

# Facebook and Privacy: The Balancing Act of Personality, Gender, and Relationship Currency

Daniele Quercia<sup>§</sup> Diego Las Casas<sup>‡</sup> João Paulo Pesce<sup>‡</sup> David Stillwell<sup>†</sup>

Michał Kosinski<sup>†</sup> Virgilio Almeida<sup>‡</sup> Jon Crowcroft<sup>§</sup>

<sup>§</sup>The Computer Laboratory, University of Cambridge

<sup>†</sup>The Psychometrics Centre, University of Cambridge

<sup>‡</sup>Universidade Federal de Minas Gerais, Brazil

daniele.quercia@cl.cam.ac.uk

## Abstract

Social media profiles are telling examples of the everyday need for disclosure and concealment. The balance between concealment and disclosure varies across individuals, and personality traits might partly explain this variability. Experimental findings on the relationship between information disclosure and personality have been so far inconsistent. We thus study this relationship anew with 1,313 Facebook users in the United States using two personality tests: the big five personality test and the self-monitoring test. We model the process of information disclosure in a principled way using Item Response Theory and correlate the resulting user disclosure scores with personality traits. We find a correlation with the trait of Openness and observe gender effects, in that, men and women share equal amount of private information, but men tend to make it more publicly available, well beyond their social circles. Interestingly, geographic (e.g., residence, hometown) and work-related information is used as relationship currency, in that, it is selectively shared with social contacts and is rarely shared with the Facebook community at large.

## 1 Introduction

Social media profiles are extraordinary windows to the self. Concealing and disclosing information is key to identity management. Facebook users might selectively reveal only the aspects of their identities that they believe others should know and, based on what they choose to disclose and conceal, they support specific narratives of the self (Back et al. 2010). Concealment and disclosure are also key to meet a user's individual need for privacy. In theory, striking the right balance between concealment and disclosure takes an extraordinary amount of knowledge and judgment. In (social-networking) practice, this translates into fiddling with privacy settings and applying restrictions on what can be viewed, and by whom.

Social-networking sites can allow their users to fiddle with increasingly more sophisticated privacy settings, but the more sophisticated the settings, the more unusable they become. In Papacharissi's words: "what renders privacy a luxury commodity is that obtaining it implies a level of computer literacy that is inaccessible to most .. As a luxury commodity, the

right to privacy, afforded to those fortunate enough to be Internet-literate becomes a social stratifier; it divides users into classes of haves and have-nots, thus creating a privacy divide." (Papacharissi 2010)

To facilitate the strategic disclosure and concealment of online identities, researchers have tried to understand what privacy actually is in the online context. As we shall see in Section 2, it has been found that privacy and what is considered private or public are ideas and practices that not only differ across cultures but also differ across individuals, and these differences depend on a variety of factors such as age, gender, and Internet literacy (Boyd and Hargittai 2010). Some researchers have suggested that personality traits might explain the variability of how individuals define "private" and "public" (Amichai-Hamburger and Vinitzky 2010; Ross et al. 2009). This suggestion has been tested on small user studies and no consistent findings have emerged. That is why we set out to study the relationship between personality traits and information disclosure and concealment. More specifically, we make the following contributions:

- We model the way 1,323 Facebook users disclose and conceal information on their profiles, and we do so by adapting the well-established technique of "Item Response Theory" to the problem at hand (Section 4). As a result, we are able to associate a *disclosure score* with each user, and to quantify which user profile fields are selectively concealed and which are widely disclosed. We find that our Facebook users tend to selectively conceal work-related information, while they freely disclose whether they are looking for a partner, their education levels, states of residence, and political and religious views.
- We study the relationship between information exposure scores and personality traits (Section 5). We find that Openness and, to a lesser extent, Extraversion are weakly correlated with disclosure attitudes.
- We finally discuss the theoretical implications of our findings on personality research, and the practical implications on designing social-networking privacy tools and social media marketing campaigns (Section 6).

## 2 Existing Studies: Multifaceted Privacy

Individual differences has been found to greatly impact information disclosure and perceptions of privacy. The traditional

division of users according to their privacy attitudes was proposed by Hofstede (Hofstede 1996): users were classified into two extreme groups, which were reminiscent of the classic sociological concepts of individualist and collectivist societies. More recently, ethnographic studies have explored privacy attitudes of social-networking users, especially of Facebook ones. After surveying the same group of students twice, once in 2009 and later in 2010, Boyd *et al.* revealed that users' confidence in changing their privacy settings strictly depended on frequency of site use and Internet literacy (Boyd and Hargittai 2010). Also, gender and cultural differences seem to affect privacy attitudes. Lewis *et al.* found that women's profiles are more likely than men's to be private and that there is a relationship between music taste and degree of disclosure of one's profile (Lewis, Kaufman, and Christakis 2008). Finally, ethnicity plays a role. Chang *et al.* studied the privacy attitudes of U.S. Facebook users of different ethnicities and found that Hispanics tend to share more pictures, Asians more videos, and Afro-Americans more status updates (Chang *et al.* 2010).

### 3 Personality and Privacy

Previous research suggests that the extent to which one discloses personal information partly depends on one's personality traits. In particular, it has been found to depend on the big five personality traits and the self-monitoring trait, all of which are discussed next.

