My Profile Is My Password, Verify Me! The Privacy/Convenience Tradeoff of Facebook Connect

Serge Egelman

University of California, Berkeley
egelman@cs.berkeley.edu

We performed a laboratory experiment to study the privacy tradeoff offered by Facebook Connect: disclosing Facebook profile data to third-party websites for the convenience of logging in without creating separate accounts. We controlled for trustworthiness and amount of information each website requested, as well as the consent dialog layout. We discovered that these factors had no observable effects, likely because participants did not read the dialogs. Yet, 15% still refused to use Facebook Connect, citing privacy concerns. A likely explanation for subjects ignoring the dialogs while also understanding the privacy tradeoff—our exit survey indicated that 88% broadly understood what data would be collected—is that subjects were already familiar with the dialogs prior to the experiment. We discuss how our results demonstrate informed consent, but also how habituation prevented subjects from understanding the nuances between individual websites' data collection policies.

Author Keywords
Privacy; Facebook Connect; user study

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces; K.4.1 Computers and Society: Public Policy Issues

General Terms
Security; Human Factors; Experimentation

INTRODUCTION
In a seminal 2007 study, Florêncio and Herley showed that the average Internet user has around 25 password-protected accounts [10]. As the web continues to grow, the number of password-protected accounts that users maintain will increase. While users may not use a unique password for each account, they must still remember which password was used for which account. Single Sign-On (SSO) systems solve this problem by allowing users to authenticate to multiple websites using a single set of credentials.

CHI 2013, April 27–May 2, 2013, Paris, France.

Facebook Connect is likely the most used SSO system. In 2010, Facebook claimed that each month 250 million people were using it to authenticate to third-party websites [15]. As of 2012, as many as eight million websites allow users to authenticate via Facebook [21]. Like other OAuth-based systems [8], Facebook Connect offers users a value proposition: the convenience of a single set of credentials in exchange for granting relying websites access to certain Facebook profile information.

When users attempt to authenticate using Facebook Connect, they are presented with consent dialogs that outline the information collected if they proceed. A dialog may indicate that a website is requesting access to minimal data, such as the user's name and gender. Alternately, websites may make requests for data beyond the defaults, such as a user's interests (e.g., political affiliation, favorite movies, or even sexual orientation). It is not clear whether the current consent dialogs make this tradeoff clear to users.

We are unaware of any researchers who have performed controlled experiments to quantify the proportion of users who accept the privacy/convenience tradeoff offered by Facebook Connect. We are also unaware of previous research that has examined the extent to which informed consent is achieved, as well as how users' decisions might change as a function of both how much information is requested and the trustworthiness of the recipient. We examined these questions by performing a laboratory experiment. We contribute the following:

- We perform a controlled experiment to quantify the proportion of users who are willing to use Facebook Connect to authenticate to various websites.
- We show that users are surprisingly cognizant of their disclosures; 88% understood the types of Facebook profile data that websites might request.
- We show that despite demonstrating a broad understanding of data collection practices, users are unlikely to notice nuances, which we believe is due to habituation. Thus, improvements are needed to highlight data collection practices that are likely to diverge from users' expectations.

BACKGROUND
Our work is informed by prior research in the areas of web single sign-on, online informed consent, and the usability of current mechanisms for information disclosure.

Web Single Sign-On

Despite the wide availability of SSO systems, websites (referred to as "relying parties") have been slow adopters until very recently. The main incentive for users is the ability to use "one account to rule them all." Sun et al. posited that the biggest barrier to adoption was a lack of incentives for relying parties [26]. For instance, websites can use registration forms to collect personal information that may be unavailable from identity providers.

The OAuth protocol has addressed some of these incentives [1]. OAuth-based SSO systems allow a relying party to request profile information from the identity provider (e.g., Facebook in the case of Facebook Connect). This provides relying parties with a strong incentive to participate, as they can now collect information about their users that they otherwise might not have been able to collect, even with lengthy registration forms.

The closest related work to our experiment was Sun et al.'s study of users' OpenID security concerns when using their webmail credentials to authenticate [27]. Forty percent of their participants were hesitant to release personal information, with 26% going so far as to request fake OpenID accounts to complete the study. In real life, this option would not be available: users unwilling to release their profile information would either have to create a new non-SSO account or discontinue the task. Thus, it is not clear how users might behave when faced with this more realistic choice. Likewise, it is unclear whether informed consent is being achieved: were participants truly unconcerned or did they simply not understand the terms of the agreement?

Online Informed Consent

As ubiquitous computing has become a reality and the perception of control over one's personal information has decreased, various researchers have proposed privacy guidelines for providing users with adequate notice about how their information may be used [2, 20]. Chief among these principles is the notion of informed consent [19]. Friedman et al. suggested that informed consent is a five-step process [11]:

1. **Disclosure**: Are the costs and benefits of providing the information presented to the user?
2. **Comprehension**: Does the user understand the disclosure statement?
3. **Voluntariness:** Is the user coerced into disclosing?
4. **Competence:** Is the user of sound mind to make a decision about disclosure?
5. **Agreement:** Is the user given ample opportunity to make a decision?

Friedman et al. first applied these principles to the web in order to raise user awareness of cookies [12]. Grossklags and Good demonstrated that informed consent was not being achieved with software end-user license agreements (EULAs) [14]. Good et al. expanded on this work through a series of studies in which they observed that comprehension problems could be decreased through the use of short summaries, which increased user attention prior to installing software [13]. However, they observed that short summaries were not a panacea: many users still proceeded with installations and then regretted those decisions afterwards. Böhme and Köpsell found that software dialogs designed similarly to EULAs were more likely to be ignored [6]. Others have since tried to improve the design of EULAs [17, 22].

