Factors Mediating Disclosure in Social Network Sites

Fred Stutzman, Robert Capra, Jamila Thompson

School of Information and Library Science
University of North Carolina at Chapel Hill
fred.stutzman@unc.edu, rcapra3@unc.edu, jamilaunc@gmail.com

Abstract

In this paper, we explore how privacy settings and privacy policy consumption (reading the privacy policy) affect the relationship between privacy attitudes and disclosure behaviors. We present results from a survey completed by 122 users of Facebook regarding their information disclosure practices and their attitudes about privacy. Based on our data, we develop and evaluate a model for understanding factors that affect how privacy attitudes influence disclosure and discuss implications for social network sites. Our analysis shows that the relationship between privacy attitudes and certain types of disclosures (those furthering contact) are controlled by privacy policy consumption and privacy behaviors. This provides evidence that social network sites could help mitigate concerns about disclosure by providing transparent privacy policies and privacy controls.

Introduction

Activity in a networked community can be stimulated by the creation and exposure of user-generated content (Erickson & Kellogg, 2000; Mynatt et al., 1998). In a social network site like Facebook, shared pictures, status updates, and links keep users interested and drive page views. However, user attitudes toward privacy may negatively impact the volume and type of content shared in a social network site (Acquisti & Gross, 2006), which may in turn have negative implications for social network site vibrancy. For example, a user who is particularly concerned about ownership or privacy of shared data may limit information disclosed in a social network. Users are in good stead to be concerned about information shared in social network sites; harms originating from inadvertent or improper disclosures include legal sanctions (Grimmelmann, 2009), unintentional exposure of personal data (Jernigan & Mistree, 2009), and physical threats including cyberbullying (Palfrey, 2009). Social network site administrators are challenged to implement technologies and policies that address user privacy concerns while enabling the free flow of content. In the design community, researchers are working to create systems that support the sharing of content in a way that reduces potential harms to users (Hawkey & Inkpen, 2006; Nov & Wattal, 2009).

Research by Cranor, Reagle and Ackerman (2000) explored attitudes towards information disclosure on the Internet. The researchers found that prior attitudes, such as conceptions about the value of identifiers, were important factors in online disclosure. Notably, the researchers also found that transparency (e.g. the posting of a privacy policy) and personal information control were important factors in online disclosure. General attitudes about privacy also play a strong role in individual online disclosure practices. If a company sufficiently addresses user privacy concerns, the role privacy concerns play in online disclosure may be mitigated.

In the following study we explore factors that potentially control the relationship between privacy attitudes and disclosure behaviors in a social network site, Facebook. As prior research suggests, the relationship between privacy attitudes and disclosure behaviors may be mitigated by education about company privacy practices, and by increasing individual control over disclosures (e.g. Ahern et al., 2007; Cranor et al., 2000; Fogel & Nehmad, 2009; Lewis et al., 2008). By controlling the effects of privacy attitudes on disclosure, social network site users may feel free to share content and engage in other community-enhancing behavior (Iriberri & Leroy, 2009).

In a social network site, individuals exert control over their disclosures through mental and technical strategies (Lampinen et al., 2009). A primary method of disclosure control is the utilization of privacy settings. We therefore treat privacy behaviors, including privacy personalization and customization, as mediators of the relationship between privacy attitudes and disclosure behavior. Transparency about company practices is another important mediator of privacy attitudes specified by Cranor, Reagle and Ackerman (2000). The delivery of privacy policy information, one such mode of transparency, may play a critical role in a variety of Internet transactions (Egelman et al., 2009). In this study, we treat the Facebook privacy policy as the vehicle for transparency. As the privacy policy outlines what happens to data users disclose in Facebook, we explore if reading the privacy policy controls the relationship between

privacy attitudes and disclosure practices. Overall, the two factors mediating the relationship between privacy attitudes and privacy behaviors we explore are 1) use of privacy settings and 2) individual reading of the privacy policy.

It is important to note the relationship between privacy attitudes and privacy behaviors is a complicated one. Often times, stated privacy attitudes and privacy behaviors do not match, in both experimental and field studies (Acquisti & Grossklags, 2004). Privacy is a normative, subjective construct. In the context of Human-Computer Interaction, privacy is a contextual and contingent information practice, and should be studied in context (Dourish & Anderson, 2006). Our study is a situated analysis, focusing on the behavior of a specific population.

