Interacting with Computers Advance Access published November 12, 2013

doi:10.1093/iwc/iwt058

Online Disclosure of Personally Identifiable Information with Strangers: Effects of Public and Private Sharing

JAYANT VENKATANATHAN[1,*], VASSILIS KOSTAKOS[2], EVANGELOS KARAPANOS[1] AND JORGE GONÇALVES[2]

[1]*Madeira-ITI, University of Madeira, 9020105 Funchal, Portugal*
[2]*Department of Computer Science and Engineering, University of Oulu, FIN-90570 Oulu, Finland*
[*]*Corresponding author: vjayant@m-iti.org*

Safeguarding personally identifiable information (PII) is crucial because such information is increasingly used to engineer privacy attacks, identity thefts and security breaches. But is it likely that individuals may choose to just share this information with strangers? This study examines how reciprocation can lead to the disclosure of PII between strangers in online social networking. We demonstrate that the widespread use of public, one-to-many, communication channels such as 'wall posts' and profile pages in online social networks poses an exception to the assumption that reciprocation happens on one-to-one channels. We find that individuals not only reciprocate and share PII when the disclosure of such information is private and directed towards them by a stranger, but also when the stranger shares this information through a public channel that is not directed towards anyone in particular. Implications for privacy and design are discussed.

RESEARCH HIGHLIGHTS

- We examine reciprocation in the disclosure of personally identifiable information instead of overall 'self-disclosure'.
- This is in the context of interactions with strangers in online social networks.
- Individuals were more likely to disclose information when reciprocating.
- Reciprocation held irrespective of disclosure channel (i.e. public or private).

Keywords: Social aspects of security and privacy; personally identifiable information; self-disclosure; social networks (social computing); social engineering attacks

Editorial Board Member: Kris Luyten

Received 19 December 2012; Revised 3 October 2013; Accepted 7 October 2013

1. INTRODUCTION

Individuals are increasingly turning to online social networks to draw support in their day-to-day activities and pursuits. These sites range from support groups for smoking cessation (Cobb *et al*., 2010) and weight loss (Hwang *et al*., 2010) to travel and accommodation (Lauterbach *et al*., 2009) and language learning (Harrison and Thomas, 2009). Individuals often tend to disclose information about themselves to each other in these interaction settings. Indeed, mutual disclosure of personal information facilitates the development of trust and bonding between individuals. However, such disclosures can also be potentially drawn and exploited by malicious parties attempting to carry out a social engineering attack.

With the increasing adoption of online social networks, and the increasing sophistication of social engineering attacks, an important research challenge is to develop an understanding of how social mechanisms and norms can be exploited for potential attacks. Such an understanding is crucial for the designers of social networking sites in order to foresee these potential attacks and put design mechanisms in place to prevent them.

In this paper, we focus on a social norm that is crucial for the understanding of information disclosure in a social network

setting—reciprocity. The paper demonstrates that reciprocity is an important factor that can leads to the disclosure of personally identifiable information (PII) in online social networks. While there exists substantial evidence showing that individuals tend to reciprocate the act of sharing information about themselves when in one-to-one situations (e.g. Archer and Berg, 1978; Barak and Gluck-Ofri, 2007), it is not clear whether this holds for online social networks where communication patterns are also one-to-many. Increasingly, however, online social networks facilitate public or 'broadcast' channels via which users disclose information in a one-to-many manner, not directed towards any particular individual. The profile page is an example of such a channel, where the user can disclose information about himself to a large audience. Another example is the 'wall post' where the user can broadcast information publicly or to an audience of friends. Such widespread use of broadcast communication channels in online social networks poses an exception to the assumption that a large body of prior work both in the context of online and face-to-face interactions is based on, which is that these disclosures are made in personal, one-to-one interactions.

Furthermore, while previous work has focused on 'self-disclosure', researchers tend to group together a broad range of information about oneself ranging from inner feelings and thoughts (of fear, vulnerability, etc.) to more mundane and factual information. In the context of online social networks, the reciprocity of PII is of particular interest because it can be used to engineer a privacy attack, identity theft or security breach. Hence, we are interested in the reciprocation of PII in the context of online social network interaction with strangers.

In examining whether a reciprocity norm exists in the disclosure of PII (such as full name, occupation, date of birth, nationality, etc.) in the online space, one might expect that the type of channel (public one-to-many vs. private one-to-one) through which the first person discloses her details can influence the other person's decision to reciprocate that disclosure. Hence, the two main research questions that the study reported in this paper addresses are: (1) Is there a reciprocity norm for the disclosure of PII in the online space? and (2) Does the initial disclosure of such information have to be one-to-one in order for the reciprocity norm to come into effect?

The rest of this paper is organized as follows: We first present an overview of prior work in the fields of social engineering attacks and self-disclosure, leading to the two fundamental research questions outlined above (Sections 2 and 3). We then present a study that we designed and conducted in an online social network in order to answer these research questions. Through the study, we demonstrate how the norm of reciprocity can be exploited by malicious parties to draw PII from unsuspecting users (Sections 4 and 5). We go on to discuss how these results enhance our prior understanding of reciprocation in online social networks and how it can be exploited for social engineering attacks. We discuss the implications for the design of social networking systems that can foresee and prevent such attacks with appropriate mechanisms in place (Section 6). Finally, we outline the limitations of this work and the ground that it sets for future work to explore (Section 7).

2. BACKGROUND

2.1. Phishing, social engineering attacks and PII

A deliberate intrusion into, by unjust means, or exploitation of an individual's information or credentials in an online context is referred to as an attack. As a hypothetical example, in a privacy attack an employer might covertly intrude into a potential employee's online social network in order to draw information for a 'background check', thus causing a violation of her privacy. The term attack can also be used while referring to the methodology used to carry out the intrusion, such as a phishing attack or a social engineering attack. A phishing attack is one in which the attacker attempts to con a victim into divulging personal information using spoofed emails and fraudulent websites. Rather than exploiting bugs in computer software, in a phishing attack the attackers attempt to directly extract sensitive information from a victim by posing as a legitimate source (Downs *et al*., 2007). Direct phishing-related losses to US financial institutions have been estimated at over a billion dollars per year (Emigh, 2005). Thus, phishing poses a significant challenge to online security.

A particularly effective form of phishing, known as spear phishing or *social engineering attacks*, involves personalized messages incorporating elements of context (O'Brien, 2005). Literature on phishing suggests that users are aware that they need to protect their computer from problems like malware, but are less aware of social engineering attacks aimed at eliciting information directly from them (Downs *et al*., 2006).

