JOURNAL OF COMPUTER-MEDIATED COMMUNICATION

Should I Share That? Prompting Social Norms That Influence Privacy Behaviors on a Social Networking Site

Erin L. Spottswood

Department of Communication, Portland State University, University Center Building, 520 SW Harrison Street, Portland, OR 97201

Jeffrey T. Hancock

Department of Communication, Stanford University, Building 120, Stanford, CA 94305

This study examines how explicit and implicit cues to social norms affect disclosure and privacy decisions in a Social Network Site (SNS) context. Study 1 revealed that participants' disclosure behavior adhered to explicit cues indicating disclosure frequency norms, while implicit social norm cues (i.e., surveillance primes) acted to increase overall disclosure frequency and affect disclosure accuracy when explicit cues discourage disclosure. Study 2 explored how these cues affected privacy-setting decisions and found that explicit cues indicating others' privacy settings could increase how strictly participants set their privacy settings, but the implicit cues had no effect. These results suggest that explicit cues about SNS norms can trigger bandwagon heuristic processing, and that, under limited circumstances, surveillance primes can affect self-disclosure.

Keywords: Privacy, Social Network Sites, Self-Disclosure, Surveillance Primes, and Social Psychology.

doi:10.1111/jcc4.12182

Deciding when and when not to disclose sensitive information has always been challenging, especially for those who join and post content onto Social Network Sites (SNSs). Privacy scholars have posited some factors that influence how people decide to share sensitive information (Acquisti, John, & Loewenstein, 2012; Nissenbaum, 2010; Petronio, 2002). For example, Petronio's (2002) Communication Privacy Management (CPM) theory posits that people decide to disclose sensitive information according to whether they think the recipient of their disclosures will or will not share their sensitive information with others. Although CPM can be used to explain how people manage their privacy in SNSs (Waters & Ackerman, 2011), it is less effective at predicting how users make disclosure and privacy-setting decisions when

Editorial Record: First manuscript received on May 7, 2015. Revisions received on November 5, 2015, March 30, 2016 and September 2, 2016. Accepted by Miriam J. Metzger on November 16, 2016. Final manuscript received on November 22, 2016.

Corrections made on 2/27/2017, after first online publications: Acknowledgements added to Scott Cambo, Jessie Taft, and the Cornell Social Media Lab.

they first join an SNS. This is due in large part to CPM's emphasis on perceptible, salient boundaries. CPM assumes that the discloser can identify who their listener is and what social rules are in play during an interaction. However, in situations where the discloser is unsure about who is listening and which rules are relevant, they may be hard pressed to determine whether it is safe to disclose various types of information.

This is the ambiguous privacy context that users face when they initially join an SNS. SNSs blur the lines between public and private communication (Baym & boyd, 2012), obscure who sees which posts and when (Litt, 2012), and have privacy policies that are difficult to navigate (Debatin, Lovejoy, Horn, & Hughes, 2009). As a result, users may disclose sensitive information in their profiles and select lenient privacy settings that fail their privacy goals. In the present research, we examine whether privacy-related behaviors could be influenced by visual cues on an SNS. We examine the effect of two types of cues: explicit cues, which display to the user the aggregate behavior of other users, and implicit cues, which present images that unconsciously prime a sense of surveillance. We report on two studies, the first examining the effect of these cues on the frequency and accuracy of self-disclosures, and the second examining their effect on the selection of privacy settings.

Study 1

Explicit Cues

Newcomers to an SNS are typically required to fill out a profile before they can browse the site and interact with others. Sometimes users disclose a great deal of information on their profiles, perhaps because they think providing more information helps them appear likable and friendly to others on the SNS (Hollenbaugh & Ferris, 2014). However, disclosing a lot of sensitive information on a profile can place them at risk for identity theft (Acquisti & Grossklags, 2005). As a result, users are motivated to learn an SNS's privacy norms so that they know how much to disclose to appear likable while avoiding the risks of overdisclosure. One way of learning an SNS's privacy norms is by continued use of the site (Lewis, Kaufman, & Christakis, 2008; Lewis, 2011; Stutzman & Kramer-Duffield, 2010). For example, Lewis and colleagues (2008) found that participants with private profiles tend to have friends and roommates who also have private profiles on the SNS Facebook. This suggests that users follow norms set by others when choosing privacy settings on SNSs. However, the origins of these norms are unclear. It is possible that these users talked with their friends and roommates offline about norms for the site, rather than deciding based on others' online behavior.

