Computers in Human Behavior 27 (2011) 590–598

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

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, USA

ARTICLE INFO

Article history:
Available online 3 November 2010

Keywords:
Privacy
Social networking
Social network sites
Facebook
Policy
Survey

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.

1. Introduction

Activity in a networked community can be stimulated by the creation and exposure of user-generated content (Erickson & Kellogg, 2000; Mynatt, O'Day, Adler, & Ito, 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 affect the volume and type of content shared in a social network site (Acquisti & Gross, 2006), which may in turn have 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, 2008). 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 positive 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 mediate 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 mediated 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, Kaufman, & Christakis, 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). Therefore, it is useful for both designers and policy-makers to understand the relationship between privacy attitudes and disclosure behaviors in social network sites. In this work, we draw on Altman's (1975) theories of boundary regulation, employing the privacy optimization process to explore if increased knowledge and privacy control potentially mediate the relationship between privacy attitudes and disclosure behaviors.

2. Literature review

The link between disclosure and privacy attitudes has been explored extensively in communication theory. Altman's (1975) original theorization of privacy posits a general optimization

* Corresponding author. Tel.: +1 919 260 8508.
E-mail addresses: fred.stutzman@unc.edu (F. Stutzman), rcapra3@unc.edu (R. Capra), jamilaunc@gmail.com (J. Thompson).

0747-5632/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2010.10.017

function, through which individuals attempt to balance the necessary disclosures of communication with individual privacy control mechanisms. The social individual must disclose, Altman argues, but disclosure is inherently tied to practical mechanisms that adjust our disclosures in relation to our privacy attitudes, goals and knowledge. To achieve an optimal state of privacy, we must know or sense the range of our disclosures. Practically, this knowledge might exist in knowing the boundary of a room (assuming people outside can not hear), knowing the trustworthiness of communication partners (who will gossip, and who will not), or knowing the data practices of a website. It is the interaction of our disclosure goals, and knowledge of the boundaries of our disclosures, that comprises Altman's boundary-regulation theory of privacy, and provides justification for this study's link between disclosure practices and privacy attitudes.

The concept of privacy as boundary regulation has been extended in a number of domains. Petronio (2002) uses boundary regulation as a central metaphor in Communications Privacy Management (CPM), an interpersonal theory of privacy regulation. Petronio argues that individuals create and apply rules that effectively manage disclosure based on goals, context, and attitudes. Indeed, Petronio places attitudes ("Motivational Criteria") as central in the privacy rule development process (Petronio, 2002, pp. 38–39). The concept of the boundary regulating disclosure is extended in the work of Derlega and Chaikin (1977), in which individuals are found to draw on attitudes and goals when constructing self- and dyadic-privacy strategies.

Although specific formulations of boundary regulation vary between the theories of Altman, Petronio, and Derlega and Chaikin, there is strong common support for the concept of knowledge informing boundary. What an individual knows about a space for disclosure, whether it be in pre-conceived knowledge or from feedback, shapes the rules placed on disclosure going forward. As Altman (1975, p. 43) writes: "It is sometimes necessary to escalate responses and make adjustments in self/other boundaries because of misestimates of the effectiveness of the boundary or because of misreadings of the social situation." In Altman's theory, knowledge of context dynamically causes change in the discloser's privacy strategy, thus identifying the important relationship between situational knowledge and privacy behaviors.

In addition to situational knowledge of a context, our privacy behaviors are shaped by the rules in which we attach to our disclosures. This concept is central to Petronio's theorization of "Boundary Coordination". Regarding interpersonal context, Petronio (2002, p.141) observes: "Boundary regulation within a relational sphere often requires calibrating such choices as to how much to tell a partner, when to tell, and how to reveal private information." Petronio sees knowing who to tell or trust as central "rules" of interpersonal disclosure, and argues that these rules fundamentally guide disclosure and are adapted over time. In the context of a social network site, privacy settings place explicit rules on disclosure. The use of these settings may allow individuals more freedom to disclose, as they provide both knowledge of context and knowledge of the disclosure's range. Therefore, we explore the boundary setting process by measuring an individual's privacy behaviors.

2.1. Understanding privacy outcomes

Boundary regulation theories of privacy identify the importance of an individual's knowledge of a communication context when making disclosure decisions. A website's privacy policy – a statement regarding the site's data use and protection practices – is a primary vehicle through which consumers can become informed about what happens to the information they disclose on or to a site. Bonneau and Preibusch (2009) surveyed 45 social network sites, both general-purpose and niche-oriented, finding that almost all had privacy policies in place.