Mechanisms for Disclosure

Recent information disclosure research has examined applications on the Facebook platform, which use consent dialogs very similar to the ones used by Facebook Connect. Besmer and Lipford examined misconceptions about how data gets shared with Facebook applications and concluded that users wish to disclose less [3]. King et al. performed a survey of Facebook application users and concluded that many only begin thinking about privacy after experiencing adverse events [18]. While many users use community ratings to decide whether an application will use data appropriately, Chia et al. found that these may not be trustworthy [7].

Others have proposed tools to allow users to limit their disclosures. Shehab et al. suggested a framework to allow users to specify their disclosure preferences [24]. Felt and Evans found that most Facebook applications functioned with a subset of the requested information and therefore proposed a proxy to limit disclosures [9]. Besmer et al. proposed fine-grained policy authoring tools so that users can specify what information they are comfortable sharing [4]. However, Wang et al. found that users are no more likely to authorize applications when given granular privacy controls [28]. Others have proposed recommender systems to help users make disclosures decisions [5, 23].

However, all of this research has examined consent for disclosing information to applications. We believe this is a different case—despite similar interfaces—from SSO authentication because the former violates Friedman et al.'s voluntariness principle [11]: users who want to use the applications have no choice but to accept the stated terms, whereas in the SSO context, users often have the option of simply creating a separate account. Thus, we believe that the question of achieving informed consent with Facebook Connect remains heretofore unexplored.

METHODOLOGY

When users attempt to log into a website using Facebook Connect, they are shown a consent dialog that indicates certain data from their Facebook profiles will be transferred to the website if they proceed (Figure 1). Users then have the choice to proceed or cancel. If they cancel, they can either use a different login method (e.g., creating an account specifically for that website or using a different SSO provider that may transmit different information) or abandon their task altogether. The initial motivation for our experiment was to examine whether informed consent was being achieved in this context.

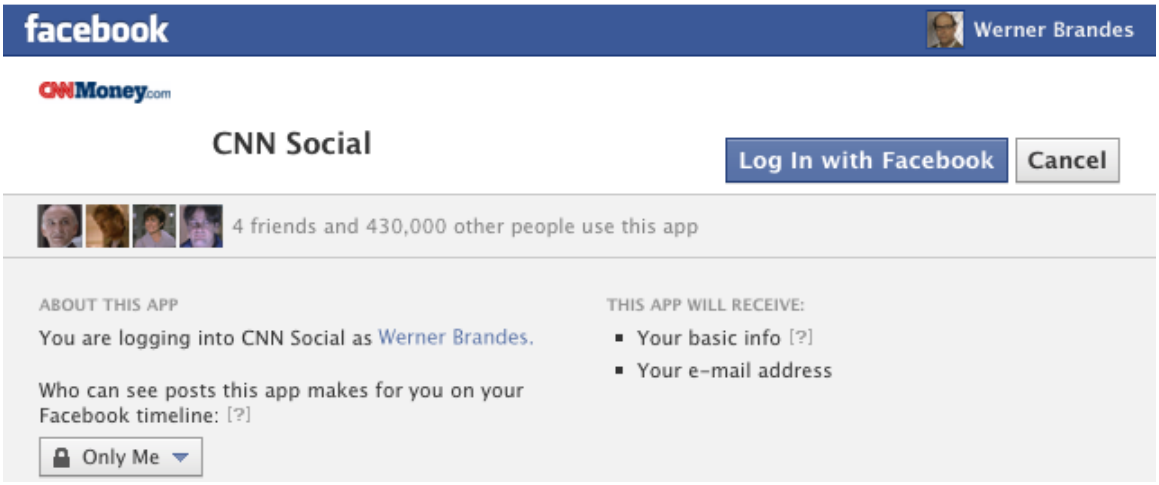

Figure 1. Screenshot of the Facebook Connect consent dialog, as seen by participants in the *control* condition.

We designed a laboratory experiment to examine the extent to which participants understood how their personal information was changing hands when using Facebook Connect. In this section we describe our experimental conditions, the websites visited, and our protocol.

Conditions

By default, websites using Facebook Connect receive "basic info." If users drag their mice over this phrase, they discover that "basic info" includes the following information from their Facebook profiles:

- Name
- Profile picture
- Gender
- Networks
- User ID
- List of friends

The information above is in addition to any other information on their profiles that is publicly viewable. For example, if a user has not changed her privacy settings, she may inadvertently allow a website to also view status updates, comments, or photo albums. Websites also have the option of requesting additional information: the Facebook API specifies permissions so that websites can request nearly any piece of information present in a user's Facebook profile, regardless of whether or not that information is viewable by other human beings; the interpersonal privacy settings do not apply to information requested through Facebook Connect.

We hypothesized that the aforementioned method of presenting privacy information to users was inadequate, and that if their relevant profile information were shown verbatim, they would be less likely to use Facebook Connect. We tested this theory by creating a GreaseMonkey[1] script that redrew the consent dialogs using data screen-scraped from each participant's Facebook profile in realtime. Thus, participants would be allowed to see their information, prior to sharing it with websites. We refer to this as the *verbatim* condition (Figure 2).