It has been suggested that as phishers get smarter, future generations of phishing attacks will incorporate more elements of context to become more effective (Jagatic *et al*., 2007). For an attacker to incorporate these elements of context in an attack, an important first step would be to obtain a user's PII. PII can be defined as information that can be used to distinguish or trace an individual's identity either alone or when combined with other information that is linkable to a specific individual (Krishnamurthy and Wills, 2009). This includes information that can by itself uniquely trace an individual's identity such as complete name, social security number or biometric records, or in combination with other data such as date of birth or mother's maiden name (Johnson, 2007). PII research has shown that individual pieces of personal information, when linked together from different sources, can be surprisingly accurate in identifying an individual. For example, a study by Acquisti and Gross (2009) demonstrates that people's social security numbers can be predicted based on other pieces of data such as birth date and birth location. Another well-known result in linking pieces of PII is that most Americans (87%) can be uniquely identified from a birth date, five-digit zip code and gender (Malin, 2005).

Thus, safeguarding PII is crucial because such information can be used to engineer privacy attacks and identity thefts (Moyer and Hamiel, 2008). Moreover, certain websites such as banks require users to enter their date of birth along with account information such as the credit card number as a fallback authentication mechanism when they forget their password (Rabkin, 2008). Hence, such PII can be potentially misused by malicious parties to gain access to users' accounts. Therefore, an important research challenge is to develop an understanding of how social mechanisms can be potentially exploited for attacks that attempt to extract PII from unsuspecting users. With such an understanding, we can foresee these potential attacks and enable the designers of these systems to put mechanisms in place that prevent them.

2.2. Online disclosure and reciprocity

There is a significant body of work on understanding the disclosure of a wide array of information about oneself, ranging from biographical data to more intimate information such as opinions, beliefs and fears (e.g. Archer and Berg, 1978; Collins and Miller, 1994). This research has mostly been clumped together under the term 'self-disclosure'. Self-disclosure has been defined as any personal information that a person communicates to another (Altman and Taylor, 1973; Collins and Miller, 1994) and it builds trust by making the discloser increasingly vulnerable to the other person (Rubin, 1975). Altman and Taylor (1973) categorize self-disclosure into three layers: peripheral (biographical data, age, etc.), intermediate (attitudes, values, opinions, etc.) and core (personal, beliefs, needs, fears and values).

The reciprocity effect (Gouldner, 1960) has been widely reported in the self-disclosure literature (e.g. Archer and Berg, 1978; Collins and Miller, 1994). People seem to give back more, and more intimate information depending on the amount and kind of information received. Further, it has been shown that people who disclose more tend to be liked more and people disclose more to those they initially like (Collins and Miller, 1994). However, the nature of the relationship between individuals is an important factor. For example, the obligation to reciprocate disclosure may be stronger between strangers than between friends (Derlega and Chaikin, 1975, p.50). Self-disclosure has been used as a tactical means to elicit information, such as in police interrogation of suspects (Alison *et al*., 2007).

Online disclosure may not involve certain vulnerabilities associated with offline disclosures, due to the relative anonymity and the ability to control which matters one wishes to reveal (Ben-Ze'ev, 2003). Hence, people seem to disclose more intimate information in Internet relationships (Parks and Floyd, 1996). Joinson and Paine (2007) remark that the relationship between self-disclosure and privacy is paradoxical—privacy is a prerequisite for disclosure, yet the process of disclosure serves to reduce privacy. On examination of prior work such as the above, we can infer that the lack of PII (which is implied by anonymity) facilitates the disclosure of more subjective information such as fear, desire and personal shortcomings in online interactions. Hence, PII is distinctly different from more subjective personal information when it comes to individuals' needs to share such information. Yet, to our knowledge, prior work on self-disclosure and the reciprocity of self-disclosure has largely failed to make an explicit separation of PII in examining the disclosure of information about oneself. This type of disclosure is, effectively, a disclosure of identity.

There is also a body of recent work in HCI examining the extent to which individuals disclose personal information, and the methods and strategies adopted by them to manage these disclosures. An early study of Facebook showed that the majority of users disclosed PII on their profile pages (Gross and Acquisti, 2005). In addition, there is often a discrepancy between people's privacy attitudes towards sharing information and their actual sharing patterns (Acquisti and Gross, 2006; Norberg *et al*., 2007, Reynolds *et al*., 2011). This behaviour has been termed the 'privacy paradox'. For instance, a study revealed a high discrepancy between stated concerns and actual behaviour towards sharing static profile information on Facebook (Acquisti and Gross, 2006). Privacy regulation in social networking sites can be considered a socio-technical activity involving interaction with the technological system and the group context. Individuals' privacy behaviour in such systems involves a mixture of technical and mental strategies. For instance, a technical strategy may involve the use of privacy settings to regulate content distribution to select audiences, such as only friends in the system (Stutzman and Kramer-Duffield, 2010), while research has also shown that considering tie strength can be another strategy for developing rules for disclosure (Wellman and Wortley, 1990). Complex group dynamics also play a role in how individuals share information. For example, individuals who occupy more central positions, in terms of the structure of the social network, tend to reveal more information (Kostakos *et al*., 2011).

A large part of the work in HCI such as the above are in social network sites that are primarily concerned with people who already know each other, and use the Internet as one way of keeping their existing social connections alive (Boyd and Ellison, 2007). While this is not surprising given that social networks such as Facebook are among the most accessed websites, there exist other important and popular social networking services in which, due to their nature and purpose, interactions can occur between strangers and often between different cultures and regions (Harrison and Thomas, 2009). These contexts of online social interaction have been largely unexplored in the HCI literature. The online social network Livemocha, with over 9 million users as of the time of writing this manuscript, is an example of such a social network where interactions are typically between strangers. Thus, what can be termed as self-disclosure in such a context can be very different from that on sites such as Facebook, as self-disclosure is not merely characterized by the information that is shared, but also

by the context of the interaction (Antaki *et al*., 2005). Moreover, sites such as Livemocha, unlike Facebook, have relatively rudimentary mechanisms for managing the level of exposure, ruling out privacy management strategies such as restricting access to only friends (Stutzman and Kramer-Duffield, 2010) or 'narrowcasting' each post only to the audience for which it is intended (Goncalves *et al*., 2013).

The reciprocity effect in 'self-disclosure' has been previously reported in online media. For example, one study observed such reciprocity in online forums where 'self-disclosure' was measured by adding together instances of disclosures of facts, thoughts and feelings about oneself (Barak and Gluck-Ofri, 2007). There has also been recent work on understanding public disclosures. For example, on Facebook, disclosure shared privately is perceived to be more intimate than disclosure shared publicly (Bazarova, 2012a,b). However, no work to our knowledge has specifically examined the reciprocity of PII disclosure, and such reciprocity in the context of broadcast disclosures.

The study described next examines reciprocity in the context of disclosing PII to strangers in online social networks. More specifically, it examines the effect of reciprocity in the disclosure of one's full name and date of birth with strangers in an online social network, both in a one-to-one and one-to-many context.

3. HYPOTHESES

Previous work suggests that people tend to reciprocate the act of disclosing a broad range of information about themselves (Joinson and Paine, 2007). Thus, in the context of online social networks, individuals may be more likely to disclose PII if they do so in reciprocation. This reasoning provides grounds for the first experimental hypothesis:

> H1: Individuals are more likely to reveal PII with a stranger in an online social network when reciprocating.