The five-factor model of personality, or the big five, consists of a comprehensive and reliable set of personality concepts (Costa and McCrae 2005; Goldberg *et al.* 2006). The idea is that an individual can be associated with five scores that correspond to five main personality traits. Personality traits predict a number of real-world behaviors. They, for example, are strong predictors of how marriages turn out: if one of the partner is high in Neuroticism, then divorce is more likely (Nettle 2007). Research has consistently shown that people's scores are stable over time (they do not depend on quirks and accidents of mood) and correlate well with how others close to them (e.g., friends) see them (Nettle 2007).

The relationship between use of Facebook and the big five personality traits has been widely studied, yet contrasting results have emerged. To see why, we will now report the results for each of the traits:

The trait of *Extraversion* is associated with descriptive terms such as sociability, activity, and excitement seeking. Individuals higher in Extraversion tend to prefer offline interactions and are thus less likely to join social-networking sites. However, if they join Facebook, they tend to "befriend" more social contacts (Golbeck, Robles, and Turner 2011) and join more groups (Ross *et al.* 2009). However, they are less likely to disclose personal information (Amichai-Hamburger and Vinitzky 2010).

The trait of *Neuroticism* is associated with descriptive terms such as emotional lability and impulsiveness. Individuals high in Neuroticism engage in a variety of behaviors across different media: they are more likely to use the Internet to avoid loneliness, to post accurate personal information in anonymous online forums (e.g., in chat rooms),

to control what information is shared when using their mobile phones (Butt and Phillips 2008), and to use Facebook Walls (their favorite feature) while being on the site (Ross *et al.* 2009). However, if one considers the process of sharing in Facebook more carefully, one is bound to come across contrasting results. Ross *et al.* reported that those high in Neuroticism are less likely to post photos and share personal information (Butt and Phillips 2008). Less than two years later, Amichai-Hamburger *et al.* reported exactly the opposite (Amichai-Hamburger and Vinitzky 2010).

The trait of *Agreeableness* is associated with descriptive terms such as trusting, altruistic and tender-minded. On Facebook, Ross *et al.* found no relation at all between Agreeableness and the use of the site (Ross *et al.* 2009), while Amichai-Hamburger *et al.* found that both women low in Agreeableness and men moderate (neither low nor high) in Agreeableness are less likely to share pictures (Amichai-Hamburger and Vinitzky 2010).

The trait of *Conscientiousness* is associated with descriptive terms such as ambitious, resourceful and persistent. Those high in Conscientiousness are less likely to use the Internet and, on Facebook, are less likely to upload pictures. These individuals generally see the use of computer-mediated forms of communication as procrastination or distraction from daily tasks (Amichai-Hamburger and Vinitzky 2010).

The trait of *Openness* is associated with descriptive terms such as imaginative, spontaneous, and adventurous. Those high in Openness are more likely to try new methods of communication, including social-networking sites (Ross *et al.* 2009), and, on Facebook, they have been reported to use a greater number of features (Ross *et al.* 2009) and share more personal information (Amichai-Hamburger and Vinitzky 2010).

To recap, the trait of Openness has been found to be associated with higher information disclosure, while Conscientiousness has been associated with cautious disclosure. Contrasting findings have emerged from the traits of Extraversion, Agreeableness, and Neuroticism. Schrammel *et al.* wrote: "personality traits do not seem to have any predictive power on the disclosure of information in online communities" (Schrammel, Köffel, and Tscheligi 2009). In the same vein, Ross *et al.* concluded that "personality traits were not as influential as expected, and most of the predictions made from previous findings did not materialise" (Ross *et al.* 2009). These researchers, however, conceded that they only considered the big five personality traits (while other traits might better explain the use of Facebook) and that their samples consisted of only undergrad students. Current studies might tackle these two limitations by, for example: 1) considering personality traits other than the big five; and 2) having participants who are not only undergrad students, who might well be more Internet savvy than older people. Next, we will describe how we partly tackle the first problem by considering the personality trait of "self-monitoring". In Section 5, we will describe how we tackle the second problem by collecting a representative sample of Facebook users in the United States.

### 3.1 Self-Monitoring

In 1974, Mark Snyder realized that some people’s personalities are more ‘fluid’ than others’. After conducting in-depth interviews, he found that some people are more prone to recognize how they are perceived by others and accordingly adjust the way they act. He called such individuals ‘high self-monitors’ (Snyder 1974). High self-monitors tend to modify their behavior (and self-presentation) to fit the situations and the people in them, while low-self monitors do not tend to alter their behavior very much across situations. Successful politicians, for example, tend to be high self-monitors (Mehra, Kilduff, and Brass 2000). When taking the self-monitoring test, high self-monitors tend to agree with statements like “I can make impromptu speeches even on topics about which I have almost no information”. By contrast, low self-monitors tend to agree with statements such as “I find it hard to imitate the behavior of other people” and “I have trouble changing behavior to suit different people and different situations”.

The self-monitoring trait impacts not only offline interactions but also online ones. In 2008, Lin studied the personal pages created by users of web portals (Lin 2008). He found that high self-monitors would display limited and generic information on their pages to represent themselves in likable ways, while low self-monitors would display more personal and in-depth information to accurately portray themselves.