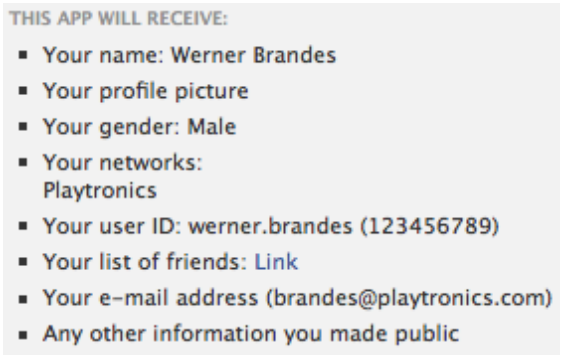

Figure 2. In the *verbatim* condition, the right side of the consent dialog listed participants' actual profile data.

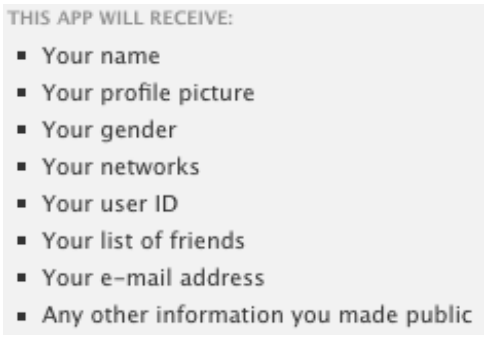

Figure 3. In the *list* condition, the right side of the consent dialog featured a list of the requested profile information.

[1]GreaseMonkey is a client-side plugin for Firefox that allows custom scripts to be executed on user-specified websites. http://www.greasespot.net/

In order to accommodate this additional information, we were forced to change the layout of the dialog into a bulleted list. Because this change resulted in a dramatic increase in the amount of text shown on the screen, and because the change might be immediately obvious to participants familiar with Facebook Connect, we created an intermediate condition to control for this. The *list* condition expanded the same information as the *control* condition into a bulleted list format (Figure 3). Thus, our three between-group conditions were as follows:

- *Control*—The layout that Facebook Connect used at the time of our experiment (Figure 1).
- *List*—The same information as the *control* condition, but expanded into a bulleted list (Figure 3).
- *Verbatim*—The layout of the *list* condition, however, each bullet contained information from participants' actual Facebook profiles (Figure 2).

The GreaseMonkey script randomly assigned each participant to one of the three between-subjects conditions at the beginning of the experiment and ensured that each participant remained in the same condition on subsequent websites throughout the experiment.

Websites

We observed participants visit three different websites that all used Facebook Connect. We chose these three websites to control for two different factors: the amount of profile information that each website requested and the extent to which participants might trust each website with access to their data.

We decided to design our tasks around retrieving information from news websites. As such, we needed two websites that requested the same amount of information, along with a third website that requested a superset of this information. Likewise, of the two websites that requested the lesser amount of data, one needed to be more trustworthy than the other. Eventually, we settled on the following three websites:

- CNN (http://www.cnn.com/)
- The Sun (http://www.thesun.co.uk/)
- Reuters (http://www.reuters.com/)

We chose these websites because CNN and Reuters are known as relatively neutral U.S. news sources, whereas The Sun is a British tabloid. CNN and The Sun both collect the "basic info" described previously, though The Sun also collects email addresses. Reuters collects the "basic info," email addresses, locations, and birthdays. Since The Sun collected email addresses, unlike CNN, and because we were concerned that Reuters did not collect enough additional information to make the contrast apparent, we used some deception. We designed our GreaseMonkey script to deceive participants into believing that more information was being requested. For example, the dialogs stated that all three websites collected email addresses, so that CNN and The Sun would appear to collect the same information (Table 1).

		CNN	The Sun	Reuters
Basic info	Name	✓	✓	✓
	Profile picture	✓	✓	✓
	Gender	✓	✓	✓
	Networks	✓	✓	✓
	User ID	✓	✓	✓
	List of friends	✓	✓	✓
	Email address	✓	✓	✓
Additional info	Birthday			✓
	Location			✓
	Hometown			✓
	Relationship status			✓
	Sexual orientation			✓
	Employment history			✓
	Education history			✓

Table 1. The amount of data that the consent dialogs indicated each website was requesting.

Finally, we also chose these three websites because in addition to allowing users to log in via Facebook Connect, they also offered the option of creating new accounts. We felt that it was critically important to offer participants alternative ways of completing each task in order to minimize the Milgram effect; if participants felt compelled to use Facebook Connect, our experiment would have been testing their ability to follow instructions, rather than their willingness to compromise privacy for convenience.

Protocol

We told participants that they would view each of the three websites in order to answer questions about the features that each offered. We asked participants what features became available once they logged in to each website. In reality, we did not care about participants' responses to these questions and instead we were only interested in whether or not they used Facebook Connect to log in or if they created new accounts on each website. We hypothesized that most participants would view the Facebook Connect consent dialogs, but that based on the experimental conditions, a subset of participants would choose not to proceed in order to protect their personal information from disclosure. We ran screen capture software on each computer to capture this data.

During August of 2012, we recruited participants from the Bay Area Craigslist, offering participants $35 to participate in a one-hour "social media" study. Prior to scheduling, we directed participants to an online screening survey to ensure that they had Facebook accounts for at least six months and were at least eighteen years old. In addition to questions to mask our screening requirements, we also determined whether or not they used the new "timeline" profile format or the previous format, since our scripts only worked on the newer format. We scheduled participants who qualified to attend one of seven laboratory sessions.