While privacy policies are often criticized as difficult or time-consuming to read (Bonneau & Preibusch, 2009; McDonald & Cranor, 2009; McDonald, Reeder, Kelley, & Cranor, 2009), there is evidence that if a website has a privacy policy, individuals are more likely to share personal information with the website (Cranor et al., 2000). In our study, we focus primarily on the simple act of becoming informed, i.e. reading the privacy policy. We explore the potential mediating effect that increased awareness of a social network site's policy privacy has on the relationship between privacy attitudes and privacy behaviors.

2.2. Identifying privacy behavior

Social network sites offer a variety of tools that allow an individual to set disclosure rules. Facebook, for example, provides friend groupings, item-level access control, block lists, and a range of other techniques for privacy management[1]. These tools allow for the creation of deterministic rules that govern where content is shared or replicated. Recent studies have identified increased use of privacy settings, particularly by the population of interest in this study (Stutzman & Kramer-Duffield, 2010).

It must be noted that privacy in social networks is not a purely deterministic function of technical rules. Lampinen, Tamminen, and Oulasvirta (2009), studying contextual privacy in Facebook, found that individuals exert control over disclosures through "mental" and "technical" strategies. This analysis revealed an interplay between technical methods of disclosure control, such as the utilization of privacy settings, and attitudinal factors that influence privacy strategies. The central challenge of modeling mental strategies is in identification. While these strategies are commonplace (e.g. limiting a disclosure based on audience perception), they are challenging to measure reliably. In this study, we focus primarily on technical strategies of privacy management, but acknowledge the range of strategies available to individual users of social network sites.

2.3. Modeling privacy in social network sites

It is important to note that the relationship between privacy attitudes and privacy behaviors is a complicated one. Often, 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 (Dourish & Anderson, 2006). Our study is therefore a situated analysis, focusing on the behavior of a specific population, with explanatory power limited to the population studied.

It should also be noted that a range of 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, Smith-Lovin, and Cook (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 (e.g. Adamic et al., 2003) is an externality; it may exert influence on the relationship between privacy attitudes and disclosure, but it

[1] http://www.facebook.com/privacy/explanation.php.

is not a variable that is generally under company control (i.e. the social network site). By focusing on variables under company control, the findings of this study may be readily applied by operators of 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. Following Altman's conception of an optimizing function, we finally propose a benefit model based on model predicted probabilities.

3. Method

3.1. 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. This listserv reaches all students that received informational messages from the university, and allows targeting based on student status. We restricted the sample to undergraduate students that used Facebook, inviting them 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. As the undergraduate population of UNC is disproportionately female (60%), and it is common for males to under-participate in web survey research (Heerwegh & Loosveldt, 2008), the patterns of gender participation are within expectation.

3.2. 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.

3.2.1. Privacy attitudes

Privacy attitudes represent the independent measure in this study. Privacy attitudes are measured by a summed 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 *identity theft, information leakage, hackers, blackmail*, and *cyberstalking*, are rooted in a general review of literature on privacy threats (e.g. Grimmelmann, 2009; Hinduja & Patchin, 2008; Palfrey, 2008) and recommendations introduced in the European Network and Information Security Agency report *Security Issues and Recommendations for Online Social Networks* (Hobgen, 2007). To validate the scale, we tested the latent structure of the items in this question using factor 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 high reliability. The responses were summed to create a measure of respondent attitudes about privacy risks stemming from sharing information on Facebook.

3.2.2. 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 are 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.

3.2.3. Privacy policy consumption

Previous work has identified the time commitment associated with reading a privacy policy (McDonald & Cranor, 2009; McDonald et al., 2009) and the effects of simply having a privacy policy (Cranor et al., 2000). Therefore, our privacy policy consumption instrument measures knowledge of the policy's existence, and amount of policy consumption. We 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]*". Reading most or all of the Facebook privacy policy indicates reading a large percentage of the policy. Scanning the policy indicates that the individual has looked at the policy at least once. Never reading is interpreted as the individual has never looked at the policy in any way. Since all respondents indicated knowledge of the policy, we collapsed the last two choices to result in three categories (percentage of responses are given in parenthesis): *read most/all (5.8%), scanned (47.1%), not read (47.1%).*

3.2.4. Disclosure behavior

Disclosure behavior is the dependent variable in this study, and is measured by the volume of identification and contact 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.