It should also be noted that other variables influence the relationship between privacy attitudes and disclosure practices. The composition of one's personal sharing network is one such variable; Adamic et al. highlighted the role of self-similarity in friendship connection in a social network site (2003), building upon the work of McPherson et al. (2001). Similar preferential attachment has been observed in Last.fm (Baym & Ledbetter, 2009) and in online dating sites (Fiore & Donath, 2005). This study focuses on privacy settings and the privacy policy because these are two variables that can be directly influenced by a social network site. Network composition is an externality; it may exert influence on the relationship between privacy attitudes and disclosure, but it is not a variable that is generally under company control. By focusing on variables under company control, the findings of this study may be more broadly applicable to other social network sites.

To explore how use of privacy settings and privacy policy consumption mediate the relationship between privacy attitudes and disclosure behaviors, we use a series of regressions to: First, validate the relationship between privacy attitudes and disclosure behaviors; Second, explore the efficacy of the control measures; Third, estimate the effects of the control measures on the relationship between privacy attitudes and disclosure behaviors. We find that while the relationship between privacy attitudes and privacy behavior is controlled for some types of disclosure, the relationship between privacy attitudes and overall disclosure is not controlled. This indicates that Facebook may need to adjust privacy controls or user privacy education in order to limit the influence of privacy attitudes on disclosure behavior. Specifically, a benefit model is proposed based on predicted probabilities.

Method

Participants and Data Collection

Participants were recruited widely from the University of North Carolina (UNC) community through an email solicitation sent to a campus-wide opt-in mass-email listserv. UNC undergraduates who used Facebook were invited to follow a link to complete a survey about privacy awareness on Facebook. The survey was hosted on the Survey Monkey on-line system and contained 16 questions about demographics, privacy attitudes, and Facebook sharing behaviors. Data collection lasted for approximately two weeks during March and April 2009. During this time, 122 respondents completed the survey. Respondents ranged in age from 18 to 23 years of age and were disproportionately female (76% of our respondents were female versus 59% of the total undergraduate population).

Measures

Four main measures from the survey are used in the analysis presented in this paper: privacy attitudes, privacy behaviors, privacy policy consumption, and disclosure practices. These are each described in more detail in the following sections.

Privacy Attitudes. Privacy attitudes represent the independent measure in this study. Privacy attitudes are measured by a unidimensional scale that asks respondents to "*Indicate [their] level of concern about the following potential privacy risks that arise when [they] share [their] personal information on Facebook*." The response categories were *very concerned*, *somewhat concerned*, and *not concerned*. The potential risk items were: *identity theft*, *information leakage*, *hackers*, *blackmail*, and *cyberstalking*. We tested the latent structure of the items in this question using principal components analysis and found that all the items loaded on a single factor (varimax rotation, eigenvalue: 3.1) accounting for 62% of the overall variance. Cronbach's alpha was 0.846, indicating acceptable reliability. The responses were summed to create a measure of respondent attitudes about privacy risks stemming from sharing information on Facebook.

Privacy Behavior. Privacy behaviors are a control measure in our study and are based on two specific questions from the survey. Respondents were asked about their current privacy settings on Facebook. The first question asked if the *respondent had changed their privacy settings from the default*; we dichotomized the responses to this question as yes or no. The second question asked if the respondent had, "*ever customized which individual friends are allowed to view your content (e.g. wall, photos, notes, etc.)?"* These responses were also dichotomized as yes or no. We refer to the first question as representing privacy *personalization* and the second question as

representing privacy *customization*. A cross-tabulation of the responses is reported in Table 1.

		Ever customized?		
		No	Yes	Total
Ever personalized?	No	22% (9)	7% (6)	12% (15)
	Yes	77% (31)	92% (74)	87% (105)
	Total	100% (40)	100% (80)	100% (120)

Table 1. Privacy Behaviors.

Privacy Policy Consumption. We then asked respondents to indicate the degree to which they had read Facebook's privacy policy: "*I have read most, or all*", "*I have scanned*", "*I know [it exists] but I've never read it*", "*I did not know [it exists]*". In our analysis, we combined the latter two choices to result in three categories (percentage of responses are given in parenthesis): *read (5.8%)*, *scanned (47.1%)*, *not read (47.1%)*.