We split 87 eligible participants into cohorts of up to eighteen. Participants in each cohort arrived at our laboratory and selected seats in front of computers separated by partitions so that each participant could not view the screens of other participants. Once participants signed

consent forms, we handed them instructions that summarized the protocol. After giving them time to read the instructions, we read the instructions aloud:

1. *In this study, you will be asked to visit three different news websites. While on each of these websites, you will need to browse around in order to answer the questions on the task description sheet. Please fill in your responses on the sheet to the best of your ability.*
2. *Some of the questions will require you to log in to the websites. You can do this by either creating a new account on each of these websites or by using "Facebook Connect." Facebook Connect allows you to log in to other websites using your Facebook account information. The method you choose is completely up to you.*
3. *On some of the websites, you may be asked to view a confirmation email after logging in or creating a new account. Please do this from within the web browser.*
4. *Once you complete a task sheet, raise your hand and the experimenter will give you the next task. Once you have completed all three tasks, you will be asked to complete an online survey about your experiences.*

We then handed participants their first task. We randomized the order in which each participant visited each of the three websites. As they completed a task, we handed them the next task until they completed all three. Finally, they completed an exit survey. Once complete, we compensated them and handed them a debriefing sheet. When participants left, we stopped the video capture software and reset the settings on each computer so as to erase all cookies and browser history.

RESULTS

We performed our laboratory experiment to test the following alternate hypotheses about Facebook Connect:

H_1: Participants who are shown verbatim examples of the data that websites request will be significantly more likely to abandon using Facebook Connect.

H_2: Participants will be significantly more likely to abandon using Facebook Connect on websites that request more data.

H_3: Participants will be significantly more likely to abandon using Facebook Connect on untrusted websites than trusted websites.

In the remainder of this section, we present our results in terms of the behaviors that we observed, participants' awareness of each website's data collection practices, the extent to which they trusted each website with their data, and whether participants engaged in other strategies to protect their personal information.

Observed Behaviors

To help explain our experimental results, our exit survey included an open-ended question about why they chose whether or not to use Facebook Connect on each of the three websites. This gave rise to a confound that we otherwise would not have identified: sixteen participants claimed that they used Facebook Connect solely because they believed it was required to participate in the study. Despite attempts to minimize the Milgram effect by offering participants an alternative authentication mechanism—creating a new account on each website—a minority still felt compelled. Thus, we were forced to remove these sixteen subjects. Another six subjects never logged in to any of the three websites,[2] which forced us to remove them as well, leaving our remaining sample size at 65.

These 65 subjects ranged in age from 18 to 59, with an average of 31 ($\sigma = 10.3$). Sixty-eight percent of our subjects were female, while 32% were male. We compared our sample's observed demographic data with the expected values from a 2012 demographic survey of Facebook users [25], and observed no statistically significant differences with regard to gender ($\chi^2_1 = 3.074$, $p < 0.080$) nor age ($\chi^2_3 = 3.545$, $p < 0.315$). However, our sample was significantly more educated than the average Facebook user ($\chi^2_3 = 46.297$, $p < 0.0001$). Regardless, we observed no significant differences based on whether or not participants used Facebook Connect with regard to any of these demographic factors.

Table 2 shows the high-level results for each website. Since some participants did not attempt to log in to some of the websites, the sample sizes were not constant across the three websites. Likewise, because the three between-subjects conditions were assigned randomly when a consent dialog was first displayed, ten participants (15% of 65) were never assigned to a condition because they never attempted to use Facebook Connect on any of the websites, proceeding directly to creating new accounts. Another two participants' condition assignments could not be determined from our screen capture videos because they accepted the dialogs before they had fully loaded.

Overall, we were surprised to discover that only one participant refused to proceed with Facebook Connect after viewing a consent dialog; the rest either proceeded with Facebook Connect regardless of what the dialogs said, or they refused to use Facebook Connect prior to seeing the dialogs. Furthermore, this participant was in the *control* condition. Thus, we observed no statistically significant differences between conditions based on how the data was presented to participants (i.e., the *control*, *list*, or *verbatim* conditions). Therefore, we cannot accept H_1 nor reject the null hypothesis.

One possible explanation for the lack of observable effect is that participants did not read the dialogs. Without using an eye tracker, it is impossible to determine this with certainty. However, we used our screen capture videos to measure the amount of time that had elapsed between the dialogs loading and participants clicking the button to proceed. Our theory was that if participants

[2]There is no reason to believe that these six subjects declined to log in due to privacy concerns. The screen capture videos indicated that they simply misunderstood the task: all of them clicked the "like" button on the websites and then claimed they had completed the task.

