[King *et al.*, 2011] Jennifer King, Airi Lampinen, and Alex Smolen. Privacy: is there an app for that? In *Proceedings of the Seventh Symposium on Usable Privacy and Security*, SOUPS '11, pages 12:1–12:20, New York, NY, USA, 2011. ACM.

[Klemperer *et al.*, 2012] Peter Klemperer, Yuan Liang, Michelle Mazurek, Manya Sleeper, Blase Ur, Lujo Bauer, Lorrie Faith Cranor, Nitin Gupta, and Michael Reiter. Tag, you can see it!: using tags for access control in photo sharing. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems*, CHI '12, pages 377–386, New York, NY, USA, 2012. ACM.

[Krasnova *et al.*, 2009] Hanna Krasnova, Oliver Günther, Sarah Spiekermann, and Ksenia Koroleva. Privacy concerns and identity in online social networks. *Identity in the Information Society*, 2:39–63, 2009.

[Krishnamurthy and Wills, 2008] Balachander Krishnamurthy and Craig E. Wills. Characterizing privacy in online social networks. In *Proceedings of the First Workshop on Online Social Networks*, WOSP '08, pages 37–42, New York, NY, USA, 2008. ACM.

[Lampe *et al.*, 2006] Cliff Lampe, Nicole Ellison, and Charles Steinfield. A Face(book) in the crowd: social searching vs. social browsing. In *Proceedings of the 20th Conference on Computer Supported Cooperative Work*, CSCW '06, pages 167–170. ACM, 2006.

[Lampe *et al.*, 2008] Cliff Lampe, Nicole B. Ellison, and Charles Steinfield. Changes in use and perception of Facebook. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, CSCW '08, pages 721–730, New York, NY, USA, 2008. ACM.

[Lampinen *et al.*, 2009] Airi Lampinen, Sakari Tamminen, and Antti Oulasvirta. All my people right here, right now: management of group co-presence on a social networking site. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, Group '09, pages 281–290. ACM, 2009.

[Lederer *et al.*, 2004] Scott Lederer, Jason Hong, Anind Dey, and James Landay. Personal privacy through understanding and action: five pitfalls for designers. *Personal Ubiquitous Comput.*, 8(6):440–454, 2004.

[Lewis *et al.*, 2008] Kevin Lewis, Jason Kaufman, and Nicholas Christakis. The taste for privacy: An analysis of college student privacy settings in an online social network. *Journal of Computer-Mediated Communication*, 14(1):79–100, 2008.

[Lipford *et al.*, 2008] Heather Richter Lipford, Andrew Besmer, and Jason Watson. Understanding privacy settings in Facebook with an audience view. In *Proceedings of the 1st Conference on Usability, Psychology, and Security*, UPSEC '08, pages 2:1–2:8, Berkeley, CA, USA, 2008. USENIX Association.

[Lipford *et al.*, 2010] Heather Richter Lipford, Jason Watson, Michael Whitney, Katherine Froiland, and Robert W. Reeder. Visual vs. compact: A comparison of privacy policy interfaces. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, CHI '10, pages 1111–1114, 2010.

[Little, 2012] Lyneka Little. Court okays facebook party photos in workers comp claim. `http://abcnews.go.com/blogs/business/2012/02/court-okays-facebook-party-photos-in-workers-comp-claim/`, February 2012.

[Liu *et al.*, 2011] Yabing Liu, Krishna P. Gummadi, Balachander Krishnamurthy, and Alan Mislove. Analyzing facebook privacy settings: user expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, pages 61–70, New York, NY, USA, 2011. ACM.

[Madden and Smith, 2010] Mary Madden and Aaron Smith. Reputation management and social media. `http://pewinternet.org/Reports/2010/Reputation-Management.aspx`, May 2010.

[Madden, 2012] Mary Madden. Privacy management on social media sites. `http://pewinternet.org/Reports/2012/Privacy-management-on-social-media.aspx`, February 2012.