The privacy policy is a complex legal document with over 5000 words. To gauge comprehension of the policy, we asked two factual questions on the survey. First, we asked if "*Facebook will collect information about you from other services (i.e. IM/AIM)*". Second, we asked, if "*Facebook will share your personal information with third-party advertisers, data aggregators, and external applications*". According to the privacy policy, both these are true. However, 47% of our respondents answered "true" to the first question, and 65% answered "true" to the second. In our multivariate analysis we did not find a significant "comprehension" effect, either as main or interaction effect. Therefore, this analysis will primarily focus on level of privacy policy consumption.

Disclosure Behavior. Disclosure behavior is measured by the kind of information a user posts on his or her Facebook profile. We asked, "*What personal information have you EVER posted on Facebook*?" and provided yes/no choices for the list of items in Table 2.

INFORMATION DISCLOSED	**% reporting yes**
Identity based disclosure	
Real name	98.4%
Birth date	95.1%
High school	94.2%
Profile picture	98.4%
Disclosure furthering contact	
Campus address	36.3%
Cell phone number	42.4%
IM Screenname	73.3%
Email address	91.0%

Table 2. Disclosure Behaviors.

In our analyses, we summarize this measure as a count of the number of items that a respondent reported disclosing (i.e. answered "yes"). However, the count distribution does not fit a Poisson distribution due to right censoring inherent in the question design. To limit errors attributable to misspecification, we treat this measure of disclosure behavior as a series of ordinal categories. Individuals with higher disclosure counts are treated as being in a higher disclosure category than those with lower counts. In an elaboration of the categorization, we divide the items in Table 2 into two groups: 1) disclosures that describe aspects of a individual's identity (e.g. real name, birth date), and 2) disclosures that may be used to further contact with an individual (e.g. address, phone number, email address). We refer to this first group as *identity based disclosures* and the second group as *disclosures furthering contact*.

Analysis

Overview

Our analysis follows the model outlined in Figure 1. We first explore the baseline association between privacy attitudes and disclosure practices (H1). Next, we test the validity of the controls by testing the relationships between the independent variable and the controls (H2 and H3) and between the controls and the dependent variable (H4 and H5). Then, we examine the final model with the controls included (H6).

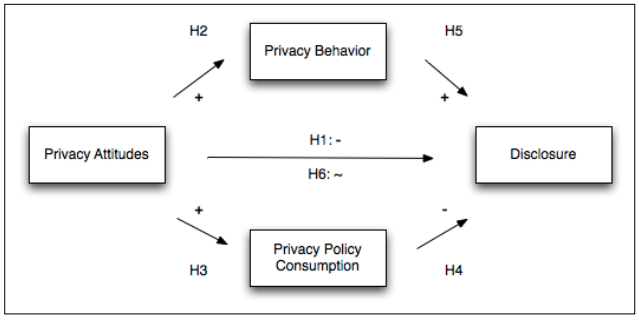

Figure 1. Overview of the Analysis Model.

Figure 1 indicates the specific hypotheses (H1-H6) we test as part of the analysis. These are explained in more detail below.

H1: People who have greater concern about privacy will disclose less information (negative relationship). This is our initial test of the baseline association between privacy attitudes and disclosure practices.

H2: People who have greater concern about privacy will be more likely to engage in privacy protecting behaviors such as personalizing privacy settings or customizing which friends can see content (positive relationship).

H3: People who have greater concern about privacy will be more likely to read the privacy policy (positive relationship).

H4: People who read more of the privacy policy will disclose less information (negative relationship).

H5: People who personalize or customize privacy settings will disclose more information (positive relationship). The rationale behind this hypothesis is that by personalizing or customizing privacy settings, people will feel more comfortable sharing information because they have greater control over who can access it.

H6: With hypothesis six we validate the control measures. We do this by testing if privacy behaviors and privacy policy consumption render the relationship between privacy attitudes and disclosure as non-significant.