	CNN									
	Control		*List*		*Verbatim*		*Unknown*		**Total**	
Used Facebook Connect	15	(94%)	10	(91%)	15	(88%)	2	(17%)	42	(75%)
Declined Facebook Connect	-		-		-		-		-	
Created new account	1	(6%)	1	(9%)	2	(12%)	10	(83%)	14	(25%)

	The Sun									
	Control		*List*		*Verbatim*		*Unknown*		**Total**	
Used Facebook Connect	16	(94%)	12	(92%)	13	(81%)	2	(17%)	43	(74%)
Declined Facebook Connect	-		-		-		-		-	
Created new account	1	(6%)	1	(8%)	3	(19%)	10	(83%)	15	(26%)

	Reuters									
	Control		*List*		*Verbatim*		*Unknown*		**Total**	
Used Facebook Connect	17	(94%)	12	(92%)	20	(95%)	2	(17%)	51	(80%)
Declined Facebook Connect	1	(6%)	-		-		-		1	(1%)
Created new account	-		1	(8%)	1	(5%)	10	(83%)	12	(19%)

Table 2. Summary of the results indicating the login method participants used on each website: proceeding with Facebook Connect, seeing the Facebook Connect dialog and then choosing to create a separate account on the website, and creating a separate account on the website without ever seeing the Facebook Connect consent dialog.

read the dialogs, those in the *verbatim* and *list* conditions would spend significantly longer than those in the *control* condition. Because we were worried about habituation effects on the subsequent dialogs after the first, we only tested this for the first dialog to which participants were exposed. We observed no statistically significant differences between the three conditions: the median reading times were $7s$ in the *control* and *list* conditions, and $6.0s$ in the *verbatim* condition ($\chi^2_2 = 0.132$, $p < 0.936$; Kruskal-Wallis test). These results suggest that participants failed to notice the changes we made to the consent dialogs. Therefore, in the remainder of this paper, we analyzed the three between-subjects conditions together.

Data Collection Perceptions

One of the two reasons for having participants visit three websites was to control for the amount of Facebook profile information that participants believed each website was requesting. Our hypothesis, H_2, was that participants would be less likely to proceed logging in with Facebook Connect on a website that requested more profile data than the others (see Table 1). We had them visit the Reuters website for this purpose. Despite observing a single participant opt out of using Facebook Connect after seeing the consent dialog on this website, we could not draw statistically significant comparisons with the other two websites. Thus, H_2 cannot be accepted.

In our exit survey, participants listed the types of data they believed each website was requesting via Facebook Connect. Only three participants (5% of 65) indicated that they believed Reuters was collecting substantially more data than the other two websites. This corroborates our theory that participants did not thoroughly read the consent dialogs in the laboratory and therefore did not notice the subtle differences between conditions.

While they may not have understood the nuances in data collection policies between the three websites, this does not mean that they were generally unaware of the privacy cost of using Facebook Connect. That is, participants may not have read the dialogs in the laboratory because they had encountered them previously, had become habituated to them, and therefore chose not to read them because they believed they already knew what the dialogs said. In the exit survey, participants reported whether they had previously used Facebook Connect on a 5-point Likert scale ("never," "annually," "monthly," "weekly," and "daily"), depicted in Figure 4. Of the 54 (83% of 65) participants who used Facebook Connect in our experiment at least once, only 26% claimed to have never used it prior to our experiment. This suggests that many participants may have already understood the basic value proposition from prior knowledge and had become habituated to future dialogs.

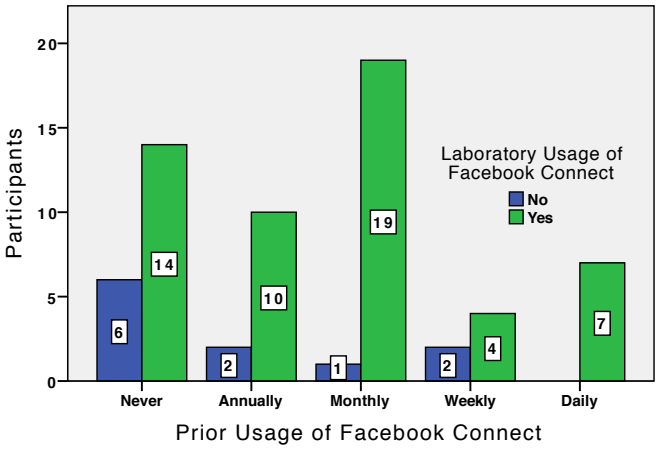

Figure 4. Participants' self-reported prior usage and observed laboratory usage of Facebook Connect.

We asked participants to list the types of data they believed each website would collect if they used Facebook Connect to log in and found that most participants understood that some amount of their profile data would be collected. Because we thought it unreasonable for them to name the complete set of items listed in Table 1, we accepted answers that mentioned a subset of this information. Examples of acceptable answers included:

- *"Email, name, gender, location, friend connections."*
- *"Email address, location, basic info."*
- *"Name and email address, probably."*
- *"Picture, profile name, basic info, age, city."*

Overall, we found that 88% of our 65 participants had a basic understanding of the privacy cost. This comprehension rate did not observably change as a function of whether or not participants used Facebook Connect in our experiment. In fact, of the 45 participants who claimed to have used Facebook Connect at least once prior to our experiment, 96% understood that they were disclosing profile information. This further corroborates our theory that participants did not read the dialogs because they were already familiar with them.

Nineteen participants (29% of 65) believed that *all* of their profile data would be transferred to the websites:

- *"All of the personal information that we submit when registering for Facebook, all of the things we have liked and whatever other information they can gather from what we have posted in the past."*
- *"Any and everything that is on your FB account."*
- *"I tend to believe the worst, so all of it."*
- *"I would think that they could access any information associated with my FB profile, even if it's not marked available to the public."*

Surprisingly, this erroneous belief did not prevent participants from using Facebook Connect: there was no observable correlation between believing a website would receive all of a participant's Facebook profile data and whether that participant used Facebook Connect to log in to that website. On CNN, nine of ten participants believed all of their Facebook data would be transferred yet used Facebook Connect anyway. On both Reuters and The Sun, this proportion was eleven out of fourteen.