[Madejski *et al.*, 2012] Michelle Madejski, Maritza Johnson, and Steven M. Bellovin. A study of privacy settings errors in an online social network. In *Proceedings of the 4th IEEE International Workshop on Security and Social Networking*, SESOC '12, 2012.

[Maxion and Reeder, 2005] Roy A. Maxion and Robert W. Reeder. Improving user-interface dependability through mitigation of human error. *International Journal of Human-Computer Studies*, 63(1-2):25 – 50, 2005.

[Palen and Dourish, 2003] Leysia Palen and Paul Dourish. Unpacking "privacy" for a networked world. In *Proceedings of the SIGCHI International Conference on Human Factors in Computing Systems*, CHI '03, pages 129–136, New York, NY, USA, April 5-10 2003. ACM.

[Reeder and Maxion, 2005] Robert W. Reeder and Roy A. Maxion. User interface dependability through goal-error prevention. *International Conference on Dependable Systems and Networks*, pages 60–69, 2005.

[Reeder *et al.*, 2007] Robert W. Reeder, Clare-Marie Karat, John Karat, and Carolyn Brodie. Usability challenges in security and privacy policy-authoring interfaces. In *Proceedings of the 11th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part II*, INTERACT '07, pages 141–155, Berlin, Heidelberg, 2007. Springer-Verlag.

[Reeder *et al.*, 2008a] Robert W. Reeder, Lujo Bauer, Lorrie Faith Cranor, Michael K. Reiter, Kelli Bacon, Keisha How, and Heather Strong. Expandable grids for visualizing and authoring computer security policies. In *Proceedings of the SIGCHI International Conference on Human Factors in Computing Systems*, CHI '08, pages 1473–1482, NY, NY, USA, 2008. ACM.

[Reeder *et al.*, 2008b] Robert W. Reeder, Patrick Gage Kelley, Aleecia M. McDonald, and Lorrie Faith Cranor. A user study of the expandable grid applied to P3P privacy policy visualization. In *WPES '08: Proceedings of the 7th ACM Workshop on Privacy in the Electronic Society*, pages 45–54, New York, NY, USA, 2008. ACM.

[Reeder *et al.*, 2011] Robert W. Reeder, Lujo Bauer, Lorrie F. Cranor, Michael K. Reiter, and Kami Vaniea. More than skin deep: measuring effects of the underlying model on access-control system usability. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*, CHI '11, pages 2065–2074, New York, NY, USA, 2011. ACM.

[Reeder, 2008] Robert W. Reeder. *Expandable Grids: A user interface visualization technique and a policy semantics to support fast, accurate security and privacy policy authoring.* PhD thesis, Carnegie Mellon University, Pittsburgh, PA, July 2008.

[Sadeh *et al.*, 2009] Norman Sadeh, Jason Hong, Lorrie Cranor, Ian Fette, Patrick Kelley, Madhu Prabaker, and Jinghai Rao. Understanding and capturing people's privacy policies in a mobile social networking application. *Personal Ubiquitous Comput.*, 13(6):401–412, August 2009.

[Saltzer and Schroeder, 1975] Jerome H. Saltzer and Michael D. Schroeder. The protection of information in computer systems. *Proceedings of the IEEE*, 63(9):1278–1308, Sept. 1975.

[Saltzer, 1974] Jerome H. Saltzer. Protection and the control of information sharing in multics. *Commun. ACM*, 17:388–402, July 1974.

[Samarati and Vimercati, 2001] Pierangela Samarati and Sabrina De Capitani di Vimercati. Access control: Policies, models, and mechanisms. In *IFIP WG 1.7 International School on Foundations of Security Analysis and Design on Foundations of Security Analysis and Design: Tutorial Lectures*, FOSAD '00, pages 137–196, London, UK, 2001. Springer-Verlag.