While previous work has reported the reciprocity effect with respect to a range of social exchanges where the initial disclosure is personal and one-to-one, this does not provide us with any grounds to hypothesize whether reciprocity can come to play when the initial disclosure is public and one-to-many. In other words, if a stranger posts his full name and date of birth on his public profile page, and then requests from another user her full name and date of birth, does this bring into play a norm of reciprocity that makes the user more likely to reveal this information? There are no clear grounds for us to suspect that such a request is as likely to result in compliance as in the case in which the stranger shares personal identification in a one-to-one message directed to the target user. This leads us to the second hypothesis:

> H2: Individuals given PII in a one-to-many interaction are less likely to reveal this information than those who are given this information in a one-to-one interaction.

In other words, H2 hypothesizes that such a reciprocity norm only holds in situations where the initial disclosure is one-to-one and directed to an individual.

4. METHOD

It is methodologically challenging to capture behaviours of users with regard to disclosure of PII in technology-mediated interactions, in a *realistic* manner and setting. Previous work has identified a discrepancy between people's attitudes and stated preference towards sharing information and their actual behaviour (Acquisti and Gross, 2006; Norberg *et al*., 2007). Thus, in order to preserve the authenticity of the setting and the validity of our results, we adopted a method to directly observe users' behaviour, as followed by (Jagatic *et al*., 2007).

Asking participants for informed consent would nullify our experiment. Thus, we opted to obtain implicit consent by giving participants an opportunity to respond (or not) to messages we sent them, and then fully debriefed and rewarded all participants at the end of the study. All participants were rewarded within the context of the social network we study, a community-based language learning website, by offering them help in language learning and providing feedback on their language exercises. An alternative approach would be to ask potential participants for informed consent for a fictional study, and then introduce our experimental stimulus. We felt this was inappropriate in our study because it may affect our results due to participants being suspicious, while at the same time it would involve lying to participants in a public online setting that could impose further stress on them.

4.1. Study design overview

We designed a study in order to test our two hypotheses. The study involved sending a message from an experimental profile to individuals in an online social network, attempting to elicit their full name and date of birth. The online social network chosen, Livemocha, is one in which interactions are typically between strangers. Owing to this, these pieces of information were to an extent privacy sensitive in the context of the social network. These target individuals from whom we attempted to elicit information were allocated to one of three conditions, and the condition determined the manner in which we attempted to elicit this information. In condition A, the experimental profile did not divulge his own full name or date of birth in his messages. This was the control condition. In condition B, the experimental profile divulged his own full name and date of birth in his messages. Hence, condition B served to test hypothesis H1. In condition C, the experimental profile did not divulge his own full name or date of birth in his private messages, but had posted

this information on his public profile page. Hence, condition C was used to test hypothesis H2.

The study was conducted on Livemocha, an online social network for language learning, with over 9 million registered users as of the time of writing this manuscript. For each language listed on the website, there are written exercises that involve writing a small paragraph in that language. A user learning a particular language can complete these exercises, and users who are speakers of that language can provide feedback on these exercises. To encourage participation, the website allows users to become 'friends', send private messages and chat with each other.

Each Livemocha user has a profile page where they can upload a profile picture, write a description and share other details about themselves such as age and location. Most people choose to upload a profile picture. Since most of the social interaction is initiated around the submission and correction of exercises, interactions on Livemocha are often between individuals from different cultures or countries, who typically have not met each other before. All profiles are visible to all users, and there are no detailed privacy mechanisms to obscure parts of one's profile to certain individuals.

Compared with Facebook, Livemocha is a much more 'low-tech' website. It lacks the dynamic interface elements found on Facebook, does not have rich media capabilities or search capabilities, and is particularly tuned to one purpose: learning languages. The benefit of this approach is that profile information and privacy settings are very explicit and easy to understand, unlike in Facebook where users often complain about not being able to understand who can see their information.

While Livemocha does not have complicated privacy mechanisms, like in Facebook, it does have certain mechanisms to help users determine credibility. For each profile, one can see the date of registration, indicating whether a user has just registered or is a seasoned veteran. In addition, users get points as a reward for being active on the site. For instance, users are awarded points when correcting an exercise submitted by another user. The total points are also visible for each profile, thus indicating the extent to which a user is a 'good citizen' on the site.

4.2. Study procedure

Our first step was to crawl 26 000 randomly selected, publicly available profiles on Livemocha, using the profiles' unique identifier as the random seed. Analysis of these data indicated to us that the most popular native language on the website was Portuguese. This led us to decide to target Brazilian Portuguese speakers, as their large presence was expected to speed up data collection. In addition, we found that, for every English speaker learning Portuguese, there were 22 Portuguese speakers learning English. This mismatch between Portuguese and English speakers suggested that if our experimental profiles spoke English, then users with complementary skills and needs are most likely to respond.

Following this initial analysis, we next created experimental profiles in Livemocha that were listed as Indian males who were English speakers. Details such as gender and nationality were identical across all the experimental profiles in order to keep results comparable between them. Each experimental profile submitted a beginner-level written exercise in Brazilian Portuguese that consisted of two simple sentences with a few simple grammatical errors. We designed the exercise submission, with the help of native speakers, to be extremely easy to correct, in order to minimize the effort that the participants would invest in our study. Subsequently, we waited for speakers of Brazilian Portuguese to provide feedback on this exercise.

Once Livemocha users responded to the exercises submitted by our experimental profiles, we sent messages from the respective experimental profile to those users, attempting to elicit their full name and date of birth. This request was made under the pretext of interest in understanding their culture (Table 1). Once users responded to this message, they were briefed about the study being conducted, via a profile belonging to one of the researchers.

Table 1. Messages that were used in the three conditions.

Messages used in condition B	Messages used in conditions A and C
Thanks a lot for correcting my exercise	Thanks a lot for correcting my exercise
I am very interested in learning about Brazilian culture. In my town in India people use their father's name as their surname. So my full name is 'Ashok Mohan' where Ashok is my name and Mohan is my father's name	I am very interested in learning about Brazilian culture. In my town in India people use their father's name as their surname. So if you saw a name like 'Vinay Mohan', Vinay is the guy's name and Mohan is actually his father's name
How is it in Brazil?	How is it in Brazil?
There are some amusing things about Indian culture. For example, I was born on 2 November in 1982, and I am considered lucky because it was the birth anniversary of a god named Krishna. When were you born? Is your date of birth special in any way?	There are some amusing things about Indian culture. For example, I have a friend who was born on 2 November in 1982, and he is considered lucky because it was the birth anniversary of a god named Krishna. When were you born? Is your date of birth special in any way?

Sampath Mukundan

Hello. My name is Sampath Mukundan. I am an Indian. I am 28 years old (born on 23 September 1982). I would like to learn Brazilian Portuguese.

Languages

Speaks:
English | Native
Learning:
Portuguese (Brazil) | Beginner

Personal Information

Gender: Male
City: Srirangam
Country: India
Member since: Sun, Jun 19th 2011

Figure 1. Screenshot of the information provided on the profile page of an experimental profile used in condition C. Experimental profiles used for conditions A and B were identical except that the description field (with full name and date of birth) was blank.