Effects of Privacy Attitudes on Disclosure (Baseline)

To analyze the relationship between privacy attitudes and disclosure (H1), we used the measures of privacy attitudes and disclosure behaviors as described in the Methods section. To test the hypothesis, we used ordinal logistic regression. Gender is retained as a control in this and all other regression models in this paper because it has been linked to differential practices on the Internet (Herring, 2003; Jackson et al., 2001; Jones et al., 2009) and in social network sites (Lewis et al., 2008; Thelwall, 2008). The ordinal logistic regression model showed that increased concern for privacy is significantly negatively associated with overall disclosures in Facebook (p=0.004). For each one-unit increase in an individual's score on our privacy attitudes scale, the individual's odds of being in a greater disclosure category decrease by 0.5227. Put another way, as an individual's privacy concern increases, they are less likely to increase their disclosures. Gender is not significant in the model (e^{β}= 1.92, p=0.103). Therefore, H1 is upheld.

Effects of Privacy Attitudes on Privacy Behavior

We collected two measures of privacy behavior, *privacy personalization* and *privacy customization*, as described in the Measures section. To examine the effects of privacy attitudes on both these behaviors, we refine H2 into two testable hypotheses, one for each dependent measure. H2a will look at the relationship between privacy attitudes and privacy personalization (i.e. changing the default privacy settings). H2b will examine the relationship between privacy attitudes and privacy customization (i.e. customizing which individual friends have access to content). Again, gender is included in our model and binary logistic regression is used for the test. For both H2a and H2b, regression was done twice, once using the combined measure of privacy attitudes as described in the Measures section and once using the individual items (e.g. identity theft, information leakage, hackers, blackmail, and cyberstalking). Based on the models, H2a was not upheld (Table 3). Notably, 87% of the respondents indicated they had personalized their Facebook privacy settings, leaving little room for explanation of variance. For H2b, the regression using the combined measure of privacy attitudes found no significant variables. However, the regression using the individual privacy attitude items found information leakage to be a significant predictor (p=0.04). A one-unit increase in concern about information leakage was found to be associated with a 2.22 times increase in the odds of privacy customization. We find this to be an intuitive relationship – an individual who fears their private information may be obtained by people they do not wish to have it ("information leakage") may be more likely to customize their privacy settings to restrict disclosures to only people they wish to see them.

DV: Privacy Behaviors	Hypothesis 2a		Hypothesis 2b	
	Privacy Personalization		Privacy Customization	
Overall Privacy	1.123 (0.545)		1.407 (0.486)	
Identity Theft		0.788 (0.425)		0.531 (0.213)
Information Leakage		2.366 (1.278)		2.222* (0.857)
Hackers		0.593 (0.325)		1.164 (0.466)
Blackmail		0.987 (0.477)		1.220 (0.417)
Cyber-Stalking		1.282 (0.693)		1.069 (0.409)
Gender (M=1)	0.611 (0.372)	0.546 (0.364)	0.677 (0.305)	0.647 (0.314)
Constant	6.412 (6.584)	3.967 (4.331)	1.149 (0.834)	0.643 (0.514)
Chi Square	0.796	4.608	2.155	8.400
Observations	120	119	121	120
Odds ratios. Standard errors in parentheses,* p<0.05.				

Table 3: Elaboration of Hypothesis 2.

We find conditional support for hypothesis H2. The combined measure of privacy attitudes does not predict either privacy personalization or privacy customization. However, the individual measure, "information leakage" does significantly predict privacy customization. Gender was not significant in any of our models.

Effects of Privacy Attitudes on Privacy Policy Consumption

Next we examine the relationship between privacy attitudes and privacy policy consumption (H3). Our privacy policy reading measure had three levels regarding the amount of the policy that the participant had read: *none*, *scanned*, and *most/all*. We hypothesize that people with greater privacy concerns will read **more** of the privacy policy. To test this, we used a multinomial logistic regression using our combined privacy attitudes measure, and the three levels of privacy policy consumption, with "*scanned*" as the base measure. As with our other models, gender is included as a control. In the regression model, we found a significant positive association between

privacy attitudes and reading "most or all" of the privacy policy as compared to "scanning" (p=0.02). A one unit of increase on our combined privacy attitudes measure resulted in a 6.35 factor increase in the odds of reading most or all of the policy rather than just scanning it (Table 4). Hypothesis H3 is supported.

DV: Level of Privacy Policy Consumption	Read none of the privacy policy	Read some/ scanned privacy policy	Read most or all of the privacy policy
Privacy attitudes	1.316 (0.445)	REFERENCE	6.354* (5.067)
Gender (M = 1)	1.193 (0.541)		1.983 (1.862)
Constant	0.563 (0.401)		0.00182** (0.00364)
Chi Square	6.509	6.509	6.509
Observations	121	121	121
Odds ratios. Standard errors in parentheses, *p<0.05, **p<0.01.			