[Sandhu *et al.*, 1996] Ravi S. Sandhu, Edward J. Coyne, Hal L. Feinstein, and Charles E. Youman. Role-based access control models. *Computer*, 29(2):38–47, 1996.

[Schneier, 2010] Bruce Schneier. A taxonomy of social networking data. *IEEE Security and Privacy*, 8:88, 2010.

[Skeels and Grudin, 2009] Meredith Skeels and Jonathan Grudin. When social networks cross boundaries: a case study of workplace use of Facebook and Linkedin. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, Group '09, pages 95–104. ACM, 2009.

[Smetters and Good, 2009] D. K. Smetters and Nathan Good. How users use access control. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, SOUPS '09, pages 15:1–15:12, New York, NY, USA, 2009. ACM.

[Smith, 2009] Justin Smith. Fastest growing demographic on Facebook: Women over 55, February 2 2009. http://www.insidefacebook.com/2009/02/02/fastest-growing-demographic-on-facebook-women-over-55/.

[Stutzman and Hartzog, 2009] Frederic D. Stutzman and Woodrow Hartzog. Boundary Regulation in Social Media. *SSRN eLibrary*, 2009.

[Stutzman and Kramer-Duffield, 2010] Fred Stutzman and Jacob Kramer-Duffield. Friends only: examining a privacy-enhancing behavior in Facebook. In *Proceedings of the 28th SIGCHI International Conference on Human Factors in Computing Systems*, CHI '10, pages 1553–1562, New York, NY, USA, 2010. ACM.

[Sweeney, 2002] Latanya Sweeney. k-anonymity: a model for protecting privacy. *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems*, 10(5):557–570, October 2002.

[Tang *et al.*, 2010] Karen P. Tang, Jialiu Lin, Jason I. Hong, Daniel P. Siewiorek, and Norman Sadeh. Rethinking location sharing: exploring the implications of social-driven vs. purpose-driven location sharing. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, Ubicomp '10, pages 85–94, New York, NY, USA, 2010. ACM.

[Tsai *et al.*, 2009] Janice Y. Tsai, Patrick Kelley, Paul Drielsma, Lorrie Faith Cranor, Jason Hong, and Norman Sadeh. Who's viewed you?: the impact of feedback in a mobile location-sharing application. In *Proceedings of the 27th SIGCHI International Conference on Human Factors in Computing Systems*, CHI '09, pages 2003–2012, New York, NY, USA, 2009. ACM.

[Tufekci, 2008] Zeynep Tufekci. Can you see me now? audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology & Society*, 28(1):20–36, 2008.

[Venkatanathan *et al.*, 2011] Jayant Venkatanathan, Denzil Ferreira, Michael Benisch, Jialiu Lin, Evangelos Karapanos, Vassilis Kostakos, Norman Sadeh, and Eran Toch. Improving users' consistency when recalling location sharing preferences. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction*, INTERACT'11, pages 380–387, Berlin, Heidelberg, 2011. Springer-Verlag.

[Wang *et al.*, 2011] Yang Wang, Gregory Norcie, and Lorrie Faith Cranor. Who is concerned about what? a study of American, Chinese and Indian users' privacy concerns on social network sites. In *Proceedings of the 4th International Conference on Trust and*

*Trustworthy Computing*, TRUST'11, pages 146–153, Berlin, Heidelberg, 2011. Springer-Verlag.

[Whitten and Tygar, 1999] Alma Whitten and J. D. Tygar. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In *Proceedings of the 8th USENIX Security Symposium*, USENIX Security '99, pages 169–184. USENIX Association, 1999.

[Wool, 2004] Avishai Wool. A quantitative study of firewall configuration errors. *IEEE Computer*, 37(6):62–67, 2004.

[Young and Quan-Haase, 2009] Alyson L. Young and Anabel Quan-Haase. Information revelation and Internet privacy concerns on social network sites: a case study of Facebook. In *Proceedings of the 4th International Conference on Communities and Technologies*, C&T '09, pages 265–274, New York, NY, USA, 2009. ACM.