The study ran between February and June 2011. A total of 10 experimental profiles were created (4 for condition C and 6 for conditions A and B together). Participants were allocated to condition A or B based on the alternating order of time at which they provided feedback to the exercise submitted by the experimental profiles used for these conditions. Since condition C required the experimental profile to have additional information in the profile page, the experimental profiles used in this condition were minimally different from those used in conditions A and B (Fig. 1). Each experimental profile was used only once to submit an exercise and subsequently message those users who provided feedback to the exercise. This was done to keep to a minimum the 'activity' level of experimental profiles, as that can introduce changes across profiles. Therefore, all experimental profiles were newly registered and had uniformly low credibility in terms of user points and teacher points awarded by the Livemocha system automatically.

A total of 99 participants provided feedback to the exercises submitted by the experimental profiles and each participant was subsequently messaged. One participant provided feedback to the exercise of two experimental profiles, and these data were discarded. The total participants were 35 (12 male) in condition A, 34 (18 male) in condition B and 30 (11 male) in condition C.

For each participant, we recorded: age, the date of joining Livemocha, gender, 'user points' and 'teacher points' as reported by Livemocha. The user points reflect the extent of the total activity of the user on the website which includes completing lessons, submitting exercises and submitting feedback on other users' exercises. The teacher points the extent of the user's activity on the website in terms of the feedback he or she has provided on others' exercises. Age and gender were optional data that the participants could fill in. Seventy-four out of the 99 participants listed their age (mean 29.25, s.d. 11.5, median 26.5) on their profile page.

A total of 59 (28 male) participants responded to the message from the experimental profile. We refer to these participants as 'respondents'. Forty-three of these respondents had listed their age (mean 28.7, s.d. 10.25, median 26).

5. RESULTS

A binary logistic regression test examining the decision to reply or not to the bait message (sent by the experimental profile) did not result in significance for any of the variables recorded: condition, age, gender, user points, teacher score ($P > 0.05$).

Subsequently, we analysed the effect of various variables on whether respondents revealed information pertaining to both kinds of PII that the experimental profile attempted to elicit from them, i.e. name and date of birth. More specifically, we consider that a participant has disclosed his full name if he mentions it in his message to the experimental profile and this mentioned name consists at least of two distinct parts (i.e. the participant mentions a first name and a last name).

When it comes to date of birth information, certain participants disclosed their birthday to the experimental profile, while certain participants, in addition to their birthday, also mentioned their year of birth. On the other hand, the year of birth of many participants could easily be inferred, given the large number of participants who mentioned their age on their profile page (75 out of 99). Therefore, it is not clear whether those who mentioned only their birthday did so with an intention to hide their year of birth or did so because it was not relevant to the significance of birth dates in Brazilian culture. Hence, for date of birth information, we consider whether a participant disclosed their birthday (not accounting for whether they revealed their year of birth) to the experimental profile.

5.1. Effect of different variables on disclosure

Participants from conditions B and C (52 and 50%, respectively) were equally likely to reveal their full name and birthday, followed by those in condition A (15%) (Fig. 2). The complete results are summarized in Table 2.

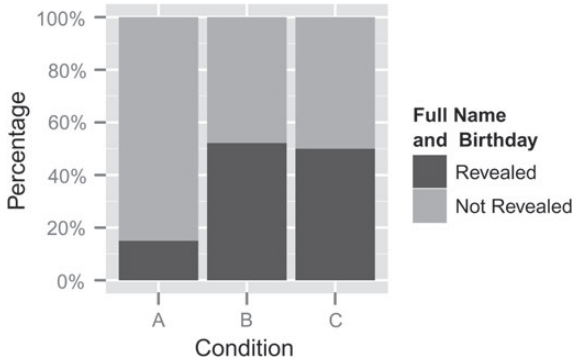

Figure 2. Respondents in conditions B and C were about equally likely to reveal PII, while those in condition A were less likely to do so.

Table 2. Summary of information revealed by users across three conditions.

	Condition A, no disclosure	Condition B, one-to-one	Condition C, one-to-many
Total users	35	34	30
Respondents	20	23	16
Revealed full name	4	16	10
Revealed birthday	8	15	9
Revealed full name and birthday	3	12	8
Revealed full name and DOB, year of birth in message	1	8	3

We conducted a hierarchical logistic regression to analyse the effect of various variables on whether respondents revealed both their full name and birthday. For our response variable in the regression, we gave 'Revealed full name and birthday' responses a value of 1 and 'Did not reveal full name and birthday' a value of 0. Our primary objective was to examine the effect of our experimental manipulation, i.e. the differences between the conditions. Therefore, for our main explanatory variable we used the condition to which respondents belonged.

In addition, we also incorporated the age, gender, time on the website, user points and teacher points as explanatory variables. Incrementally adding blocks of variables to the model allowed us to examine whether the newly incorporated variables provided improved prediction ability over the preceding model. However, given our sample size, we must interpret the results pertaining to these additional variables with caution. The primary objective and contribution of this work is to examine the effect of the experimental manipulation across the three conditions, and further variables are examined only to draw additional insights into disclosure patterns.

Table 3 shows the parameters for the logistic regression and our resulting analytical model of sharing decisions. In the first stage of our model, we examined if the condition in which the participants were allocated affected their decision to reveal their full name and birthday. The results showed that the condition to which the participants belonged offered significant predictive power to our model ($P < 0.01$). The model also included a significant negative constant (intercept) component ($B = -1.73$, $P < 0.01$), indicating that by default our participants did not exhibit an inclination to reveal their full name and birthday unless other variables motivated them to do so. A likelihood ratio test showed that the improvement of this model over the null model was statistically significant ($\chi^2(2) = 8.43$, $P < 0.05$).

Table 3. Details of binary logistic regression modelling the factors involved in the prediction of sharing decisions.

	B (SE)	z-value	$P(>\|z\|)$	exp(B)
Step 1				
Condition (1)	2.41 (1.12)	2.15	0.032	11.13
Condition (2)	2.89 (1.16)	2.50	0.014	17.99
Intercept	−2.89 (1.03)	−2.81	0.005	0.056
Step 2				
Condition (1)	2.59 (1.24)	2.01	0.037	13.33
Condition (2)	3.18 (1.29)	2.46	0.014	24.05
Gender	−1.06 (0.84)	−1.25	0.210	0.346
Age	0.049 (0.040)	1.29	0.219	1.05
Step 3				
Condition (1)	1.76 (0.80)	2.19	0.028	5.81
Condition (2)	2.66 (0.97)	2.73	0.006	14.3
Time on website	0.000 (0.000)	1.145	0.147	1.00
User points	0.000 (0.000)	−0.652	0.515	1.00

In the second stage of our model, we examined if participants' demographics could account for any variation in their choice to reveal their full name and birthday. We found that age and gender did not offer significant influence ($P > 0.05$) within our model, and were hence removed from the subsequent stage.