Thus, while participants did not pay attention to the details of the consent dialogs during the experiment, almost all of them understood that some amount of their Facebook profile data would be released to the requesting websites upon logging in with Facebook Connect.

Trusting Data Recipients

The second reason why we had participants visit three different websites was so that we could control for whether or not participants trusted the websites with their data. We chose The Sun to test this hypothesis, H_3, for two reasons. First, since it is based in a foreign country, we reasoned that many participants may simply be less familiar with it than the other two U.S.-based websites. Second, for participants familiar with The Sun, we reasoned that they may trust it less because it is a tabloid. We validated this design decision in the exit survey by asking participants to use a 5-point Likert scale ("strongly disagree," "disagree," "unsure," "agree," and "strongly agree") to rate the extent to which they trust each website with their Facebook profile data.

Using a Wilcoxon Signed Ranks test, we observed that we were correct: significantly more people trusted CNN ($Z = 4.673$, $p < 0.0005$) and Reuters ($Z = 4.960$, $p < 0.0005$) than The Sun, while there was no observable difference between CNN and Reuters ($Z = 0.426$, $p < 0.670$). The median response for both CNN and Reuters was "unsure," whereas it was "disagree" for The Sun. Despite this varying level of trust, we observed no effect on participants' decisions to use Facebook Connect. Thus, we cannot accept H_3.

We examined whether our results could have been confounded by participants who already had accounts on one or more of the three websites. While five participants claimed to have already had accounts on the CNN website, we observed no correlation with whether they chose to use Facebook Connect in our experiment. None of our participants had accounts on either Reuters or The Sun's website. We therefore conclude that neither participants' prior relationship with each website nor the extent to which they trusted each website with their profile data had an observable impact on their decisions to grant those websites access to their Facebook profile data.

Privacy-Preserving Strategies

Ultimately, we were curious why participants chose whether or not to use Facebook Connect on each website. In our exit survey we explicitly asked this open-ended question about each of the three websites visited. Of the participants who refused to use Facebook Connect, almost all of them explicitly mentioned privacy: 12 participants on CNN (84% of 14), 14 participants on The Sun (93% of 15), and 12 participants on Retuers (92% of 13). Examples of these explanations included:

- *"I don't want CNN knowing my information."*
- *"Dozens of times I've clicked on Facebook Connect. But when I see the dialogue box that says something about giving permission to share content or share my information or share something, I ALWAYS cancel."*
- *"I don't like volunteering my information just for an easier way to log in."*
- *"I don't use Facebook if there's a manual sign up option. I don't think what I do on Facebook is the business of The Sun or any other site."*
- *"It's a garbage newspaper and I don't want them spamming me or finding info out about me...they are not trusted!"*

The corollary is that those who chose to proceed with using Facebook Connect, despite understanding the privacy implications, did so out of convenience: 34 participants on CNN (81% of 42), 36 participants on The Sun (84% of 43), and 38 participants on Reuters (75% of 51). Examples of these explanations included:

- *"Easier because I already have a Facebook account."*
- *"Easier to log in, I also trust CNN."*
- *"I didn't want to spend the time to fill out a big form to be a new user and have to remember [a] new password and username."*

	Everyone	Friends	Only Me	Custom
CNN	5 (12%)	17 (40%)	12 (29%)	8 (19%)
The Sun	5 (12%)	19 (44%)	12 (28%)	7 (16%)
Reuters	4 (8%)	23 (45%)	16 (31%)	8 (16%)

Table 3. The parties with whom posts originating on each of the three websites would be shared.

- *"It is more convenient. After the Facebook login was invented, I never went back to creating a new account because you have to think of a creative username, which might be taken, and a password, and then confirm all that information with your email. Facebook Connect is just one click of a button."*
- *"It is faster to use Facebook Connect and not much of a commitment as you can disconnect at any time."*

While participants who were more concerned with convenience than privacy were more likely to use Facebook Connect, this does not mean that they did not take additional steps to limit their information exposure. In addition to specifying what profile information would be accessible to websites, the consent dialogs also allowed users to modify who could see posts made to their Facebook profiles originating from these websites (Figure 1). The choices available to users were "everyone," "friends," "only me," "custom," and any user-defined lists of friends. If users do not change their default privacy settings from within Facebook, the default is "everyone," which was the case for only six participants (11% of the 54 who logged in with Facebook Connect). In fact, 34 participants (63% of 54) had the default set to "friends," thirteen (24% of 54) had this set to "custom," while a single participant had this set to "only me." This indicates that 89% had previously modified their Facebook privacy settings, which corroborates Johnson et al.'s findings that most users are adept at limiting strangers from accessing their profiles [16]. But these are just the defaults that appeared when the consent dialog was first displayed: a majority of participants (59% of 54) changed these defaults to further restrict access.

Table 3 depicts the parties with whom websites' posts would be shared, for the participants who used Facebook Connect. Of the 32 participants who changed their defaults during the experiment, every single one of them selected "only me." While we cannot say anything about those who had "custom" as their default setting, our results indicate that for participants who accepted the privacy tradeoff involved with using Facebook Connect, they took steps to mitigate the flow of information to additional parties, which indicates informed consent.