[Zurko *et al.*, 1999] Mary Ellen Zurko, Richard T. Simon, and Tom Sanfilippo. A user-centered, modular authorization service built on an RBAC foundation. In *Proceedings of the 1999 IEEE Symposium on Security and Privacy*, pages 57–71, 1999.

# Appendix A

# User Study Material for Empirical Evaluation of the Correctness of Users' Facebook Privacy Settings

1. Choose one of the following options to represent what you believe is the most important reason for online privacy.
    - Economic Security: To prevent identity theft and protecting browsing habits from advertisers and third parties.
    - Reputation Security: To hide information to protect my social reputation.
    - Physical Security: To protect me physically, by hiding my face, location, and/or contact information from strangers.
2. Are you concerned with online privacy as related to "economic security?" Economic security refers to preventing identity theft and protecting browsing habits from advertisers and third parties.
    - Why would I be concerned?
    - I'm not concerned.
    - I'm a little concerned.

- I am concerned.
- I'm very concerned.

3. Are you concerned with online privacy as related to "reputation security?" Reputation security refers to hiding information to protect my social reputation.

   - Why would I be concerned?
   - I'm not concerned.
   - I'm a little concerned.
   - I am concerned.
   - I'm very concerned.

4. Are you concerned with online privacy as related to "physical security?" Physical security refers to protecting me physically, by hiding my face, location, and/or contact information from strangers.

   - Why would I be concerned?
   - I'm not concerned.
   - I'm a little concerned.
   - I am concerned.
   - I'm very concerned.

5. Do you feel your Facebook settings reflect your attitude related to privacy?

   - Yes.
   - No.
   - I am not concerned with privacy.

6. If not, why not?

   - Either my settings reflect my privacy attitude or I do not care about privacy.
   - I know how to change the settings but I don't have the time to do it.
   - I don't know how to change the settings.

- Facebook doesn't have the privacy controls I want.

7. For which reasons have you untagged (not HIDDEN) photos? Please rank each reason.

   - I have chosen not to post any photos.
   - I have never untagged a photo.
   - The photo displayed my face or location, which I have chosen to keep secret to protect my physical security.
   - I didn't like the photo of me (it was unattractive or not flattering).
   - The photo displayed behavior I did not to be associated with (something that could be embarrassing if others saw it).

8. For which reasons have you hid posted factual data (e.g. Birthday, hometown, gender, etc)?

   - The information could potentially be used for identity theft.
   - I do not feel safe with that information out there since I believe I could be potentially stalked, found, and/or harmed.
   - I don't want other people to know how old I am or where I am from for professional or social reasons.
   - I have not hid any information.
   - I have chosen not to enter factual data on my profile.

9. Why do you use Facebook? Check all that apply for the respective groups: People you don't personally know, Friends, Friends of Friends, Network Members.

   - Keep people informed about my life (i.e. status updates, photo uploads)
   - Finding information on about people (i.e. profile watching).
   - Finding information on people's daily lives (i.e. newsfeed, status updates).
   - Personal communication (i.e. messages, walls, etc).
   - Being socially informed (i.e. events, groups).

10. Have you ever had an accidental leak of information on Facebook that had a negative impact? If so, what happened? If there have been multiple leaks, please pick the one with the largest impact.

    - I never had an accidental leak of information.
    - I was a victim of identity theft or my account was hacked into.
    - I was physically harmed, stalked, or contacted by someone I did not want to see due to the release of information.
    - Information sensitive to my social reputation was viewed by a friend or coworker who I did not want see the information.

11. What type of information was accidentally leaked?

    - Fact-type information (gender, birthday) available on my profile.
    - Activity (photo, status update) information posted by me.
    - Activity (photo, status update) information posted by others.