In the third stage of our model, we examined if respondents' experience of using the website affected their decision to reveal their full name and birthday. The teacher points variable was not included in this equation as it had a high correlation with the user points variable (Pearson's correlation $= 0.88$, $P < 0.001$). The results show that the time since people registered on the website or their user points did not significantly affect their decision to reveal this information ($P > 0.05$). Figure 3 also illustrates that user points had no effect.

Since we did show the age of experimental profiles used in condition C on the profile page, we checked for the effect of this on participants' disclosure. To avoid suspicion, we had varied the age reported on the profile page of these experimental profiles. We reported the age of the four experimental profiles used in this condition as 28, 28, 29 and 21, respectively. The experimental profile with age 21 had the highest rate of respondents who disclosed this information (4 out of 4), but the ages of these respondents greatly varied (17, 35, 60 and 33). However, the effect of the experimental profiles' age on

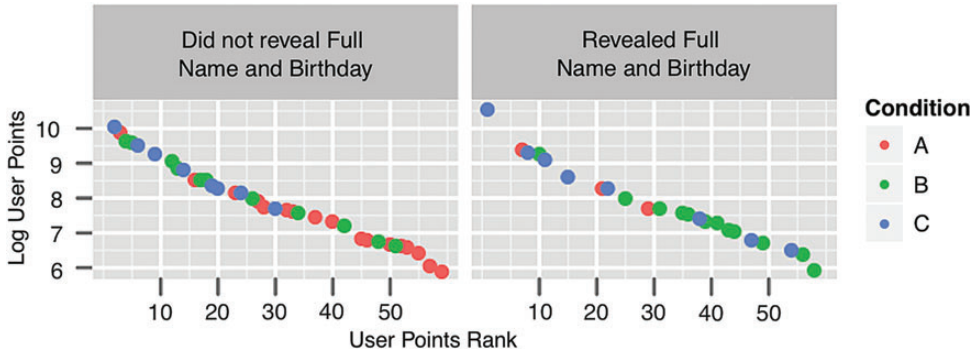

Figure 3. User points had no effect on whether respondents revealed PII or not. On the x-axis is the rank of user points among respondents. The points are spread out across the entire range among both the group of users who revealed this information and those who did not.

whether participants revealed their full name and birthday was not significant ($\chi^2(6) = 8.0$, $P = 0.24$). Profiles in condition B too revealed their age in the private message, but this was constant in all messages sent by experimental profiles in this condition (28 years).

We also examined the effect of various variables on whether participants mentioned their full name and complete date of birth, including year of birth, in their message to the experimental profile. A binary logistic regression showed no statistically significant difference between conditions A and C in this analysis ($P > 0.05$), although participants in condition B were significantly more likely to mention this information ($P < 0.05$). As in the analysis of the disclosure of full name and birthday (not accounting for year of birth), gender, age or experience on the website did not have any significant effect on the likelihood of participants revealing this information.

5.2. Qualitative information in responses and follow-up questions

We received responses in both English and Portuguese across participants. Some participants mentioned that they had used an online translation tool since they were not good at English. One respondent from condition A asked for the birthday of the experimental profile

> I was born on <date removed> ... How about you? When is your birthday?

Some respondents did not divulge their own name but gave examples instead. Instances of these were found in all three conditions. For example, one participant wrote

> Let's imagine my mother's name is <name removed>, and the name of my father is <name removed> ... the child's name might look like <name removed>.

Some participants merely explained how the full name is derived from the mother's and father's names without giving an example. Some responses to mentioning their name and date of birth were brief and to the point, while others were elaborate. While the text in most messages pertained to the explanation of names and the significance of dates of birth, some respondents divulged other details such as interests and employment.

Overall, responses from conditions B and C tended to be longer (mean 130 words (s.d. 90) and mean 115 words (s.d. 72), respectively) than those from condition A (mean 102 words (s.d. 61)).

At the end of the study, all participants (including those who did not respond to the message of the experimental profile) were informed about the study being conducted. We apologized for needing to have communicated with them through an experimental profile, and explained why it was necessary for us to have done that in order to observe responses in a valid manner. As a gesture of appreciation, we offered them help with their English exercises. We were interested in understanding better their behaviour with the experimental profile, and seven users offered give us further feedback through an optional questionnaire. Five of these users had responded to the experimental profile's message while two had not. From this feedback, we learnt that all were active users of other social networks such as Facebook and Orkut, and some had used Internet banking and made online transactions. Thus, this subset of participants were to an extent seasoned users of the Internet.

Overall these participants initially felt that there was some genuineness in the experimental profile's interest in Brazilian culture. They found it interesting for an outsider to be interested in their culture, and wanted to help such a person in learning about it. When asked why they did or did not share their full name or date of birth with the experimental profile, one of the users wrote that she was tricked by the 'complete casualness' of the message into sharing her details. Finally, those who shared any information with the experimental profile reported to have shared accurate information.

6. DISCUSSION

Our results show that users were much more likely to reveal their full name and date of birth when the experimental

profile revealed his own. This suggests that individuals tend to reciprocate the act of sharing PII (more specifically full name and date of birth information), confirming hypothesis H1.

On the other hand, individuals who could see the full name and date of birth information of the experimental profile on his public profile page were more likely to reveal their information than those who were not given this information. Since condition C was identical to A in terms of the message received by the user, the only factor that can explain the significant difference in the disclosures in this condition is that these users subsequently went to the profile page of the experimental profile and saw the additional information posted there. As a result of seeing additional details posted on the profile page, these users were more willing to share their details.

Moreover, there was no difference between conditions B and C when it came to disclosure of full name and birthday, leading us to reject H2. That is, participants were equally willing to reveal this information irrespective of whether the experimental profile shared his information in a private message or in a broadcast manner. This provides evidence that the reciprocity norm implied by H1 also applies to the case where the initial disclosure is one-to-many.

6.1. The norm of reciprocity

This paper set out to address two fundamental questions with regard to the sharing of PII with strangers in an online social network. The first is whether individuals reciprocate the sharing of such information. Our results indicate that the answer to this question is yes. This is in agreement with prior work on 'self-disclosure' taken as a disclosure of a broad range of personal information (e.g. Barak and Gluck-Ofri, 2007).

More surprising, however, is the finding that the reciprocation occurs even when the information is broadcast, such as through a public profile page, where it is not directed at a particular user. This is especially interesting in the light of recent findings that public disclosures on Facebook were perceived less intimate than private disclosures (Bazarova, 2012a,b). Our findings suggest that in stranger interactions, there might be no difference between public and private disclosures of personal identifiable information in terms of willingness to reciprocate such disclosures.

It must be noted that the users who were sent a message had all first provided feedback on an exercise submitted by the experimental profile. This was done in order to increase the rate of response to the messages. In addition, the fact that users might have perceived the experimental profile to be able to help them with learning English might have increased response rates overall. Consequently, the reciprocation that we have observed is over and above these effects. However, since these factors apply equally to all three conditions, the conclusions drawn from our results remain valid.