More recently, Gogolinski posited that the same would hold on Facebook. She carried out a study among 134 college students who were active Facebook users (81% females) to test the extent to which different degrees of self-monitoring would affect what information is displayed on profiles (Gogolinski 2010). Since Facebook profiles are forms of self-expression, self-monitoring users would be expected to gauge what is appropriate or inappropriate to display on their profiles and, in so doing, they would represent themselves in likable ways. She concluded that her results confirmed just that: high self-monitors had less detailed and more cautious pages to ensure more agreeable profiles, while low self-monitors preferred more detailed profiles.

## 4 Modeling Privacy Attitudes

Having identified the personality traits that have been associated with information disclosure in the literature, we now need to model the process of disclosure itself. By doing so, we will be able to quantify a user’s disposition to disclose her/his personal information. One measure of information disclosure is the count of fields (e.g., birthday, hometown, religion) a user shares on her/his profile. The problem of this approach is that some information fields are more revealing than others, and it is questionable to assign arbitrary (relevance) weights to those fields. An alternative way is to resort to a psychometric technique called Item Response Theory (*IRT*), which has already been used for modeling information disclosure in social media (Liu and Terzi 2009). Next, we detail what *IRT* is, why choosing it, and when and how it works.

**What IRT is.** *IRT* is a psychometric technique used to design tests and build evaluation scales for those tests. It extracts patterns from participants’ responses and then creates a math-

ematical model upon the extracted patterns, and this model can then be used to access an estimate of user attitude (de Bruin and Buchner 2010) (MacIntosh 1998) (Vishwanath 2006). To ease illustration, say that in our case we have a binary *user-by-field* matrix that reflects whether user  $j$  has disclosed field  $i$ . By applying a 2-parameter *IRT* to the matrix, one obtains two parameters that characterize field  $i$  (*difficulty*  $\alpha_i$  and *discrimination power*  $\beta_i$ ), and a third variable derived from the two parameters that characterizes user  $j$  (the user’s *disclosure* attitude  $\theta_j$ ). Then, to consider not only the amount of personal information users have disclosed but also the extent to which they have made that information visible, we build two models upon two information sources that entail different levels of visibility:

*Community Privacy Model.* This model is built upon what users share with the entire Facebook community.

*Social Circle Privacy Model.* This model is built upon what users share only with their social contacts and *not* with the community at large, and it reflects the extent to which users fiddle with their privacy settings and, as a result, share personal information only with their contacts and *not* with the public at large.

**Why IRT.** The most desirable property of *IRT* is its group invariance. This means that the extent to which a user-specified field is sensitive and possesses discriminative power (i.e.,  $\alpha_i$  and  $\beta_i$  scores) does not hold only for the users under analysis but is applicable to any user (potentially to users of a variety of social-networking platforms, assuming that social norms will not greatly differ): one can predict the disclosure score of *any* user based on the  $\alpha_i$  and  $\beta_i$  scores corresponding to the fields the user has disclosed and concealed.

**When IRT works.** The *IRT* model fits the experimental data but, in so doing, it makes two assumptions: 1) a factor called *latent trait* accounts for the covariances between user-specified fields and is assumed to be linearly related to the observed responses; and 2) the model’s parameters best describe any experimental data at hand. In our case, this translates into assuming the existence of underlying behaviors (attitudes) with which users disclose and conceal personal information on their profiles.

**How IRT works.** The basic random variable of *IRT* is the probability that user  $j$  discloses field  $f_i$ . Given that  $f_i$  is ‘0’ (field  $f_i$  undisclosed) or ‘1’ ( $f_i$  disclosed), this probability can be translated as:

$$P_{ij} = \frac{1}{1 + e^{-\alpha_i(\theta_j - \beta_i)}} \quad (1)$$

The model consists of two parameters ( $\alpha_i, \beta_i$ ) from which two steps can derive a third variable  $\theta_j$ :

*Step 1. Estimate the item parameters  $\alpha$  and  $\beta$ .* This is commonly done by ‘Maximum Likelihood Estimation’ (*MLE*). Since there are different versions of *MLE*, we select the hybrid ‘Expectation-Maximization/BFGS’ parameter estimation (as it is the most widely used) and compute  $\alpha$  and  $\beta$  with it. Initially, arbitrary values for  $\theta$  are set.

*Step 2. Estimate the user’s disclosure score  $\theta$ .* One way to estimate this variable would be to use again *MLE*, as Liu

and Terzi did (Liu and Terzi 2009). However, to estimate  $\theta$ , researchers have often found ‘Expected A Posteriori’ (EAP) to produce more accurate results (Yang 2006), and we will thus use it instead of MLE.

Upon termination of these two steps, the model is able to characterize each user-specified field  $i$  with variables  $\alpha_i$  and  $\beta_i$ , and each user  $j$  with  $\theta_j$ . By ‘characterize’ items and users, we mean that:  $\alpha_i$  reflects field  $i$ ’s *discrimination power* and accounts for the fact that not all private Facebook profile fields are equally private;  $\beta_i$  reflects field  $i$ ’s *sensitivity* - fields with a high  $\beta$  tend to be disclosed by users who are willing to expose themselves; and  $\theta_j$  reflects user’s  $j$  willingness to be exposed and we thus call it *user disclosure score*.