12. What do you think was the cause of the information leak?

    - I'm not sure. I didn't think the person would be able to view it based on my privacy settings.
    - I had not changed any of my privacy settings.
    - I had not changed my privacy settings for that type of information.
    - I didn't remember that information was on my profile.
    - I didn't expect my friend to post that information about me on Facebook.

13. If you suffered from a leak, did you alter your behavior? How? Check all that apply.

    - I paid closer attention to privacy on Facebook on a per-activity basis, such as modifying the privacy level of each individual status update.
    - I did not alter my behavior.
    - I became more selective about the information I put on Facebook.

- I disabled some features of Facebook, like my wall.
- I made changes to my privacy settings so that it wouldn't happen again.
- I deleted the piece of information.
- I deleted that friend or put them on a limited profile view.

14. Have you heard anything regarding Facebook privacy in the news lately? Check all that apply.

    - I haven't heard anything.
    - I haven't heard anything but, then again, I don't really read the news.
    - I read a headline or heard something briefly I didn't really care to investigate further.
    - I heard something about Facebook and it seemed negative but I don't know any further details.
    - I heard something about Facebook and it seemed positive but I don't know any further details.
    - Facebook released my private information to advertisers.
    - Facebook released my private information to the general public.
    - Facebook has released an improved privacy interface.
    - Facebook has released a new privacy interface.

15. Where did you hear it from? Check all that apply.

    - Somebody told me in person/over the phone.
    - General news source.
    - On Facebook (such as from a user's status).
    - Privacy-related source.

16. Has the media affected your behavior on Facebook at all on Facebook? If not, why? Check all that apply.

- It has not affected my behavior at all.
- I became more selective about the information I posted on Facebook.
- I deleted some Facebook friends.
- I modified my privacy settings to be less private.
- I modified my privacy settings to be more private.
- I double-checked my privacy settings but didn't change them.

# Appendix B

# User Study Material for Survey of Facebook Users' Privacy Concerns and Mitigation Strategies

## B.1 Survey Questions about Facebook Usage and Privacy Preserving Behaviors

The following questions were presented one at a time with the options listed below each question.

1. How long have you had a Facebook account?
   - Less than 1 year.
   - Between 1 and 2 years.
   - More than 2 years.
2. How often do you use Facebook?
   - Several times a day.
   - Once a day.
   - Once every few days.

- Once a week.
- Once a month.
- Less than once a month.

3. About how much time do you spend on Facebook reading your news feed each week?

    - Less than 1 hour.
    - Between 1 and 2 hours.
    - Between 2 and 4 hours.
    - 4 hours or more.

4. About how much time do you spend on Facebook posting information and updating your profile each week?

    - Less than 1 hour.
    - Between 1 and 2 hours.
    - Between 2 and 4 hours.
    - 4 hours or more.

5. About how much time do you spend on Facebook browsing your friend's profiles or photos each week?

    - Less than 1 hour.
    - Between 1 and 2 hours.
    - Between 2 and 4 hours.
    - 4 hours or more.

6. How frequently do you use Facebook for the following?

    - To look up information about a friend.
    - To communicate with a friend.
    - To upload photos.

- To view photos your friends uploaded.
- To share a link to a news story.
- To browse the profiles of people that are friends with your Facebook friends.
- To browse the profiles of people that you do not know.
- To find new friends.

7. Are you Facebook friends with:

   - Members of your immediate family (parents/siblings).
   - Members of your extended family.
   - Coworkers.
   - People you know from high school/college/grad school.
   - People you met through friends.
   - People you have not met in person.

8. How many Facebook friends do you have? If you're unsure, an estimate is fine.

   - 0-99
   - 100-299
   - 300-599
   - 600 or more

9. Do you have more than one Facebook account?

   - Yes.
   - No.

10. How many accounts do you have?

    - 2
    - 3
    - 4

11. Have you ever used the option to change the privacy settings of a single status update?

    - Yes.
    - No.
    - Please describe an instance where you changed the settings of a single update.