Table 4: Relationship between privacy attitudes and privacy policy consumption.

Because of the complex nature of the privacy policy, we examined the correlation between levels of privacy policy consumption (*none, scanned, most*) and being able to correctly answer factual questions about the policy. We found a positive (0.18) and marginally significant (p=0.53) correlation between higher levels of privacy policy reading and providing more correct answers to the two factual questions about the policy.

Effects of Privacy Policy Consumption on Disclosure

In hypothesis H4, we predict that increased privacy policy consumption is associated with disclosing less information in Facebook. As a person learns more about what happens to information disclosed on Facebook, we predict they may disclose less. We test this hypothesis using ordered logistic regression using the three levels of privacy policy reading (none, scanned, most/all) and the combined measure of disclosure behavior described in the Measures section. Gender is included as a control. In the regression model, we found that increased reading of the Facebook privacy policy is significantly (p=0.04) and negatively associated with overall disclosures, indicating that a one-unit increase in privacy policy reading is associated with a .558 factor decrease in the odds of being in a higher disclosure category (Table 5). Gender was not significant. Therefore, H4 is supported. Notably, in hypothesis H3, we found that privacy attitudes are associated with greater levels of privacy policy consumption, so it is possible that privacy attitudes are a latent construct that is acting through privacy policy consumption to lead to the effect observed here in H4. We will test for this possibility in hypothesis H6 using a nested regression model.

Effects of Privacy Behavior on Disclosure

In hypothesis H5, we examine the relationship between privacy behaviors (i.e. personalization and customization) and disclosure behaviors. Privacy personalization is measured as a dichotomous variable (yes/no) indicating if the respondent has changed their Facebook privacy settings. Privacy customization is also a dichotomous variable (yes/no) that indicates if the respondent has ever customized which individual friends have access to content. In Facebook, these variables are orthogonal; when using them as independent variables, they can be included together in a single regression model without confound.

We use the same combined disclosure measure described in previous analyses as our dependent variable and use ordinal logistic regression to examine the effect. In the model, no effect of privacy personalization was observed, but privacy customization was found to be significantly and positively (p=0.007) associated with increased disclosures, indicating that people who have customized who can see their content are approximately 2.5 times more likely to be in a higher disclosure category on Facebook (Table 5). In addition, gender was found to be significant (p=0.03), with males sharing more than females. Hypothesis H5 is supported, but as with the previous hypothesis, privacy attitudes could be a latent construct acting through privacy customization (due to H2). We will test for this possibility in hypothesis H6 using a nested regression model.

DV: Combined disclosure measure	Hypothesis 4	Hypothesis 5
Privacy policy reading	0.558* (0.160)	
Gender (M=1)	2.170 (0.875)	2.534* (1.055)
Privacy personalization		0.623 (0.309)
Privacy customization		2.822** (1.093)
Chi Square	7.858	10.95
Observations	111	110
Odds ratios. Standard error in parentheses, ** p<0.01, * p<0.05. Cut points not reported to preserve space.		

Table 5: Relationship between privacy policy consumption, privacy behaviors, and disclosure.

Effects of Privacy Attitudes on Disclosure (including controls)

With H2-H5 upheld (conditionally for H2), the control variables (e.g. privacy behaviors and privacy policy consumption) represent potential mitigating factors in the relationship between privacy attitudes and disclosure behaviors (H6). We conduct a nested ordinal logistic regression to simultaneously evaluate the effects of privacy attitudes, privacy behaviors, and privacy policy

consumption on disclosure behavior. We run the regression models for both the combined disclosure measure and for the two disclosure groups described in the Measures section: *identity based disclosures*, and *disclosures furthering contact*.

Before running the regression model for H6, we examine the relationship between the control variables to ensure low covariance. The correlation between privacy personalization and privacy policy reading is significant, but low (r = 0.2). The correlation between privacy customization and privacy policy reading is not significant (r = -0.02).

The nested logistic regression to examine the relationship between privacy attitudes and disclosure behaviors allows the estimation of effects based on grouped predictors using the likelihood ratio test. We use the following grouped predictors: 1) gender, 2) combined measure of privacy attitudes, 3) privacy personalization and privacy customization, and 4) privacy policy consumption.