DISCUSSION

Facebook Connect offers users a privacy/convenience tradeoff: use Facebook credentials to log into third party websites (convenience benefit) while disclosing personal information (privacy cost), or protect personal information (privacy benefit) by creating separate accounts on each website (convenience cost). We observed that most users understood this tradeoff: those who cared more about convenience than privacy used Facebook Connect, whereas those who cared more about privacy did not. But at the same time, our data suggests that participants were acting out of bounded rationality: most participants opted not to read the details of each website's data collection policy because they felt they already had an idea of what the policies said, and therefore participants did not understand the differences between the policies of varying websites. We discuss our results within the context of Friedman et al.'s principles of informed consent [11], and then conclude with design implications.

Disclosure

We examined whether users understood the privacy cost of Facebook Connect by creating three between-subjects conditions to vary how the information was presented. We found that when given additional details, participants were no less likely to use Facebook Connect. Our exit survey showed that participants had a broad understanding of data collection policies, likely from prior exposures to the consent dialogs. Thus, during the tasks they did not pay enough attention to the dialogs to notice nuances between them, likely because they were habituated to them. This reflects a potential shortcoming in the informed consent process: participants may be failing to notice disclosures that diverge from their expectations.

Comprehension

We observed that 88% of our participants exhibited a basic understanding that their Facebook profile information would be transferred to the websites. They may have gained this knowledge from previous exposures or through other sources, such as media stories or word of mouth. Regardless of how they learned about Facebook Connect's value proposition, participants demonstrated comprehension of the default data collection policies (i.e., requests for "basic info"); our data suggests that informed consent failures occurred due to participants not noticing additional disclosures, rather than noticing but not understanding disclosures.

Voluntariness

Despite documented problems with similar dialogs for granting application permissions [3], our participants understood the privacy/convenience tradeoff when using Facebook Connect. While we cannot directly compare our results with other studies that were performed at different times and under different conditions, a key difference between our results and previous research appears to be context. When the dialogs are used for application permissions, the user has but one choice: she must accept the privacy cost to use the application. In the context of SSO, the user makes a decision: accept the privacy/convenience tradeoff of Facebook Connect or maintain privacy and create a separate account. Thus, when used for SSO, the dialogs succeed at the voluntariness principle where the same dialogs failed when used for application permissions. However, a side-by-side comparison under controlled conditions is still needed.

Competence

While we did not specifically determine whether participants were competent enough to make the decisions needed to complete the tasks, we measured whether their observed behaviors matched their stated privacy preferences. In the exit survey, we asked participants questions in order to classify them within the Westin Index of privacy preferences (i.e., "privacy fundamentalists," "privacy pragmatists," or "privacy unconcerned") [29]. We believe that the dialogs met the competence criterion because we observed a significant correlation between participants' Westin Index classifications and whether they chose to use Facebook Connect ($r = -.360$, $p < 0.003$; Spearman correlation); "privacy fundamentalists" were significantly more likely to opt out. Thus, the dialogs allowed participants to act competently: those who had privacy concerns were able to avoid making disclosures.

Agreement

The agreement principle states that users should be given opportunity to reconsider their decisions. Within the context of Facebook Connect, this corresponds to the ability to revoke websites' access to Facebook profile data. While Facebook provides users this ability, explicitly testing whether or not they knew how to use it was beyond the scope of our experiment. At the same time, two participants, unprompted, volunteered that they knew they could revoke their decisions later: *"I can choose later to disconnect myself from the site."* The proportion of participants who understood this is a subject for future work.

Design Implications

We believe that during our experiment, informed consent was largely achieved: in general, participants who were comfortable making disclosures used the system, whereas those who were not completed the task using other means. At the same time, our data point to a potential shortcoming of the Facebook Connect system: users are incorrectly viewing the consent dialogs as static warnings, mistakenly believing that they all communicate similar data collection policies. It is unclear whether users of other OAuth-based systems also believe this. Therefore, designers need to improve these dialogs so that users understand that the terms of a data collection policy may drastically change from one website to another.

We observed that participants made privacy decisions based on a coarse understanding of the types of data that websites might collect; in most cases, participants were either correct or cynically believed that data recipients were collecting more information than in reality. However, this level of understanding may pose problems when users incorrectly believe that websites are collecting less data than they actually are. Future designs should address this issue by examining user expectations and then doing a much better job of highlighting situations that are likely to diverge from these expectations. For instance, when a website goes beyond collecting "basic info," icons or colored text could be used to draw attention to these additional items. Future studies are needed to determine the most effective techniques for overcoming habituation when users encounter "unexpected" privacy policies, as well as to establish how often these situations arise.

ACKNOWLEDGMENTS

Thanks to Christopher Thompson, Rowilma del Castillo, Miho Tanaka, Stephanie Peña, Ashley Lopez, and Thant Hein for help conducting the experiment; Adrienne Porter Felt, Maritza Johnson, Heather Lipford, and David Wagner for feedback. This work was supported by the Intel Science and Technology Center for Secure Computing.