12. Have you ever turned down a friend request?

    - Yes.
    - No.

13. Why did you turn down the request?

    - You did not know the person.
    - You knew the person but did not want them to see your profile.
    - You suspected the profile was fake.

14. Have you ever unfriended someone?

    - Yes.
    - No.

15. Why did you unfriend them?

    - You were no longer friends in real life.
    - You did not want to share your Facebook profile with them any longer.
    - You were unsure whether you knew them.

16. Have you ever sent a friend request to someone you did not know in person?

    - Yes.
    - No.

17. Why did you friend them?

    - They were a friend of one of your Facebook friends.

- You had common interests.
- You were interested in meeting them offline.

18. Have you changed your privacy settings such that some of your Facebook friends have limited access to your profile?

    - Yes.
    - No.

19. How do you know the friends that have limited access to your profile?

    - Members of your immediate family (parents/siblings).
    - Members of your extended family.
    - Coworkers.
    - People you know from high school/college/grad school with.
    - People who you met through friends.

20. Have you ever untagged yourself in a photo that was posted by a friend?

    - Yes.
    - No.

21. Why did you untag yourself in your friend's photo?

    - You didn't want anyone to see it.
    - You didn't want a specific person to see it.
    - You didn't like it.

22. Have you ever deleted a photo you uploaded to Facebook?

    - Yes.
    - No.

23. Why did you delete the photo you uploaded?

- You didn't want anyone to see it.
- You didn't want a specific person to see it.
- You didn't like it.

24. Have you ever asked a Facebook friend to delete a photo they uploaded of you?

    - Yes.
    - No.

25. Why did you ask your Facebook friend to delete the photo?

    - You didn't want anyone to see it.
    - You didn't want a specific person to see it.
    - You didn't like it.

26. Have you ever posted a status update or comment and deleted it later?

    - Yes.
    - No.

27. Why did you choose to delete the post?

    - You didn't want anyone to see it.
    - You didn't want a specific person to see it.
    - You didn't like it.

28. Have you ever deleted a comment that was posted by a Facebook friend?

    - Yes.
    - No.

29. Why did you delete the comment?

    - You didn't want anyone to see it.
    - You didn't want a specific person to see it.
    - You didn't like it.
    - You suspected the comment was spam.

## B.2 General Scenarios

The following questions were displayed one at a time as Likert items on a 5-point scale with the anchor points unconcerned, indifferent, and concerned.

1. A stranger will see an inappropriate photo or comment on my profile.

2. A family member will see an inappropriate photo or comment on my profile.

3. A coworker will see an inappropriate photo or comment on my profile.

4. An employer will see an inappropriate photo or comment on my profile.

5. Your employer using Facebook to monitor your conduct while you're at work.

6. Your employer using Facebook to monitor your conduct while you're away from work.

7. Thieves using Facebook to track, monitor, locate, and identify you as a potential victim.

8. Your employer using the information on your Facebook profile to assess your suitability for the company.

9. Law enforcement using Facebook to track illegal activities (illegal drug use, underage drinking, etc.).

10. Your university using Facebook postings, personal information, and images to identify you as a university code violator.

11. Sexual predators using Facebook to track, monitor, locate, and identify you as a potential victim.

12. Political parties using Facebook to target you through the use of advertisements and data mining.

## B.3 Questions about Profile Information

The following questions were presented one at a time. The possible answers were presented as percentages from 0% to 100% in increments of ten.

1. What percentage of your Facebook friends do you trust with access to your profile and shared information?

2. What percentage of the NETWORK-NAME network members do you trust with access to your profile data?

The following question was asked for each of the participant's custom friend lists.

1. Why did you create the friend list LIST-NAME and what do you use it for? Friend lists are a Facebook feature for grouping friends. If you cannot remember, enter that as your response.

### B.3.1 Individual Facebook Friends

The following questions were asked for nine of the participants' Facebook friends.