In the first regression using the combined disclosure measure as the dependent variable, two blocks are significant: the combined measure of privacy attitudes and privacy customization (Table 6, H6a). Privacy attitudes exert a significant (p=0.03) and negative effect on overall disclosures. Privacy customization is significant (p=0.006) – an individual who has customized privacy settings is likely to share more than an individual who has not by a factor of 2.905. Thus, in the first regression, we do not see the relationship between privacy attitudes and disclosure behavior fully mitigated by the control variables.

To explore the relationship between privacy attitudes and disclosure further, we refine the analysis of disclosure by considering two subscales of disclosures: identity based disclosures and disclosures furthering contact (see Table 2 for list of items). Using nested ordinal logistic regression, we evaluate the relationship between our predictors and the identity based disclosure subscale (Table 6, H6b). We do not find any significant blocks, and therefore the null model is not improved. Identity based disclosures are very common in Facebook, so there is little variance to explain.

Next we repeat this analysis for the disclosures furthering contact (Table 6, H6c). In this model, we find that privacy behaviors and privacy policy consumption are significant, additive steps. With regard to privacy behavior, privacy customization is the significant predictor (p=0.002). An individual that has customized privacy settings is 3.34 times as likely to share more disclosures furthering contact than someone that has not customized privacy settings. Privacy policy consumption is also significant (p<.031), with reading more of the privacy policy associated with less disclosures furthering contact by a factor of .527. Notably, privacy attitudes are not a significant predictor, indicating that the relationship between privacy attitudes and disclosures furthering contact is mitigated by the specified control variables.

	H6a	H6b	H6c
DV: Amount of disclosed information	Combined disclosure measure	Identity based disclosures	Disclosures furthering contact
Privacy attitudes scale	0.513* (0.157)	0.827 (0.447)	0.575 (0.175)
Privacy person-alization	0.833 (0.420)	2.958 (2.372)	0.678 (0.357)
Customizing privacy	2.905** (1.125)	0.740 (0.498)	3.304** (1.295)
Reading of FB Privacy Policy	0.567 (0.170)	0.781 (0.396)	0.527* (0.157)
Gender (M = 1)	2.320* (0.962)	1.909 (1.568)	2.018 (0.828)
Chi Square	19.27	2.468	19.79
Observations	110	119	111
Standard error in parentheses, ** p<0.01, * p<0.05. Cut points not reported to preserve space.			

Table 6: Evaluation of controlled model.

The model evaluation allows explication of how privacy attitudes affect disclosure. Overall, privacy attitudes and privacy behaviors are significantly associated with disclosure behavior in Facebook. In the case of disclosures furthering contact, it is demonstrated that privacy policy consumption and privacy behaviors control the relationship between privacy attitudes and disclosures. This is important insight for administrators of social network sites, providing evidence that both transparency and control can mitigate concerns about disclosure, and it is in line with the findings of Cranor et al. (2000).

Benefit Analysis

Administrators of social network sites are challenged to address user privacy concerns; in this study, we have explored how privacy behaviors and privacy policy consumption may mitigate privacy concerns. To explore the impact these controls have on disclosure behavior, we present a benefit analysis, using predicted probabilities. Predicted probabilities allow us to demonstrate the impact of a manipulation of a single variable, holding all other values constant. Using the disclosures furthering contact subscale as the dependent measure, we manipulate privacy customization and privacy policy consumption levels.

The benefit analysis can be interpreted as the effect of manipulating the control variables on the probability of being in the particular disclosure category. As demonstrated in Table 7, engaging in privacy customization increases the probability of being in the highest disclosure categories. At the same time, privacy policy consumption lowers the probability of being in the highest disclosure category. Comparing the probabilities, the effect of privacy customization is roughly double the effect of privacy policy consumption. While increased