To measure the extent to which users make sensitive information visible, we will compute disclosure scores for the two privacy models (community and social circle). To then measure the extent to which users with different personality traits disclose sensitive information differently, we will correlate disclosure scores with personality data.

## 5 Personality and Information Disclosure

We will now break our discussion down into four parts (subsections) and detail: 1) how we collect our Facebook data; 2) how we study the relationship between information disclosure and personality; 3) whether we are able to predict user disclosure scores with personality; and 4) which Facebook profile fields tend to be more private than others.

### 5.1 Dataset

In Section 3, we have reviewed the literature on the relationship between personality and disposition of disclosing personal information. We have found contrasting results, and that is largely because participants in previous studies were often few hundreds undergraduates (Henrich, Heine, and Norenzayan 2010).

To fix this problem, a new Facebook application called myPersonality tries to collect personality scores of large samples of users whose behavioral measures are recorded directly as the users are on the site. Users who have installed the application have been able to take a variety of genuine personality and ability tests. Users are not paid to install the application and are solely motivated by the prospect of receiving reliable personality test results. The application ensures high test result validity by removing the protocols that may be a product of inattentive, language incompetent, or randomly responding individuals. The resulting quality of the responses is high: the scales’ reliabilities are on average higher than reported in test manuals<sup>1</sup> and the discriminant validity (average  $r = .16$ ) is better than those obtained using traditional samples (average  $r = .20$  (John and Srivastava 1999)). myPersonality users can give their consent to share their personality scores and profile information, and around 40% of them choose to do so. That allows us to gather the profiles fields users share with their own contacts (with their social circles). To then gather what users share with the Facebook community at large, we run a web crawler and collect the fields users publicly share with everyone. Critics might rightly say that users

<sup>1</sup><http://www.mypersonality.org/wiki/>

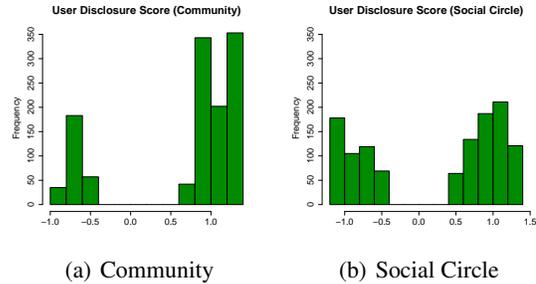


Figure 1: Disclosure scores computed upon profile fields that are made visible (a) to the Facebook community at large; or (b) to one’s social circle (i.e., friends) only.

Variable	User Disclosure Score	
	Community	Social Circle
O	<b>0.14</b>	<b>0.10</b>
C	-0.01	0.04
E	<b>0.05</b>	<b>0.05</b>
A	-0.03	-0.02
N	-0.03	-0.02
Self-Monitoring	<b>0.10</b>	<b>0.07</b>
Male	<b>0.15</b>	0.03
Contacts ( <i>log</i> )	<b>0.14</b>	<b>0.10</b>
Age ( <i>log</i> )	<b>-0.12</b>	<b>-0.08</b>

Table 1: Pearson correlation coefficients  $r$  of the disclosure scores ( $\theta_j$  values) computed upon information shared either with Facebook users at large (first column) or exclusively with Facebook ‘friends’ (second column). Highlighted are those results that are statistically significant (i.e., p-values are  $p < 0.05$  at most).

who take the personality tests but choose not to disclose their test results publicly are more privacy sensitive. That might well be the case, and we have partly addressed this concern in two ways. First, the data sharing consent form clearly states that personal data will be anonymized before research is carried out on it. Second, we compare the participants who disclosed their personal information and those who did not. Statistical tests indicate that there is not any difference in the two groups’ distributions of personality traits and sharing of information with the public at large. Also, as for age and gender distributions, our participants do not differ from typical American Facebook users<sup>2</sup>.

We take a sample of 1,313 Facebook users who live in United States and have taken the big five and the self-monitoring tests. Their number of social contacts is between 32 and 998, age range is between 18 and 60, and median age is 24. They are 701 women (58%) and 514 men (42%).

### 5.2 Information Disclosure and Personality

To begin with, we study the relationship between personality traits and information disclosure. The big five traits are all

<sup>2</sup><http://www.facebook.com/press/info.php?statistics>

normally distributed, as one would expect. Interestingly, the frequency distributions of disclosure scores for both privacy models are *bimodal* (Figure 1). Disclosure scores are not centered on the mean, and a large disparity exists between the scores for privacy-conscious users and those in the pragmatic majority. This result differs from those reported in privacy studies that have found not two but three (clusters) types of users. The seminal work in this area was carried out by Westin in 1991 (Westin 1991). After administering privacy surveys and analyzing the results, he identified three groups of respondents: privacy fundamentalists (minority), the pragmatic majority, and marginally concerned. Since that time, Westin has been creating several privacy indexes from surveys and has studied how those indexes have changed over time. Year after year, he noticed a steady decrease for the fraction of marginally concerned (from 18% in 1991 to 8% in 2001) and commented: “what this documents is something that makes good sense in terms of what we see happening all around us - that unconcern about privacy among the public has dropped ... But it also suggests that privacy fundamentalism is not increasing.” A similar trend has been observed on Facebook. In 2005, Gross and Acquisti found that “limiting privacy preferences are hardly used; only a small number of members change the default privacy preferences, which are set to maximize the visibility of users profiles” (Gross and Acquisti 2005). Five years later, boyd and Hargittai observed that “modifications to privacy settings have increased during a year in which Facebook’s approach to privacy was hotly contested” (Boyd and Hargittai 2010). In a similar way, upon our sample, our methodology has identified two types of users: those who are privacy-concerned (minority) and those who belong to the pragmatic majority.