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 properly fit a Poisson distribution, commonly employed in the analysis of counts, due to right censoring in the question design. To limit errors attributable to misspecification, we treat the dependent measure of disclosure behavior as a series of ordinal categories. In collecting this information, we selected disclosure behaviors that respondents could recall quickly and accurately based on a review of the individual's profile. 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

Table 1
Privacy behaviors.

		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 2
Disclosure behaviors.

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 screen name	73.3
Email address	91.0

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*. These categories are influenced by previous work exploring disclosure patterns in social network sites, particularly those focused on privacy harms related to disclosure in social network sites (Acquisti & Gross, 2006, 2009; Joinson, Reips, Buchanan, & Schofield, 2010; Tufekci, 2008) We feel that items related to identification and contact with the person are most likely affected by privacy attitudes; this allows for a more robust test of the proposed control mechanisms.

4. Analysis

4.1. Overview

Our analysis follows the model outlined in Fig. 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).

Fig. 1 indicates the specific hypotheses (H1–H6) we test as part of the analysis. These hypotheses are influenced by aspects of boundary regulation theories and prior research on privacy policies. Below, we explain each hypothesis in more detail and provide references to influencing prior research.

H1: Boundary regulation theories of privacy have identified the importance of prior individual attitudes in disclosure formulation (Altman, 1975; Petronio, 2002). We hypothesize that people who have greater concern about privacy will disclose less information in Facebook (negative relationship). This is our initial test of the baseline association between privacy attitudes and disclosure practices.

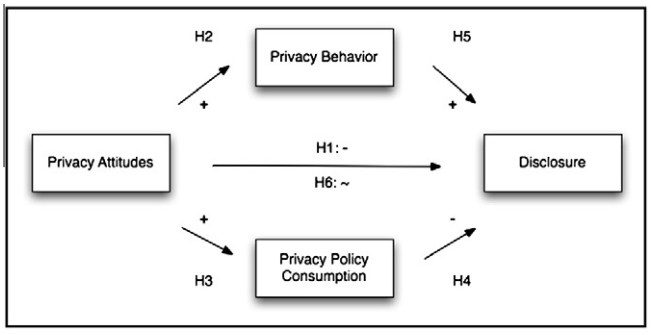

Fig. 1. Overview of the analysis model.

H2: Disclosure rule construction was identified by Petronio (2002) as an important factor in privacy management. In a social network site, the use of privacy controls represents an observable instance of rule formation. We hypothesize that 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: Boundary regulation theories of privacy specify the importance of contextual knowledge about a disclosure environment (Altman, 1975; Petronio, 2002; see also Nissenbaum (2004, 2010)). We hypothesize that people who have greater concern about privacy will be more likely to read more of the privacy policy (positive relationship).

H4: Studies have demonstrated that existence of a privacy policy may increase disclosure by website visitors. Facebook's privacy policy, however, has often been characterized as overly aggressive and invasive (Boyd, 2008; Grigoriadis, 2009). Therefore, we hypothesize that people who read more of the privacy policy will disclose less information (negative relationship).

H5: Altman (1975) and Petronio (2002) suggest that exerting control over disclosure through rule making can allow an individual new freedom to engage in disclosure. Therefore, we hypothesize that 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 predicted by H4 and H5. We do this by testing if privacy behaviors and privacy policy consumption render the relationship between privacy attitudes and disclosure as non-significant.

The specified hypotheses will be tested with regression employing variants of the logit model: logistic regression, as well as ordered and multinomial logit regression. We do this for two reasons. First, the dependent variables in the model are categorical variables, whose distributions best fit logit models. Second, logit models employ a maximum likelihood estimator to derive parameters, which is robust for smaller cell counts (Hosmer & Lemeshow, 2000). By employing logit models, we are able to more reliably derive parameter estimates than if we relied on ordinary least squares estimation techniques.

4.2. 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 Section 3. To test the hypothesis, we employed ordinal logistic regression. Ordinal logistic regression is appropriate in this case, because the dependent variable is an ordered categorical variable. 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, Ervin, Gardner, & Schmitt, 2001; Jones, Johnson-Yale, Millermaier, & Perez, 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 ($e\beta = 1.92$, $p = 0.103$). Therefore, H1 is upheld.