Probability of being in *disclosure furthering contact* category	Moving from no **privacy customization** to some privacy customization	Moving from no **privacy policy consumption** to some privacy policy consumption	Moving from some **privacy policy consumption** to most privacy policy consumption
Lowest Category, Pr(y=0|x):	-0.0781 (-0.1509, -0.0054)	0.0325 (-0.0016, 0.0666)	0.0559 (-0.0232, 0.1351)
Pr(y=1|x):	-0.1363 (-0.2388, -0.0338)	0.0671 (0.0021, 0.1321)	0.0806 (-0.0017, 0.1629)
Pr(y=2|x):	-0.0682 (-0.1323, -0.0041)	0.0588 (-0.0047, 0.1223)	0.0036 (-0.0575, 0.0647)
Pr(y=3|x):	0.1274 (0.0320, 0.2228)	-0.0608 (-0.1215, -0.0001)	-0.0778 (-0.1516, -0.0039)
Highest Category, Pr(y=4|x):	0.1552 (0.0598, 0.2506)	-0.0976 (-0.1902, -0.0050)	-0.0623 (-0.0991, -0.0255)
95% confidence intervals in parentheses, using analytical derivatives.			

Table 7: Predicted probabilities of increased disclosure, negative effects shaded.

privacy policy consumption may exert negative influence on disclosure, the effect is more than offset through the use of privacy customization.

Discussion

Social network sites, such as Facebook, thrive on user-contributed content. However, many users report apprehension about the risks that may result from sharing content in social network sites (Acquisti & Gross, 2006). By increasing transparency - educating users about their personal data with a privacy policy, and by providing privacy controls, social network sites may be able to alleviate some of the privacy concerns that affect the contribution of user data to the site. In this paper, we explore the relationship between privacy attitudes and disclosure behavior, as mediated by privacy behavior and privacy policy education.

Hypothesis H1 establishes baseline relationship between privacy attitudes and disclosure practices. Hypotheses H2 and H3 explore the relationship between privacy attitudes and privacy behavior/privacy policy consumption, finding significant associations. In hypothesis H4 and H5, we explore the relationship between privacy behavior and disclosure, and privacy policy consumption and disclosure. The hypotheses outlined in H4 and H5 are upheld. In the full analysis (H6), we explore the extent to which privacy education (via privacy policy consumption) and privacy controls (via privacy customization and personalization) mitigate the relationship between privacy attitudes and disclosure. In the first subscale, identity-based disclosure, no predictors are significant due to the homogeneity of this disclosure category. In the second subscale, disclosures furthering contact, results are promising. Privacy attitudes are not significantly associated with disclosure behaviors, their effect mediated by privacy policy consumption and privacy behaviors.

To address the negative impact of privacy attitudes on disclosure, research suggests that social network sites should increase transparency though the privacy policy, and by allowing privacy control. Using predicted probabilities, we demonstrate that the use of privacy controls offsets the negative impact on disclosure of increased privacy policy consumption. This demonstrates that social network sites can mitigate the effect of privacy attitudes on disclosure practices while ultimately encouraging greater levels of sharing, a mutually beneficial outcome for the users and the site.

Limitations and Conclusions

There are a number of important limitations of this study. First, the data examined are self-reported, which is a source of potential error. While we utilized a solicitation method that allowed access to a diverse population, the use of convenience sampling limits the generalizability of the results. Finally, as this data was cross-sectional, it is not possible to make causal claims about the results. While there is substantial evidence that privacy attitudes influence disclosure, we would need to run a longitudinal analysis to demonstrate causality.

This paper makes a number of contributions to our understanding of privacy in social media environment. First, we demonstrated the significant, negative association between privacy attitudes and disclosures practices. We then identified privacy behaviors, and privacy policy consumption as valid control measures mitigating the relationship between privacy attitudes and disclosure. Finally, we elaborated the analysis to include differing types of disclosure. In doing so, we identified how the relationship between privacy attitudes and these specific disclosure types can be effectively mitigated. This important empirical finding will provide insight for designers and maintainers of social media as they consider implementation of privacy features.

Privacy is a contextual and contingent information practice (Dourish & Anderson, 2006). By examining the behaviors of users in context, this study contributes to the growing body of work exploring privacy practice in large-scale social media and network sites. As more users join social network sites, mitigating privacy concerns while encouraging sharing becomes a chief concern for site administrators. The importance of this work lies in the mutual benefit it demonstrates; when social network sites address the privacy and transparency needs of users, the users share more freely in the social network.

References

Acquisti, A. and Grossklags, J. (2004). Privacy attitudes and privacy behavior: Losses, gains, and hyperbolic discounting. In Camp, J. and Lewis, R. (Eds.), *The Economics of Information Security* (pp. 1-15). Kluwer Academic Publishers.