Next, we study the Pearson product-moment correlation between user  $j$ ’s disclosure score  $\theta_j$  and the user’s five personality scores, plus three additional attributes, namely sex, number of social contacts, and age. We considered the logarithms of the last two attributes because their distributions are skewed. Pearson’s correlation  $r \in [-1, 1]$  is a measure of the linear relationship between two random variables. Table 1 summarizes the results, which are consistent with preliminary findings reported across multiple studies in the literature. Weak correlations are found with Openness for both privacy models (0.14 and 0.10) and with the self-monitoring trait (0.10 and 0.07), and very weak with Extraversion (0.05). These correlations are moderate and weak correlations are also found for the logarithms of: number of contacts (0.14 and 0.10) and age (0.12 and -0.08). This suggests that older and not-so-popular users tend to disclose significantly fewer privacy-sensitive fields than what younger and popular individuals do. One interesting finding is that women’s *public* sharing tends to be more cautious: men and women equally share private-sensitive information with their Facebook social contacts, but women tend to share less privacy-sensitive information with the Facebook community at large (when sharing publicly, the contribution of being a man translates into a correlation of 0.15).

Variable	User Disclosure Score	
	Community	Social Circle
O	<b>0.15</b>	<b>0.13</b>
C	-0.02	0.05
E	0.01	0.04
A	<b>-0.07</b>	-0.07
N	-0.03	0.00
Self-Monitoring	<b>0.01</b>	0.01
Contacts ( <i>log</i> )	<b>0.01</b>	<b>0.00</b>
Age ( <i>log</i> )	<b>-0.01</b>	<b>-0.01</b>
Male	<b>0.22</b>	0.04

Table 2:  $\gamma$  coefficients for the linear regression between the variables listed above and the disclosure scores computed upon information shared either with Facebook users at large or exclusively with Facebook ‘friends’. Highlighted are those results that are statistically significant (i.e., p-values are at least  $p < 0.05$ ).

### 5.3 Predicting Information Disclosure with Personality

So far we have measured how each personality trait is *independently* related to information disclosure. To control for interaction effects among traits and between traits and other variables such age, sex, and number of contacts, we build a regression model that predicts a user’s disclosure score based on personality variables. That is, we model the user  $j$ ’s disclosure score  $\theta_j$  as a linear combination of the big five personality scores, self-monitoring score, plus sex, and the logarithms of number of contacts and age. Our regression coefficients are reported in Table 2. From them, we learn that, after controlling for all the remaining factors, Openness is still a significant predictor (0.15), while no contribution is offered now by self-monitoring, number of social contacts, and age. That is largely because both self-monitoring and number of social contacts have a significant interaction effect with Extraversion. Also, there is no gender effect when sharing privately with friends; whereas, there is a gender effect when sharing publicly with everyone: men are more likely than women to publicly share privacy-sensitive fields (0.22).

The extent to which the regression predicts  $\theta_j$  is reflected in a measure called  $R^2$  - the higher  $R^2$ , the better the fit of the model. In our case,  $R^2$  are 0.07 (community) and 0.03 (social circle), which are both very low. However, the distributions of disclosure scores are bimodal (Figure 1), suggesting the existence of two types of users who might be called, in line with the literature, privacy-conscious and pragmatic majority. Now, if we classify each user as being either of the two types, our classification problem becomes a binary one: we just need to predict which users are privacy conscious and which belong to the pragmatic majority, and we need to do so based on their personality scores. To this end, we consider a Naive Bayes classifier and perform a 10-fold cross-validation. We obtain that 62% of the users are correctly classified in their privacy categories by the model. The problem is that a baseline that classifies everyone as a “pragmatic majority” performs just as well as it would correctly classify 100% of

Field Name	Field Sensitive Score	
	Community	Social Circle
looking	1.4	1.3
education	1.8	1.7
<b>residence</b>	<b>1.9</b>	<b>1.4</b>
political	2.1	2.1
religion	2.2	2.2
<b>hometown</b>	<b>2.3</b>	<b>1.6</b>
<b>position</b>	<b>8.4</b>	<b>7.7</b>
<b>employer</b>	<b>9.9</b>	<b>8.1</b>

Table 3: Sensitive scores for user-specified fields ( $\beta$  values). Highlighted are those fields (and corresponding values) whose sensitive scores change depending on whom they are shared with (i.e., rows whose two column values differ) - whether they are shared with the Facebook community at large or with one’s own social contacts. One useful property of *IRT* is that these scores are comparable with one another.

the pragmatic majority and 0% of the privacy-conscious users. The same goes for a baseline that learns the distribution of the two types of users on test data - it would classify 60.4% of the times users as pragmatic and 39.6% of the times as privacy-conscious. The individual results for our classifier slightly differ - it correctly classifies 73.4% of the pragmatic majority and 39.7% of the privacy-conscious users. So personality data explains only a limited part of the variation of disclosure scores, and predicting these scores with personality data is practical only for a specific class of applications.