4.3. Effects of privacy attitudes on privacy behavior

We collected two measures of privacy behavior, *privacy personalization* and *privacy customization*, as described in Section 3.2. 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, as the dependent variable in this case is a dichotomous representation of privacy behavior. For both H2a and H2b, regression was done twice, once using the combined measure of privacy attitudes as described in Section 3.2 and once using the individual items (e.g. identity theft, information leakage, hackers, blackmail, and cyberstalking). By running the regression with the scales and individual items, we are able to examine the validity of our construct, and identify individual privacy items that may particularly leverage the construct. This exploratory work will allow for refinement of the privacy scale in future research.

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.

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.

4.4. 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. Although our categories are naturally ordered, multinomial logit regression is appropriate because category distance (e.g. the difference between none/scanned and scanned/all) is not clearly defined. The multinomial logit allows for comparisons between categories, with scanning set as the reference category. 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.

4.5. 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) treated as a ordered categorical predictor, with the outcome being the combined measure of disclosure behavior described in Section 3.2. 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.

4.6. 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

Table 3
Elaboration of hypothesis 2.

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
Pseudo R^2	0.0088	0.0511	0.0140	0.0550

* Odds ratios. Standard errors in parentheses, $p < 0.05$.

Table 4
Relationship between privacy attitudes and privacy policy consumption.

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

Pseudo R^2: 0.0308.
* Odds ratios. Standard errors in parentheses, $p < 0.05$.
** Odds ratios. Standard errors in parentheses, $p < 0.01$.

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

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
Pseudo R^2	0.0222	0.0313

Cut points not reported to preserve space.
** Odds ratios. Standard error in parentheses, $p < 0.01$.
* Odds ratios. Standard error in parentheses, $p < 0.05$.

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.

4.7. 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 scale, privacy behaviors (personalization and customization, binary measures), and privacy policy consumption (an ordered categorical predictor) representing on disclosure behavior. We run the regression models for both the combined disclosure measure and for the two disclosure groups described in Section 3.2: *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

Table 6
Evaluation of the controlled model.

DV: amount of disclosed information	H6a Combined disclosure measure	H6b Identity-based disclosures	H6c Disclosures furthering contact
Privacy attitudes scale	0.513* (0.157)	0.827 (0.447)	0.575 (0.175)
Privacy personalization	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
Pseudo R^2	0.0551	0.0587	0.0265

Cut points not reported to preserve space.
** Standard error in parentheses, $p < 0.01$.
* Standard error in parentheses, $p < 0.05$.

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.

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).

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

Fig. 2 presents a visualization of the benefit model. At each stage of privacy customization and privacy policy consumption, an individual has a specific probability of being in a disclosure category. The visualization identifies the changes in probability associated with changes in privacy behavior. As we can see, individuals are more likely to be within the highest bands (highest disclosure categories) as they increase privacy customization. Conversely, individuals are more likely to be in the lower bands (lowest disclosure category) as they increase privacy policy consumption.

6. Discussion

Social network sites, such as Facebook, thrive on user-contributed content. However, many users report apprehension about

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

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.

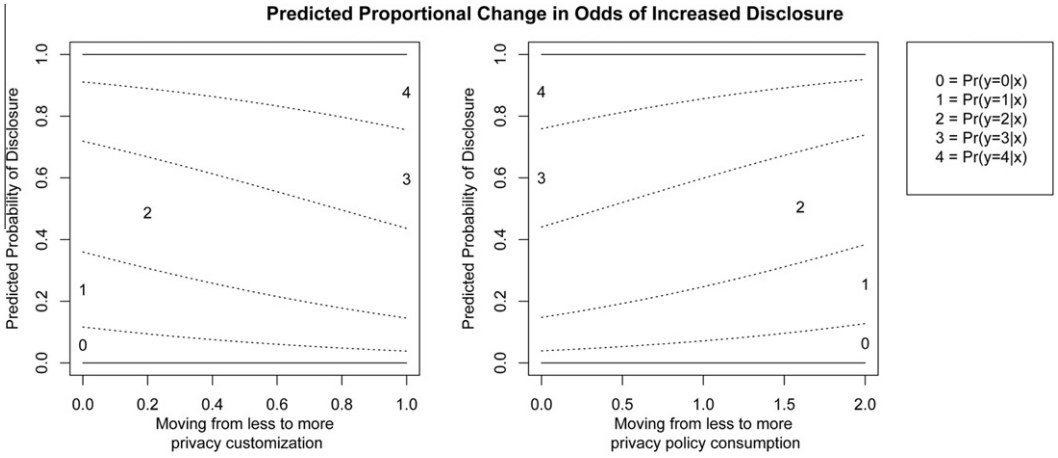

Fig. 2. Visualization of the benefit model. Size changes in the bands indicate shifts in the probability of disclosure group participation, as individuals shift privacy behaviors.