The simplest interpretation of our results is that the sharing of the full name and date of birth affected the compliance of the recipient when it came to revealing his own full name and date of birth because the recipient felt obligated to reciprocate the act. By sharing PII, an individual communicates a certain degree of trust on the recipient, and it is an unspoken obligation of the recipient to reciprocate this act of trust when required to do so (Derlega and Chaikin, 1975). Hence, the reciprocity of disclosing PII can also be viewed fundamentally as a reciprocity of a display of trust. Interestingly, this display of trust can be towards a group or community of people as a whole and the norm of reciprocity still holds when an explicit request is made to an individual from this group.

6.2. Effects beyond reciprocity

Even though a reciprocity norm is a plausible explanation for the increased compliance observed in our results, one cannot rule out other causes. We next discuss possible alternative explanations for the results we have obtained, and show whether or not they are plausible. While the following list is not meant to be exhaustive, we believe these are the most important alternative factors that can explain the observed results.

Credibility: It can be hypothesized that the act of revealing their full name and date of birth made the experimental profile seem more credible. Therefore, using credibility as a guiding concern (e.g. Andrade *et al*., 2002), respondents showed increased compliance in conditions B and C. However, we argue that if credibility was indeed the guiding concern, all conditions would have observed a low level of compliance. This is due to the fact that their credibility was actually quite low: all profiles were newly created, with extremely low user points and teacher points, indicating a person who is not an active or trusted member of the community. Hence, we can rule out credibility as the guiding concern of respondents, as they all responded to overall low-credibility profiles.

Imitation: Studies have shown that humans have a tendency to imitate the behaviour of others (e.g. Meltzoff and Moore, 1977). Along similar lines, it is plausible that respondents in conditions A and B tended to replicate the behaviour of the experimental profile in their response by hiding or disclosing their full name and date of birth in their message. However, this hypothesis does not account for the behaviour of respondents in condition C. If these respondents were simply imitating, then they should not be more likely to disclose their details than respondents in condition A, since the message they received was identical in both conditions. On the other hand, imitation might partially explain why participants in condition B were more likely to explicitly mention their year of birth in the message, since the experimental profile mentioned his year of birth in the message too. Participants in condition C, however, were possibly less disposed to do so, as year of birth was probably irrelevant to explaining the significance of their birthday, and the experimental profile himself did not mention his own year of birth in his message. Nevertheless, while imitation might possibly explain the difference between conditions B and C in

terms of disclosure of year of birth, it does not fully explain our results.

Liking: Research has shown a link between 'self-disclosure' and liking, which can in turn lead to self-disclosure in return (Collins and Miller, 1994). Here, the motivation for disclosure is not a feeling of obligation. Rather, this explanation posits that because in conditions B and C the experimental profiles shared personal information, respondents felt that they like this profile. As a result, they chose to also share their personal details. While we cannot completely rule out this hypothesis, there is evidence against it. Primarily, all profiles were mostly identical: the nationality, gender and approximate age of the experimental profiles were identical, therefore equally contributing to a respondent's liking of the profile. It is true that in conditions B and C the profile shared a date of birth, which may have had an impact on respondents' liking of the profile. While we cannot rule it out, this explanation asserts that the reciprocity effect we have observed is indirect. In either case, the impact of our experimental manipulation is existent, whether direct or indirect.

Erroneous norm-inference: Visibility of actions allows individuals to observe others' behaviour and infer social norms (Erickson and Kellog, 2000). Thus, the experimental profile publicly revealing his personal information might have suggested to new users that sharing such information is a norm on the website. As a result, respondents in condition C might have been more willing to share this information. However, the data suggest that this is not the case. First, our observations showed that sharing such private information is in fact not a norm on this website. However, one might expect that new users may not be aware of this, and could potentially be 'tricked' into believing this behaviour is a norm. It is also possible that technology savviness and prior experience with the web may have a role to play in this. While we do not have data for technology savviness or overall web usage in our sample, our analysis of experience on the website (time since registration, teacher points, user points) showed no relationship with whether users shared their information. While we caution the reader to interpret with care the results from variables in addition to our experimental conditions, at least within our sample respondents who shared their information were at various levels of experience on the website. This can also be visually verified by Fig. 3—the dots representing users in condition C appear across the range of user point values. Hence, this explanation is unlikely to explain our results.

6.3. Implications for privacy

The experiment described in this paper demonstrates the vulnerability of users against attempts to trick them into revealing information by exploiting this social norm. Inferring or linking personal information such as that obtained in the current study would typically be an important first step in a malicious party's attempt to exploit a user. For example, the malicious party might use elements of context inferred from the site such as the user's interest in learning a language or the people that the individual has friended in the social network. The unsuspecting user might then be sent a spam message or email incorporating these elements of context on his birthday for an advertisement of a language learning product or even a link to a virus. Such context-aware spam messages are known to have higher click-through rates (Brown *et al*., 2008) and are thus likelier to trick the user.

With people increasingly interacting with strangers on various social networking platforms, there is a need for mechanisms to help them identify such attackers apart from genuine users. Exploiting social norms and trust is a well-understood mechanism for social engineering attacks (Jagatic *et al*., 2007). What our results show, however, is that whereas such attacks were targeted in a one-on-one fashion, users are also vulnerable to easier and cheaper one-to-many attacks.

While systems must support the development of ties between individuals, and mutual disclosure of personal information is an integral element of such a bonding process, it is important to distinguish between genuine individuals forming a relationship and malicious parties. The challenge is then to provide mechanisms that help users identify such malicious parties in a manner that does not hinder the sharing of information between genuine users.

In looking for a solution to this problem, we might take inspiration from the theory of social translucence (Erickson and Kellog, 2000). The authors of that work suggest that making certain activity or information visible ('translucence') to relevant individuals can cause those directly involved to act differently. It does so through supporting *mutual awareness* among all individuals ('they know that others know') and this brings our social rules into play and therefore a sense of accountability on the part of those who are acting. Clearly, there is a tension between visibility and privacy, and the goal is not to take away the privacy of the environment but rather to understand that privacy simply supports certain types of behaviour and inhibits others. Drawing from this idea, one solution to protect users from such attacks would be to provide a public communication channel for each profile, similar to the Facebook 'wall'. This allows an individual who is approached by a stranger attempting to elicit private information (under a pretext such as interest in culture, as in our study) to move their interaction to this public space where she can address his supposed interest without divulging in personal information. This provides a certain amount of visibility of the users' interactions to the community. When other users view the stranger's wall, they know that the stranger has been doing this with multiple users and thereby exercise caution in their interactions with him. While such a solution does not eliminate the risk of this kind of attack, it serves as a means for users to support each other and reduce its chances.

Another approach would be to automatically monitor newly created profiles and profiles that have not invested much effort

in the activities of the community (in our case, users with low teacher and user points). If such a user sends messages with similar content to multiple recipients within a short time span, the low credibility level of the user can be highlighted to the recipients, so that they can make an informed decision to exercise caution. However, such a solution must be thoughtfully implemented, as it might result in disproportionate costs for genuine newcomers and thus the community as a whole, since newcomers are crucial for the vitality of online communities (Kraut *et al*., 2010). It is therefore important to keep perspective of genuine forms of interaction so as to ensure that the solution does not inhibit them.