REFERENCES

1. OAuth. `http://oauth.net/`. Accessed: September 17, 2012.
2. V. Bellotti and A. Sellen. Design for privacy in ubiquitous computing environments. In *Proceedings of the third conference on European Conference on Computer-Supported Cooperative Work*, pages 77–92, Norwell, MA, USA, 1993. Kluwer Academic Publishers.
3. A. Besmer and H. R. Lipford. Users' (mis)conceptions of social applications. In *Proceedings of Graphics Interface 2010*, GI '10, pages 63–70, Toronto, Ont., Canada, Canada, 2010. Canadian Information Processing Society.
4. A. Besmer, H. R. Lipford, M. Shehab, and G. Cheek. Social applications: exploring a more secure framework. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, SOUPS '09, pages 2:1–2:10, New York, NY, USA, 2009. ACM.
5. A. Besmer, J. Watson, and H. R. Lipford. The impact of social navigation on privacy policy configuration. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, pages 7:1–7:10, New York, NY, USA, 2010. ACM.
6. R. Böhme and S. Köpsell. Trained to accept?: a field experiment on consent dialogs. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, pages 2403–2406, New York, NY, USA, 2010. ACM.
7. P. H. Chia, Y. Yamamoto, and N. Asokan. Is this app safe?: a large scale study on application permissions and risk signals. In *Proceedings of the 21st international conference on World Wide Web*, WWW '12, pages 311–320, New York, NY, USA, 2012. ACM.
8. Facebook. Authentication. `http://developers.facebook.com/docs/authentication/`, 2012. Accessed: September 11, 2012.

9. A. Felt and D. Evans. Privacy protection for social networking platforms. In *Workshop onWeb 2.0 Security and Privacy*, 2008.

10. D. Florencio and C. Herley. A large-scale study of web password habits. In *WWW '07: Proceedings of the 16th International Conference on the World Wide Web*, pages 657–666, New York, NY, USA, 2007. ACM Press.

11. B. Friedman, E. Felten, and L. I. Millett. Informed consent online: A conceptual model and design principles. Technical Report UW-CSE-00-12-02, University of Washington, 2000. `ftp://128.95.1.178/tr/2000/12/UW-CSE-00-12-02.pdf`.

12. B. Friedman, D. Howe, and E. Felten. Informed consent in the mozilla browser: Implementing value sensitive design. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS'02)*, page 247, Washington, DC, USA, 2002. IEEE Computer Society.

13. N. S. Good, J. Grossklags, D. K. Mulligan, and J. A. Konstan. Noticing notice: a large-scale experiment on the timing of software license agreements. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 607–616, New York, NY, USA, 2007. ACM.

14. J. Grossklags and N. Good. Empirical studies on software notices to inform policy makers and usability designers. In *Proceedings of the 11th International Conference on Financial cryptography and 1st International conference on Usable Security*, FC'07/USEC'07, pages 341–355, Berlin, Heidelberg, 2007. Springer-Verlag.

15. J. V. Grove. Each month 250 million people use facebook connect on the web. `http://mashable.com/2010/12/08/facebook-connect-stats/`, December 2010. Accessed: September 11, 2012.

16. M. Johnson, S. Egelman, and S. M. Bellovin. Facebook and privacy: it's complicated. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, SOUPS '12, pages 9:1–9:15, New York, NY, USA, 2012. ACM.

17. M. Kay and M. Terry. Textured agreements: re-envisioning electronic consent. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, pages 13:1–13:13, New York, NY, USA, 2010. ACM.

18. J. King, A. Lampinen, and A. 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.

19. M. Langheinrich. Privacy by design - principles of privacy-aware ubiquitous systems. In *Proceedings of the 3rd international conference on Ubiquitous Computing*, UbiComp '01, pages 273–291, London, UK, UK, 2001. Springer-Verlag.

20. S. Lederer, A. K. Dey, and J. Mankoff. A conceptual model and a metaphor of everyday privacy in ubiquitous. Technical report, University of California at Berkeley, Berkeley, CA, USA, 2002.

21. J. Mann. Is facebook connect mark zuckerberg's ace in the hole? `http://beta.fool.com/wealthlift/2012/06/28/facebook-connect-mark-zuckerbergs-ace-hole/6367/`, June 2012. Accessed: September 11, 2012.

22. O. Schneider and A. Garnett. Consentcanvas: automatic texturing for improved readability in end-user license agreements. In *Proceedings of the ACL 2011 Student Session*, HLT-SS '11, pages 41–45, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.

23. M. Shehab, S. Marouf, and C. Hudel. Roauth: recommendation based open authorization. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*, SOUPS '11, pages 11:1–11:12, New York, NY, USA, 2011. ACM.

24. M. Shehab, A. C. Squicciarini, and G.-J. Ahn. Beyond user-to-user access control for online social networks. In *Proceedings of the 10th International Conference on Information and Communications Security*, ICICS '08, pages 174–189, Berlin, Heidelberg, 2008. Springer-Verlag.

25. A. Skelton. Social demographics: Who's using today's biggest networks. `http://mashable.com/2012/03/09/social-media-demographics/`, March 2012. Accessed: September 14, 2012.

26. S.-T. Sun, Y. Boshmaf, K. Hawkey, and K. Beznosov. A billion keys, but few locks: the crisis of web single sign-on. In *Proceedings of the 2010 workshop on New security paradigms*, NSPW '10, pages 61–72, New York, NY, USA, 2010. ACM.

27. S.-T. Sun, E. Pospisil, I. Muslukhov, N. Dindar, K. Hawkey, and K. Beznosov. What makes users refuse web single sign-on?: an empirical investigation of openid. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*, SOUPS '11, pages 4:1–4:20, New York, NY, USA, 2011. ACM.

28. N. Wang, J. Grossklags, and H. Xu. An online experiment of privacy authorization dialogues for social applications. In *CSCW '13: Proceedings of the 2013 ACM Conference on Computer Supported Cooperative Work*. ACM, 2013.

29. A. F. Westin. *E-Commerce & Privacy: What Net Users Want*. Privacy & American Business, Hackensack, NJ, 1998. `http://www.pwcglobal.com/gx/eng/svcs/privacy/images/E-Commerce.pdf`.