#### 5.4 Sensitive and Discriminative Fields

After differentiating users depending on their personality traits, we now need to differentiate Facebook profile fields, as not all private fields are equally private.

In her latest book “Islands of Privacy”, sociologist Christena E. Nippert-Eng explored different ways that privacy is understood by a sample group of Chicago residents (Nippert-Eng 2010). She studied how her subjects managed secrets, phone calls, e-mails, the perimeters of their homes, and interactions with neighbors. She looked at information about the self that her subjects offered to, or withhold from, others. One of her aims was to have a concrete understanding of how people thought about what is more private and what is more public. She discovered that “more private” things are those that are personally and emotionally precious; have significant implications for social status; or invite harm if one lost control of them (especially for information whose access might facilitate identity theft). On the other hand, she found that “more public” things are those that we either do not care about or those that are shared regularly with others.

In a similar way, it would be interesting for us to understand which Facebook fields users consider to be “more private” and which “more public”. One of the reasons we chose *IRT* is that it returns  $\beta$  (field sensitivity) values: one such value quantifies the extent to which a given user-specified field is shared widely (or, symmetrically, is shared sparingly) on Facebook. We therefore consider the following user-

specified profile fields: whether a user is looking for a partner (field *looking*), his/her *education* level, state of *residence*, *political* views, *religion*, *hometown* state, and work-related information (*position* and *employer*). We then compute their sensitivity values and report them in Table 3. By looking at the single columns of  $\beta$  values in the table, one can see two distinctive clusters. The first contains “less sensitive” information (fields) with lower  $\beta$ : looking, education, residence, politics, religion, and hometown. The second cluster contains “more sensitive” information with higher  $\beta$ : work position and employer. This suggests that our Facebook users tend to selectively conceal (or do not talk about) work-related information, while they freely disclose whether they are looking for a partner, their education levels, states of residence, and political and religious views. By then considering the two work-related fields and comparing their sensitive scores when sharing them publicly and when sharing them privately with one’s social contacts, these fields become less sensitive when shared among “friends” (the differences between the  $\beta$  values are 0.7 for work position and 1.8 for employer), suggesting that users do share work-related information, but they do so with their Facebook social contacts more freely than they do with the Facebook community at large. The same applies to past and current states of residence (the differences between the  $\beta$  values are 0.5 for residence and 0.7 for hometown): users share their states of residence and hometown states more freely with their social contacts than with the community.

If one looks at these results under the lens of Nippert-Eng’s work (Nippert-Eng 2010), one could speculate that, for our Facebook users, work-related information seems to be both personally meaningful and have significant implications for social status. By contrast, information disclosing whether one is looking for a partner, education levels, political views, and religion is shared regularly with others and is thus seen to be “more public”. This does not mean that Facebook users do not care about their education, political or religious views. It simply suggests that this type of information is the “relationship currency” with which users maintain their Facebook relationships. The concept of “relationship currency” has been introduced by Nippert-Eng to explain how managing privacy translates into managing social relations. She considered one of the most private pieces of information, namely, secrets and observed that offering up secrets to another person or institution is a way of decreasing social distance, while withholding secrets is a way of increasing it. The use of secrets represents a social currency that is traded back and forth and that re-negotiates our relationships. The idea of relationship currency does not apply only to secrets but to any kind of information, as social observer Michael Schrage suggested more than a decade ago. In 1997, when asked by Merrill Lynch to analyze how “new” technologies would transform businesses, Schrage concluded that the shift did not herald an “information revolution” as much as a “relationship revolution” and added: “Whenever you see the word ‘information’ ... substitute the word ‘relationship’ to more fully understand its uses and its consequences” (Schrage 1997).

We should stress, however, that privacy and what is considered private or public are culturally specific ideas and prac-

tices that it would be inappropriate to export these findings to any other social-networking service, let alone to individuals in the offline world. In fact, Facebook norms might differ from another site's: finding work-related information to be sensitive information is a case in point. A large number of LinkedIn users fill out their work information as LinkedIn is a networking site for professionals. Instead, people generally use Facebook for dating and meeting people, and so perhaps work information just does not matter. So our results are best understood within the Facebook context, but, given limited biases in our sample, it is likely that what we have learned about our users may be generalized to the Facebook population, at least in the United States.

## 6 Discussion

We will now discuss the limitations of our dataset and of our study, the novelty of our information disclosure model, and the theoretical and practical implications of this work.

**Limitations of our data.** Critics might rightly put forward one important issue with our data: some users might have responded untruthfully to personality tests. However, validity tests on the personality data suggest that users have responded to the personality questions accurately and honestly, and that is likely because they installed the application motivated primarily by the prospect of taking and receiving feedback from a high quality personality questionnaire. Critics might then add that, since our sample consists of self-selected users who are interested in their personality, these users might be more active than the general Facebook population. We have looked into this matter and found that our users do not seem to be more active than average - we have reported that the number of contacts for the average user in our sample is 124 whereas Facebook reports an average of 130<sup>3</sup>.