1. What is your relationship to FRIEND-NAME? Select the answer that fits best.

   - A member of your immediate family (parent/sibling).
   - A member of your extended family.
   - A coworker.
   - Someone you know from high school/college/grad school.
   - A friend of a friend.
   - Someone you have not met in person.
   - Someone you socialize with in person.
   - Not sure.

2. How do you feel about FRIEND-NAME viewing all the information you have uploaded to Facebook?

### B.3.2 Individual Posts

The following questions were asked for ten of the participants' posts.

1. How would you feel if a stranger saw this?

2. How would you feel if a member of your immediate family saw this?

3. How would you feel if a member of your extended family saw this?

4. How would you feel if a coworker saw this?

5. How would you feel if someone you know from high school/college/grad school saw this?

6. How would you feel if a friend of a friend saw this?

7. How would you feel if someone you have not met in person saw this?

8. How would you feel if someone you socialize with in person saw this?

# Appendix C

# User Study Material for Empirical Evaluation of the Effect of Viewership Feedback

## C.1 Pre-study Questionnaire

This set of questions were completed by all participants after they were approved for participation and assigned to a group.

### C.1.1 Measure of Technical Skills

The following questions were displayed one at a time as Likert items on a 5-point scale with the anchor points strongly disagree, neutral, and strongly agree.

1. I am confident that I can correctly change the privacy settings of my Facebook account.

2. I am confident that I can correctly post a comment on a blog.

3. I am confident that I can correctly vote on the quality of content available on sites where users rate content (such as YouTube).

4. I am confident that I can correctly upload a video to a video-sharing site (such as YouTube).

5. I am confident that I can correctly submit a review about a product or service (on sites such as Amazon or Yelp).
6. I am confident that I can correctly create a quiz or poll for friends to take online.
7. I am confident that I can find advice in an online discussion group when I need help with something.
8. I am confident I know the difference between http and https.
9. I am confident I could edit the information on an existing Wikipedia entry.

### C.1.2 Privacy Concerns and Facebook Use

The following questions were displayed one at a time as Likert items on a 5-point scale with the anchor points strongly disagree, neutral, and strongly agree.

1. I am often concerned that Facebook could store my information for the next couple of years.
2. I am often concerned that Facebook could share the information I provide with other parties (e.g. marketing, HR or government agencies).
3. I am often concerned other parties (e.g. marketing, HR, government agencies) could actually collect my publicly available information on Facebook.
4. It often worries me that other users might purposefully write something undesired about me on Facebook.
5. I am often concerned that other users might take advantage of the information they learned about me through Facebook.
6. I am often concerned that I don't have control over the actions of other users.
7. I find time to keep my profile up-to-date.
8. When I have something to say, I like to share it on Facebook.
9. I am always honest in the information I provide on Facebook.

10. When I express myself on Facebook, I always consider who can see the information I publish.
11. I think carefully how much I reveal about myself on Facebook.
12. I am often concerned that someone I don't expect (e.g. a stranger, my "ex", my parents, teacher, boss) could view my profile on Facebook.
13. I feel uncomfortable that many people might follow changes in my profile.
14. If I was in a job application process I would make many changes to my profile.
15. I feel confident changing the privacy settings of my Facebook account.
16. I feel confident that the privacy settings of my Facebook account accurately reflect my attitude toward sharing on Facebook.
17. Since creating my Facebook account, I've changed my privacy settings:
    - Never
    - Once
    - 2-3 times
    - 4 or more times