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) mediate 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.

Our research has implications for how privacy may be addressed in social network sites and relevance for research on disclosure and privacy in computer mediated communication. 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. Prior research has suggested that users may be more willing to disclose personal information (such as income level and postal code) to web sites that have privacy policies but that users are less likely to share information when they are also required to submit personally identifiable information such as name and address (Cranor et al., 2000). Our study showed a negative impact on disclosures on Facebook related to increased privacy policy consumption. Considered together, these results suggest that users need understandable privacy policies and usable privacy controls. While understanding privacy policies may increase concerns about disclosure, we describe below how the use of privacy controls may help users establish an environment in which they make increased disclosures. We suggest that simplifying privacy policies and their presentation (e.g. as suggested by Kelley, Cesca, Bresee, and Cranor (2010)), is an important part of helping users to feel more confident that they understand the range and implications of their disclosures.

Using predicted probabilities, we demonstrate that the use of privacy controls offsets the negative impact on disclosure of increased privacy policy consumption. The use of privacy controls in our study was linked with greater disclosure, which has important implications for both theory and design. Designers should create privacy controls that are easy for users to understand, and furthermore, users should be able to create their own rules within the system. This design implication strongly follows the logic of Petronio's rule-based-management system (2002). It is interesting to note that since our study was conducted, there were calls for simplified Facebook privacy controls, and Facebook responded with redesigned systems and interfaces[2].

Our results suggest 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. From the perspective of research on communication and disclosure, our work represents an application of boundary regulation theories to disclosure behaviors in social network systems, and this work supports both the substantive and predictive facets of this theory.

7. 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. Second, while we utilized a solicitation method that allowed access to a diverse population, a low response rate and evidence of potential gender bias due to nonresponse limits the generalizability of the results. Furthermore, unique aspects of the population studied (college students in the US Southeast) may limit generalization to other colleges. Third, this study only explores a single social network site, Facebook. It is therefore possible that contextual factors related to the specific site may limit generalizability to other social network sites. Fourth, 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.

This paper provides empirical support for the application of Boundary Regulation theories of privacy (e.g. Altman, 1975; Petronio, 2002) in socio-technical contexts. In the fast-changing domain of social media, the application of Boundary Regulation theories of privacy stands to provide a common ground for privacy researchers. 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 may share more freely in the social network.

References

Acquisti, A., & Grossklags, J. (2004). Privacy attitudes and privacy behavior: Losses, gains, and hyperbolic discounting. In J. Camp & R. Lewis (Eds.), *The economics of information security* (pp. 1–15). Kluwer Academic Publishers.

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

Acquisti, A., & Gross, R. (2009). Predicting social security numbers from public data. *Proceedings of the National Academy of Sciences, 106*(27), 10975–10980.

Adamic, L., Buyukkokten, O., & Adar, E. (2003). A social network caught in the web. *First Monday, 8*(6).

Ahern, S., Eckles, D., Good, N., King, S., Naaman, M., & 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* (pp. 357–366). New York, NY: ACM Press.

Altman, I. (1975). *The environment and social behavior*. Monterey, CA: Brooks/Cole.

[2] http://www.nytimes.com/2010/05/27/technology/27facebook.html.