**Limitation of our study.** Our study has four main limitations. First, we have computed exposure scores only upon user-specified fields, while we should have ideally considered additional elements (e.g., photos, wall comments, fan page memberships, segmentation of one's social contacts into friend groups). We were not able to do so as we did not have full access to user profiles - we could access only the elements that we have been studying here. So we looked at sharing of profile fields, which is the simplest instance of sharing, making it an ideal starting point. Yet, even by considering only profile fields, we have learned that our users are far from behaving in the same way - very different privacy attitudes emerge (Figure 1). Second, we have considered Facebook users who live in the United States. Since cultural guidelines clearly exist about what is more and less private and public, one should best consider that our results are likely to hold for Facebook users in the United States. We would refrain from generalizing our results to other social-networking platforms (e.g., Twitter) or to any other society - or even to certain subcultures within the United States or within Facebook. To partly tackle this limitation, we are currently studying users of countries other than United States and platforms other than Facebook (i.e., Twitter). Third, information disclosure

and concealment might be likely confounded by Facebook activity and one thus needs to control for it. We did not do so because activity information is not readily available from Facebook's API, and future studies should propose reasonable proxies for Facebook activity<sup>4</sup>. Fourth, information disclosure might be confounded by the general desire to complete one's profile: fields like "interested in" are quick to fill out, whereas "work information" just takes more effort and is not generally relevant to Facebook interactions. Yet, this reflects what the "relationship currency" of Facebook is (Nippert-Eng 2010) - that is, it reflects which fields are used to maintain relationships and which are instead concealed without causing any harm.

**Novelty of Information Disclosure Model.** *IRT* has already been used to model privacy attitudes of social-networking users (Liu and Terzi 2009). Therefore, our methodology closely followed what has been proposed before. However, since *IRT* has never been applied to large-scale social-networking data, we have run into a number of problems. First, algorithms used for estimating *IRT*'s parameters turned out to have a greater impact on result accuracy than what one might expect. We found that best results are achieved if scores are computed in ways different than those proposed in (Liu and Terzi 2009) and similar to those proposed in psychometrics research (Baker 2001). Second, for a large number of users, errors associated with their  $\theta$  exposure scores are unacceptably high. We had to filter those users out, yet we worked with a sample of 1,323 users. However, for studies starting off with a smaller initial sample, this problem should be addressed.

**Theoretical Implications.** Studies that have attempted to understand how individuals conceptualize privacy offline and online have often resorted to in-depth interviews. This methodology has enabled scholars to better ground their work on exactly what people *say* about privacy in their daily experiences. We have proposed the use of a complementary methodology: we have studied what people *do* on Facebook by modeling the work of disclosure and concealment which their profiles reflect. As a result, we have quantified the extent to which personality traits affect information disclosure and concealment, and we have done so upon large-scale Facebook data that has been collected *unobtrusively* as the users were on the site. We have found that Openness is correlated, albeit weakly, with the amount of personal information one discloses. By contrast, after controlling for Extraversion, the self-monitoring trait's contribution disappears, suggesting that high self-monitors might present themselves in likable ways, as the literature would suggest, but they do not tend to share more private information or to make that information more visible. In addition to differentiating users based on their personality traits, we have also found important differences among profile fields and quantified the extent to which not all private fields are equally private: for example, work-related (job status) information is perceived to be more private than information on whether one is looking for a partner.

<sup>3</sup><http://www.facebook.com/press/info.php?statistics>

<sup>4</sup>In our analysis, we did control for number of Facebook social contacts, which might be a reasonable yet biased proxy for activity.

**Practical Implications.** There are two areas in which our findings could be practically applied in the short term. The first is social media marketing. Marketing research has previously found that individuals high in Openness tend to be innovators and are more likely to influence others (Amichai-Hamburger and Vinitzky 2010). Since we have found that these individuals also tend to have less restrictive privacy settings, social media marketing campaigns would be able to identify influentials by determining which users have less restrictive privacy settings. The second area is privacy protection. Our results suggest that, by simply using age and gender, one could offer a preliminary way of personalizing default privacy settings, which users could then change. One could also imagine building privacy-protecting tools that inform users about the extent to which they are exposing information that is generally considered to be sensitive by the community. This tool could also exploit the independence assumption behind *IRT* to parallelize the estimation of the model's parameters and thus be able to monitor who is exposing privacy sensitive information in real-time.

## 7 Conclusion

As the desire to track users continues to outstrip privacy features in social-networking sites, individuals are actively using privacy settings, and, in part, they are doing so depending on their personality traits. We have found two types of users: those who are privacy-conscious and those who belong to the pragmatic majority. These users tend to have specific personality traits - the more privacy-conscious they are, the higher their traits of Openness and Extraversion. Men and women share equal amount of personal information; however, women tend to be more cautious and make information less visible. Finally, work-related information (especially that reflecting job status) is selectively shared, while information related to whether one is looking for a partner is shared more widely and is used as "social currency" to maintain Facebook relationships. Software tools that personalize our privacy settings are likely to be developed in the future. Until then, the more we know about the kind of privacy work in which social-networking users engage to achieve comfortable amounts of publicity and privacy, the better.

**Acknowledgment.** We thank EPSRC for its financial support through the Horizon Digital Economy Research grant (EP/G065802/1).