Acquisti, A. and Gross, R. (2006). Imagined communities: awareness, information sharing, and privacy on the Facebook. In *PET, Heidelberg, 2006* (pp. 36-56). Springer-Verlag.

Adamic, L., Buyukkokten, O., and Adar, E. (2003). A Social Network Caught in the Web. *First Monday*, 8(6).

Ahern, S., Eckles, D., Good, N., King, S., Naaman, M., and Nair, R. (2007). Over-Exposed? Privacy Patterns and Considerations in Online and Mobile Photo Sharing. In *Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, 2007* (pp. 357-366). ACM Press.

Baym, N. K. and Ledbetter, A. (2009). Tunes that Bind? *Information, Communication & Society*, 12(3), 408-427.

Cranor, L. F., Reagle, J., and Ackerman, M. S. (2000). Beyond Concern: Understanding Net Users' Attitudes About Online Privacy. In Vogelsang, I. and Compaine, B. M. (Eds.), *The Internet Upheaval: Raising Questions, Seeking Answers in Communications Policy* (pp. 47-70). Cambridge, MA: MIT Press.

Dourish, P. and Anderson, K. (2006). Collective Information Practice: Exploring Privacy and Security as Social and Cultural Phenomena. *Human-Computer Interaction*, 21(3), 319-342.

Egelman, S., Tsai, J., Cranor, L. F., and Acquisti, A. (2009). Timing is everything?: the effects of timing and placement of online privacy indicators. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems, New York, NY, USA, 2009* (pp. 319-328). ACM Press.

Erickson, T. and Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction*, 7(1), 59-83.

Fiore, A. T. and Donath, J. S. (2005). Homophily in Online Dating: When Do You Like Someone Like Yourself? In *Proceedings of CHI 2005*. ACM Press.

Fogel, J. and Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, 25(1), 153-160.

Grimmelmann, J. T. (2009). Facebook and the Social Dynamics of Privacy. *Iowa Law Review*, 95(4).

Hawkey, K. and Inkpen, K. M. (2006). Keeping up appearances: understanding the dimensions of incidental information privacy. In *Proceedings of the SIGCHI conference on Human Factors in computing systems, New York, NY, USA, 2006* (pp. 821-830). ACM Press.

Herring, S. (2003). Gender and Power in On-Line Communications. In Holmes, J. and Meyerhoff, M. (Eds.), *The Handbook of Language and Gender* (pp. 202-228). Oxford University Press.

Iriberri, A. and Leroy, G. (2009). A life-cycle perspective on online community success. *ACM Computing Surveys*, 41(2), 1-29.

Jackson, L. A., Ervin, K. S., Gardner, P. D., and Schmitt, N. (2001). *Gender and the Internet: Women Communicating and Men Searching*. Sex Roles, 44(5), 363-379.

Jernigan, C. and Mistree, B. (2009). Gaydar: Facebook friendships expose sexual orientation. *First Monday*, 14(10).

Jones, S., Johnson-Yale, C., Millermaier, S., and Perez, F. S. (2009). U.S. College Students' Internet Use: Race, Gender and Digital Divides. *Journal of Computer-Mediated Communication*, 14(2), 244-264.

Lampinen, A., Tamminen, S., and Oulasvirta, A. (2009). All My People Right Here, Right Now: management of group co-presence on a social networking site. In *GROUP '09: Proceedings of the ACM 2009 international conference on Supporting group work, New York, NY, USA, 2009* (pp. 281-290). ACM Press.

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*, 14(1), 79-100.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1), 415-444.

Mynatt, E. D., O'Day, V. L., Adler, A., and Ito, M. (1998). Network Communities: Something Old, Something New, Something Borrowed.... In *1998 Computer Supported Cooperative Work (CSCW)* (pp. 123-156). Springer.

Nov, O. and Wattal, S. (2009). Social computing privacy concerns: antecedents and effects. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems, New York, NY, USA, 2009* (pp. 333-336). ACM Press.

Palfrey, J. (2008). Enhancing Child Safety and Online Technologies. *Internet Safety Task Force.* Retrieved January 10, 2009 from http://cyber.law.harvard.edu/pubrelease/isttf/.

Thelwall, M. (2008). Social networks, gender and friending: An analysis of MySpace member profiles. *Journal of the American Society for Information Science and Technology*, 59(8), 1321–1330.