Baym, N. K., & Ledbetter, A. (2009). Tunes that bind? *Information Communication & Society, 12*(3), 408–427.
Bonneau, J., & Preibusch, S. (2009). The privacy jungle: On the market for data protection in social networks. In *The eighth workshop on the economics of information security* (*WEIS 2009*).
Boyd, D. (2008). Facebook's privacy trainwreck. *Convergence, 14*(1), 13–20.
Cranor, L. F., Reagle, J., & Ackerman, M. S. (2000). Beyond concern: Understanding net users' attitudes about online privacy. In I. Vogelsang & B. M. Compaine (Eds.), *The internet upheaval: Raising questions, seeking answers in communications policy* (pp. 47–70). Cambridge, MA: MIT Press.
Derlega, V., & Chaikin, A. (1977). Privacy and self-disclosure in social relationships. *Journal of Social Issues, 33*(3), 102–115.
Dourish, P., & Anderson, K. (2006). Collective information practice. Exploring privacy and security as social and cultural phenomena. *Human–Computer Interaction, 21*(3), 319–342.
Erickson, T., & 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., & Donath, J. S. (2005). Homophyly in online dating: When do you like someone like yourself? In *Proceedings of CHI 2005*. ACM Press.
Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior, 25*(1), 153–160.
Grigoriadis, V. (2009). Do you own facebook? Or does facebook own you? *New York Magazine, 42*(12), 24–29.
Grimmelmann, J. T. (2009). Facebook and the social dynamics of privacy. *Iowa Law Review, 95*(4), 1–52.
Hawkey, K., & 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* (pp. 821–830). New York, NY, USA: ACM Press.
Herring, S. (2003). Gender and power in on-line communications. In J. Holmes & M. Meyerhoff (Eds.), *The Handbook of Language and Gender (pp. 202–228)*. Oxford University Press.
Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population: Differences in response quality. *Public Opinion Quarterly, 72*(5), 836–846.
Hinduja, S., & Patchin, J. W. (2008). Personal information of adolescents on the Internet: A quantitative content analysis of MySpace. *Journal of Adolescence, 31*(1), 125–146.
Hobgen, G. (October 25, 2007). *Security issues and recommendations for online social networks*. European Network and Information Security Agency. <http://www.enisa.europa.eu/>; 2008 Accessed 07.04.08.
Hosmer, D., & Lemeshow, S. (2000). *Applied logistic regression*. New York, NY: Wiley.
Iriberri, A., & 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., & Schmitt, N. (2001). Gender and the internet: Women communicating and men searching. *Sex Roles, 44*(5), 363–379.
Jernigan, C., & Mistree, B. (2009). Gaydar: Facebook friendships expose sexual orientation. *First Monday, 14*(10).
Joinson, A. N., Reips, U.-D., Buchanan, T., & Schofield, C. B. P. (2010). Privacy, trust, and self-disclosure online. *Human–Computer Interaction, 25*(1), 1–24.
Jones, S., Johnson-Yale, C., Millermaier, S., & Perez, F. S. (2009). US college students' internet use: Race, gender and digital divides. *Journal of Computer-Mediated Communication, 14*(2), 244–264.
Kelley, P. G., Cesca, L., Bresee, J., & Cranor, L. F. (2010). Standardizing privacy notices: An online study of the nutrition label approach. In *Proceedings of the 28th international conference on human factors in computing systems*, Atlanta, Georgia, USA, April 10–15, 2010.
Lampinen, A., Tamminen, S., & 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* (pp. 281–290). New York, NY, USA: ACM Press.
Lewis, K., Kaufman, J., & 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., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*(1), 415–444.
McDonald, A. M., & Cranor, L. F. (2009). The cost of reading privacy policies. *ISJLP, 4*, 543–897.
McDonald, A., Reeder, R., Kelley, P., & Cranor, L. (2009). A comparative study of online privacy policies and formats. In *Proceedings of privacy enhancing technologies* (pp. 37–55). Springer.
Mynatt, E. D., O'Day, V. L., Adler, A., & Ito, M. (1998). Network communities: Something old, something new, something borrowed. In *1998 Computer Supported Cooperative Work (CSCW)* (pp. 123–156). Springer.
Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington law review, 79*(1), 119–158.
Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford Law Books.
Nov, O., & 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* (pp. 333–336). New York, NY, USA: ACM Press.
Palfrey, J. (2008). Enhancing child safety and online technologies. In *Internet safety task force*. <http://cyber.law.harvard.edu/pubrelease/isttf/>; 2009 Accessed 10.01.09.
Petronio, S. (2002). *Boundaries of privacy: Dialectics of disclosure*. Albany, NY: State University of New York Press.
Stutzman, F., & Kramer-Duffield, J. (2010). Friends only: Examining a privacy-enhancing behavior in facebook. In *Proceedings of CHI 2010* (pp. 1553–1562).
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.
Tufekci, Z. (2008). Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science Technology and Society, 28*(1), 20–36.