## References

- [Amichai-Hamburger and Vinitzky 2010] Amichai-Hamburger, Y., and Vinitzky, G. 2010. Social network use and personality. *Journal of Computers in Human Behavior*.
- [Back et al. 2010] Back et al. 2010. Facebook Profiles Reflect Actual Personality, Not Self-Idealization. *Psychological Science*.
- [Baker 2001] Baker, F. B. 2001. *The Basics of Item Response Theory*. ERIC Clearing-house.
- [Boyd and Hargittai 2010] Boyd, D., and Hargittai, E. 2010. Facebook privacy settings: Who cares? *First Monday*.
- [Butt and Phillips 2008] Butt, S., and Phillips, J. G. 2008. Personality and self-reported mobile phone use. *Journal of Computers in Human Behavior*.
- [Chang et al. 2010] Chang, J.; Rosenn, I.; Backstrom, L.; and Marlow, C. 2010. ePluribus: Ethnicity in Social Networks. In *AAAI ICWSM*.
- [Costa and McCrae 2005] Costa, P., and McCrae, R. 2005. *The Revised NEO Personality Inventory (NEO-PI-R)*. Handbook of Personality Theory and Testing (SAGE).
- [de Bruin and Buchner 2010] de Bruin, G. P., and Buchner, M. 2010. Factor and item response theory analysis of the Protean and Boundaryless Career Attitude Scales. *Journal of Industrial Psychology*.
- [Gogolinski 2010] Gogolinski, T. B. 2010. Effects Of Self-Monitoring and Public Self-Consciousness on Perceptions of Facebook Profiles. *CAAUR Journal*.
- [Golbeck, Robles, and Turner 2011] Golbeck, J.; Robles, C.; and Turner, K. 2011. Predicting personality with social media. In *Proc. of the 29<sup>th</sup> ACM CHI*.
- [Goldberg et al. 2006] Goldberg, L.; Johnson, J.; Eber, H.; Hogan, R.; Ashton, M.; Cloninger, R.; and Gough, H. 2006. The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*.
- [Gross and Acquisti 2005] Gross, R., and Acquisti, A. 2005. Information revelation and privacy in online social networks. In *Proc. of the ACM Workshop WPES*.
- [Henrich, Heine, and Norenzayan 2010] Henrich, J.; Heine, S.; and Norenzayan, A. 2010. The weirdest people in the world? *Journal of Behavioral and Brain Sciences*.
- [Hofstede 1996] Hofstede, G. 1996. *Cultures and Organizations, Software of the Mind: Intercultural Cooperation and its Importance for Survival*. McGraw-Hill.
- [John and Srivastava 1999] John, O. P., and Srivastava, S. 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. Guilford Press.
- [Lewis, Kaufman, and Christakis 2008] Lewis, K.; Kaufman, J.; and Christakis, N. 2008. The Taste for Privacy: An Analysis of College Student Privacy Settings in an Online Social Network. *Journal of Computer-Mediated Communication*.
- [Lin 2008] Lin, C. S. 2008. Exploring the Personality Trait of Self-Monitoring on Technology Usage of Web Portals. *Journal of CyberPsychology and Behavior*.
- [Liu and Terzi 2009] Liu, K., and Terzi, E. 2009. A Framework for Computing the Privacy Scores of Users in Online Social Networks. In *Proc. of IEEE ICDM*.
- [MacIntosh 1998] MacIntosh, R. 1998. Global Attitude Measurement: An Assessment of the World Values Survey Postmaterialism Scale. *American Sociological Review*.
- [Mehra, Kilduff, and Brass 2000] Mehra, A.; Kilduff, M.; and Brass, D. 2000. The Social Networks of High and Low Self-monitors: Implications for Workplace Performance. *Administrative Science Quarterly*.
- [Nettle 2007] Nettle, D. 2007. *Personality: What Makes You the Way You Are*. Oxford University Press.
- [Nippert-Eng 2010] Nippert-Eng, C. E. 2010. *Islands of Privacy*. University of Chicago Press.
- [Papacharissi 2010] Papacharissi, Z. 2010. Privacy as a luxury commodity. *First Monday*.
- [Ross et al. 2009] Ross, C.; Orr, E. S.; Sisc, M.; Arseneault, J. M.; Simmering, M. G.; and Orr, R. R. 2009. Personality and motivations associated with Facebook use. *Journal of Computers in Human Behavior*.
- [Schrage 1997] Schrage, M. 1997. The Relationship Revolution. *Merrill Lynch Forum*.
- [Schrammel, Köffel, and Tscheligi 2009] Schrammel, J.; Köffel, C.; and Tscheligi, M. 2009. Personality traits, usage patterns and information disclosure in online communities. In *Proc. of the 23<sup>rd</sup> ACM BCS-HCI*.
- [Snyder 1974] Snyder, M. 1974. Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*.
- [Vishwanath 2006] Vishwanath, A. 2006. The Effect of the Number of Opinion Seekers and Leaders on Technology Attitudes and Choices. *Journal of Human Communication Research*.
- [Westin 1991] Westin, A. F. 1991. Harris-Equifax Consumer Privacy Survey 1991.
- [Yang 2006] Yang, X. 2006. Effects of Estimation Bias on Multiple-Category Classification With an IRT-Based Adaptive Classification Procedure. *EPM Journal*.