6.4. Relevance to Facebook research

Unlike the most popular social networks such as Facebook that are better explored in HCI, where ties are mostly between individuals who share some offline element (Boyd and Ellison, 2007), in the social network examined in this study interactions are typically between strangers. In Facebook, users connect with people from different aspects of their lives, including family, friends, schoolmates and co-workers. Thus, the issue of context collapse (Boyd, 2008)—the process by which various kinds of individuals' ties become grouped together under generic terms such as 'Friends'—and how users manage the merging of these different contexts is important to understand in sites such as Facebook.

One might argue that in sites such as Livemocha the issue of context collapse is relatively less complex, as users in the social network do not share information with people with different aspects from their lives but rather connect with people for the purpose of language learning and to explore cultures. Nevertheless, reciprocity can be an important factor in information disclosure on Facebook as is on sites such as Livemocha. For example, if a user on Facebook shares her phone number with her friend on a wall post or comment that is visible to a large audience, is her friend likely to feel compelled to share her phone number too on the same thread? While work has examined how users perceive and interpret disclosures (Bazarova, 2012a,b), it would also be interesting to study how they perceive *non-disclosures* such as the refusal to directly reciprocate. For example, if the above friend does not share her phone number in the wall post, or possibly chooses to rather share her number in a private message, how would the user and the audience to which the wall post is visible perceive this? It would be interesting and fruitful for future work to explore how the processes of reciprocity and context collapse operate together, and perhaps contrast this between online and face-to-face social networks (Kostakos and Venkatanathan, 2010).

7. LIMITATIONS AND FUTURE WORK

A methodological drawback of the study is not that participants were chosen at random from the population of users but rather that participants were self-selected. Therefore, we cannot rule out the presence of a non-response bias in our sample, whereby those users who chose not to correct our exercise or reply to our message might have behaved differently from our observed sample. While the implications for privacy remain unchanged, the extent of reciprocity observed in our results might possibly differ from that of a truly random sample. However, this methodology is more valid in our case than those that rely on self-reported data from users. For example, it might be unrealistic to expect to obtain credible data by asking users questions such as 'Would you reveal your complete name and date of birth to a stranger in an online social network?'.

On the other hand, our own and others' recent work in understanding social engineering attacks (e.g. Jagatic *et al*., 2007) has resorted to using a post-consent technique, to directly observe users. While such naturalistic experiments must be executed with caution and avoided where possible, there is an important case for them in understanding online fraud (Jakobsson *et al*., 2008). With the increasing sophistication of social engineering attacks (Jagatic *et al*., 2007) it might be important for researchers to develop and test alternative lab methods, such as role play, to understand online fraud. Nevertheless, because the approach used in this paper offers the most accurate picture, the need to better understand how people interact with computers makes such research worthwhile (Jakobsson *et al*., 2008).

While discussing privacy aspects, a valid question to ask is whether the personal information used in the study is privacy sensitive within the context of which these disclosures took place. One can argue that information such as full name is easily accessible nowadays, as opposed to a credit card number or a social security number, and hence users' perceptions of privacy threats might have been low. On the other hand, the Livemocha social network is largely anonymous where users have never met before and typically share no mutual friends or other social support mechanisms that they can use to socially verify each other's credibility. In such a context, one might expect users to build trust over time and multiple interactions. Hence, it is of concern that such information can be elicited within one or two brief interactions, as was observed in this study. While the findings on reciprocity hold irrespective of the extent to which such information is privacy sensitive within the current context, future work must explore how users behave when it comes to more sensitive information and how far one can take this before the reciprocity effect breaks down.

We also highlight that for the majority of participants we did not verify whether the date of birth they reported was correct. Some users deliberately use fake personal details online to minimize their exposure to fraud, and it is possible that some respondents adopted this strategy when responding to our bait message. It is certainly interesting to investigate further the extent of this behaviour and its consequences.

Finally, we note that our experimental profiles were all Indian males who listed themselves as English speakers and our sample

consisted only of native Brazilians. This was an explicit decision we made to make our experimental profiles more attractive, as these profiles could provide help in English learning. This is likely to have resulted in a potential power imbalance, which might have had an effect on the participants' willingness to respond. Future work can examine the effect of such power imbalances in the context of information disclosure. Further, although it is expected that the reciprocity observed in this study also holds for a more general population of users, clearly there might be differences across cultures in finer details, such as the extent to which such a norm is adhered to. These potential cultural differences set a fertile ground for future work to explore.

ACKNOWLEDGEMENTS

The authors would like to thank Filipe Quintal and Lucas Pereira for their support over the period of the study.

FUNDING

This work was funded by the Portuguese Foundation for Science and Technology (FCT) grant CMU-PT/SE/0028/2008 (Web Security and Privacy). Additional support was provided by the Academy of Finland and TEKES.

REFERENCES

Acquisti, A. and Gross, R. (2006) Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook. In Proceedings of the 6th international conference on Privacy Enhancing Technologies (PET'06), pp. 36–56. Springer-Verlag, Berlin, Heidelberg.

Acquisti, A. and Gross, R. (2009) Predicting Social Security Numbers from Public Data. Proc. Natl Acad. Sci. USA, 106, 10975.

Alison, L., Kebbel, M. and Leung, J. (2007) Facet analysis of police officers' self-reported use of suspect interviewing strategies and their discomfort with ambiguity. Appl. Cogn. Psychol., 22, 468–481.

Altman, I. and Taylor, D. (1973) Social Penetration: The Development of Interpersonal Relationships. Holt, Rinehart, & Winston, New York.

Andrade, E.B., Kaltcheva, V. and Weitz, B. (2002) Self-disclosure on the web: the impact of privacy policy, reward, and company reputation. Adv. Consum. Res., 29, 350–353.

Antaki, C., Barnes, R. and Leudar, I. (2005) Self-disclosure as a situated interactional practice. Br. J. Soc. Psychol., 44, 181–199.

Archer, R.L. and Berg, J.H. (1978) Disclosure reciprocity and it's limits: a reactance analysis. J. Exp. Soc. Psychol., 14, 527–540.

Barak, A. and Gluck-Ofri, O. (2007) Degree and reciprocity of self-disclosure in online forums. Cyber Psychol. Behav., 10.3, 407–417.

Bazarova, N.N. (2012a) Contents and Contexts: Disclosure Perceptions on Facebook. In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12). pp. 369–372. Seattle, Washington, WA, USA.

Bazarova, N.N. (2012b) Public intimacy: disclosure interpretation and social judgments on Facebook. J. Commun., 62, 815–832.

Ben-Ze'ev, A. (2003) Privacy, emotional closeness, and openness in cyberspace. Comput. Hum. Behav., 19, 451–467.

Boyd, D. (2008) Taken Out of Context: American Teen Sociality in Networked Publics. School of Information. University of California, Berkeley, Berkeley, CA.