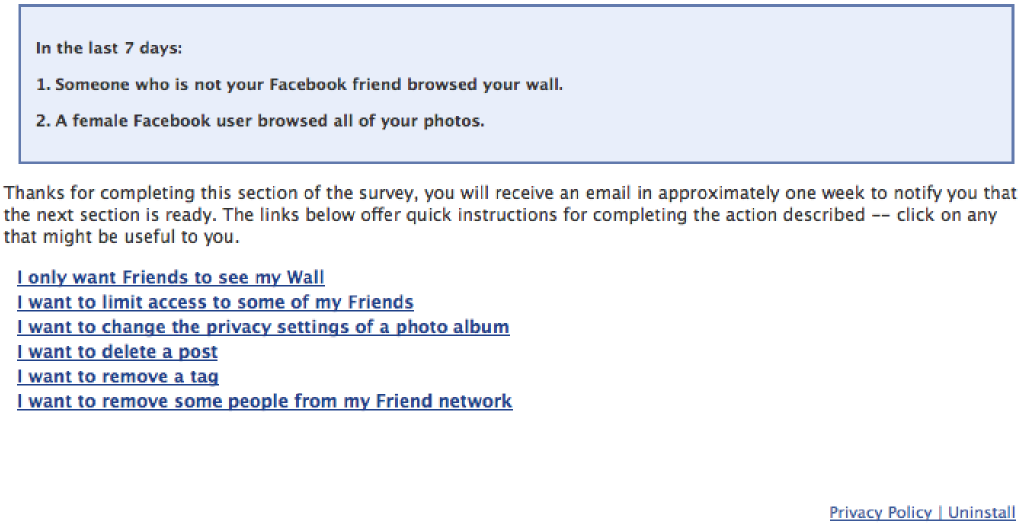
## C.2 Questions about Viewership Feedback

The following questions were displayed one at a time as Likert items on a 5-point scale with the anchor points strongly disagree, neutral, and strongly agree.

1. This information is what I expected.
2. I am a little surprised by this.
3. I am comfortable with this information.

## C.3 Instructions for Privacy Preserving Behaviors

After the participants responded to the Likert items about each round of feedback, the application displayed a list of privacy-preserving behaviors for their consideration.



**In the last 7 days:**

**1. Someone who is not your Facebook friend browsed your wall.**

**2. A female Facebook user browsed all of your photos.**

Thanks for completing this section of the survey, you will receive an email in approximately one week to notify you that the next section is ready. The links below offer quick instructions for completing the action described -- click on any that might be useful to you.

I only want Friends to see my Wall
I want to limit access to some of my Friends
I want to change the privacy settings of a photo album
I want to delete a post
I want to remove a tag
I want to remove some people from my Friend network

Privacy Policy | Uninstall

## C.4 Post-study Questionnaire

The following questions were displayed one at a time as Likert items on a 5-point scale with the anchor points strongly disagree, neutral, and strongly agree.

1. I am often concerned that Facebook could store my information for the next couple of years.

2. I am often concerned that Facebook could share the information I provide with other parties (e.g. marketing, HR or government agencies).

3. I am often concerned other parties (e.g. marketing, HR, government agencies) could actually collect my publicly available information on Facebook.

4. It often worries me that other users might purposefully write something undesired about me on Facebook.

5. I am often concerned that other users might take advantage of the information they learned about me through Facebook.

6. I am often concerned that I don't have control over the actions of other users.
7. I find time to keep my profile up-to-date.
8. When I have something to say, I like to share it on Facebook.
9. I am always honest in the information I provide on Facebook.
10. When I express myself on Facebook, I always consider who can see the information I publish.
11. I think carefully how much I reveal about myself on Facebook.
12. I am often concerned that someone I don't expect (e.g. a stranger, my "ex", my parents, teacher, boss) could view my profile on Facebook.
13. I feel uncomfortable that many people might follow changes in my profile.
14. If I was in a job application process I would make many changes to my profile.
15. I feel confident changing the privacy settings of my Facebook account.
16. I feel confident that the privacy settings of my Facebook account accurately reflect my attitude toward sharing on Facebook.
17. Since creating my Facebook account, I've changed my privacy settings:
    - Never
    - Once
    - 2-3 times
    - 4 or more times
18. Since enrolling in this study, I've changed my privacy settings:
    - Never
    - Once
    - 2-3 times
    - 4 or more times