Boyd, D. and Ellison, N. (2007) Social networks: definition, history, and scholarship. J. Comput.-Mediat. Commun., 13, 210–230.

Brown, G., Howe, T., Ihbe, M., Prakash, A. and Borders, K. (2008) Social Networks and Context-Aware Spam. In Proceedings of the 2008 ACM conference on Computer supported cooperative work (CSCW '08), pp. 403–412. ACM, San Diego, CA, USA.

Cobb, N.K., Graham, A.L. and Abrams, D.B. (2010) Social network structure of a large online community for smoking cessation. Am. J. Public Health., 100, 1282–1289.

Collins, N.L. and Miller, L.C. (1994) Self-disclosure and liking: a meta-analytic review. Psychol. Bull., 116, 457–475.

Derlega, V.J. and Chaikin, A.L. (1975) Sharing Intimacy. What we Reveal to Others and Why? Prentice-Hall, Inc., Englewood Cliffs, NJ.

Downs, J.S., Holbrook, M.B. and Cranor, L.F. (2006) Decision Strategies and Susceptibility to Phishing. In Proceedings of the second symposium on Usable privacy and security (SOUPS '06), pp. 79–90. Pittsburgh, PA, USA.

Downs, J.S., Holbrook, M.B. and Cranor, L.F. (2007) Behavioral response to phishing risk. In APWG 2nd Annual eCrime Researchers Summit, pp. 37–44. Pittsburgh, Pennsylvania, PA, USA.

Downs, J.S., Holbrook, M., & Cranor, L.F. (2007). Behavioral response to phishing risk. In Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit, pp. 37–44. ACM, Pittsburgh, PA, USA.

Emigh, A. (2005) Online Identity Theft: Phishing Technology, Chokepoints and Countermeasures. Identity Theft Technology Council Report. http://www.antiphishing.org/ Phishing-dhs-report.pdf (accessed October 3, 2005).

Erickson, T. and Kellog, W. (2000) Social translucence: an approach to designing systems that support social processes. Trans. Comput. Hum. Interact., 7, 59–83.

Goncalves, J., Kostakos, V. and Venkatanathan, J. (2013) Narrowcasting in Social Media: Effects and Perceptions. Proc. of ASONAM'13, Niagara Falls, Canada. IEEE.

Gouldner, A.W. (1960) The norm of reciprocity: a preliminary statement. Am. Soc. Rev., 25, 161–178.

Gross, R. and Acquisti, A. (2005) Information Revelation and Privacy in Online Social Networks. Workshop on Privacy in the Electronic Society, Alexandria, VA. ACM Press.

Harrison, R. and Thomas, M. (2009). Identity in online communities: social networking sites and language learning. Int. J. Emerg. Technol. Soc., 7, 109–124.

Hwang, K.O., Ottenbacher, A.J., Green, A.P., Cannon-Diehl, M.R., Richardson, O., Bernstam, E.V. and Thomas, E.J. (2010) Social support in an Internet weight loss community. Int. J. Med. Inform., 79, 5–13.

Jagatic, T.N., Johnson, N.A., Jakobsson, M. and Menczer, F. (2007) Social phishing. Commun. ACM 50, 10 (October 2007), 94–100.

Jakobsson, M., Finn, P. and Johnson, N. (2008) Why and how to perform fraud experiments. IEEE Secur. Priv., 6, 66–68.

Johnson III, C. Safeguarding Against and Responding to the Breach of Personally Identifiable Information, May 22 2007. Office of Management and Budget Memorandum.

Joinson, A.N. and Paine, C. (2007) Self-disclosure, Privacy and the Internet. In Joinson, A.N., McKenna, K.Y.A., Postmes, T. and Reips, U. (eds), Oxford Handbook of Internet Psychology, pp. 237–252. Oxford University Press, New York.

Kostakos, V. and Venkatanathan, J. (2010) Making Friends in Life and Online: Equivalence, Micro-Correlation and Value in Spatial and Transpatial Social Networks. SOCIALCOM '10, pp. 587–594. IEEE.

Kostakos, V., Venkatanathan, J., Reynolds, B., Sadeh, N., Toch, E., Shaikh, S.A. and Jones, S. (2011) Who's Your Best Friend? Targeted Privacy Attacks in Location-sharing Social Networks. UbiComp '11: Ubicomp, pp. 177–186. ACM, Beijing, China.

Kraut, R., Burke, M. and Riedl, J. (2010) Dealing with New Comers. In Kraut, R.E. and Resnick, P. (eds), Evidencebased Social Design Mining the Social Sciences to Build Online Communities: 1 42. MIT Press.

Krishnamurthy, B. and Wills, C.E. (2009) On the Leakage of Personally Identifiable Information Via Online Social Networks. ACM SIGCOMM Workshop on Online Social Networks (WOSN).

Lauterbach, D., Truong, H., Shah, T. and Adamic, L. (2009) Surfing a Web of Trust: Reputation and Reciprocity on CouchSurfing.com. 2009 Int. Conf. Computational Science and Engineering, pp. 346–353. IEEE.

Malin, B. (2005) Betrayed by my shadow: learning data identify via trail matching. J. Priv. Technol.

Meltzoff, A.N. and Moore, K.M. (1977) Imitation of facial and manual gestures by human neonates. Science, 198, 75–78.

Moon, Y. (1998) Impression management in computer-based interviews: the effects of input modality, output modality, and distance. Public Opin. Quart., 62.

Moyer, S. and Hamiel, N. (2008) Satan is on my friends list: attacking social networks. http://www.blackhat.com/html/bh-usa-08/bh-usa-08-archive.html.

Norberg, P.A., Daniel, R.H. and David, A.H. (2007) The privacy paradox: personal information disclosure intentions versus behaviors. J. Consum. Aff., 41, 100–126.

O'Brien, T.L. (2005) Gone spear-phishin'. The New York Times (4 December). http://www.nytimes.com/2005/12/04/business/yourmoney/04spear.html?pagewanted=1&ei=5088&en=2f313fc4b55b47bf&ex=1291352400&partner=rssnyt&emc=rss.

Parks, M.R. and Floyd, K. (1996) Making friends in cyberspace. J. Commun., 46, 80–97.

Reynolds, B., Venkatanathan, J., Goncalves, J. and Kostakos, V. (2011) Sharing Ephemeral Information in Online Social Networks: Privacy Perceptions and Behaviours. INTERACT, pp. 204–215. Springer, Berlin.

Rubin, Z. (1975) Disclosing oneself to a stranger: reciprocity and its limits. J. Exp. Soc. Psychol., 11, 233–260.

Stutzman, F. and Kramer-Duffield, J. (2010) Friends Only: Examining a Privacy-Enhancing Behavior in Facebook. Proc. Conf. Human Factors and Computing Systems: CHI 2010, pp. 1553–1562. ACM Press.

Wellman, B. and Wortley, S. (1990) Different strokes from different folks: community ties and social support. Am. J. Sociol., 96, 558–588.