One possible method of helping newcomers learn about an SNS's privacy norms is by displaying visual cues on the site's pages that explicitly show current users' disclosure and privacy-setting behavior. According to Sundar's (2008) MAIN model, users attend to visual cues to help them evaluate content they encounter online. Sometimes these visual cues explicitly show an aggregation of previous users' thoughts and behaviors, which can trigger bandwagon heuristic processing (e.g., if other people think that something is good or safe, then I should too). For example, Apple lets users review and give "star" ratings to podcasts, and displays these ratings and reviews alongside every podcast on their application. The aggregation of star ratings is an explicit cue showing the quality of the podcast. This explicit cue can trigger bandwagon heuristic processing (e.g., if a lot of other users enjoy listening to this podcast, then I should listen to it too), motivating the user to listen to the podcast.

In this study, we define explicit cues as visual signals that display disclosure and privacy setting trends in an SNS context. Explicit cues that indicate disclosure trends should trigger bandwagon heuristics for privacy-related behaviors. For example, Acquisti and colleagues (2012) examined the effect of explicit cues showing an aggregation of previous users' disclosure behavior on an online survey. Some

participants were shown explicit cues indicating that a majority of previous survey takers had answered sensitive questions; others were shown that only a minority of previous survey takers had answered these questions; and the remaining participants were not shown any explicit cues. Consistent with the bandwagon heuristic, participants who were showed explicit cues indicating a majority of previous survey takers answered sensitive questions were far more likely to answer the sensitive questions themselves than participants who were shown explicit cues that only a minority had answered the same questions. In this view, the explicit cues triggered bandwagon heuristic processing and influenced disclosure behaviors in the survey in accordance with the norms described by the explicit cues.

Here we extend the application of the bandwagon heuristic (Sundar, 2008) to the SNS context, and replicate Acquisti and colleagues' (2012) prior observation that explicit cues can affect disclosure rates. This is important for two reasons. The first is a boundary extension of the disclosure comparative framework (Acquisti et al., 2012) to the SNS context. Second, given the recent replication controversy in social psychology and related disciplines (e.g., Open Science Collaboration, 2015), replication of effects has become more important. Thus, we hypothesized that explicit cues indicating disclosure rates affect disclosure frequency in an SNS context:

H1: Explicit cues that indicate high disclosure rates increase disclosure frequency and explicit cues that indicate low disclosure rates decrease disclosure frequency.

While research has shown that explicit cues can affect disclosure frequency (Acquisti et al., 2012), there is no research that has explored how explicit cues affect the accuracy of disclosures. Although a user may disclose more in order to adhere to a disclosure norm, they may alter the accuracy of their disclosures if they are concerned about their privacy. Sometimes people disclose inaccurate information in order to guard their sensitive information (DePaulo, Wetzel, Weylin Sternglanz, & Wilson, 2003). In fact, disclosing inaccurate information is a privacy-enhancing strategy when the discloser is unsure when or by whom their information could be accessed online (Fuster, 2010). Teenagers sometimes use coded or deceptive language on SNSs in order to keep the true meaning of their messages secret from their parents or other adults (Marwick & boyd, 2014). To the best of our knowledge, no studies have examined how explicit cues that influence disclosures affect the accuracy of those disclosures. How will a cue indicating that most users disclose a great deal of information on their profile affect how accurately they disclose the information on their profile? Given the lack of prior work on how explicit cues can affect the accuracy of self-disclosure, we pose a research question:

RQ1: How will explicit cues depicting disclosure frequency rates affect disclosure accuracy?

Implicit Cues

Although past research has examined how explicit cues affect disclosure behavior (Acquisti et al., 2012), to date there has been little examination of how implicit or unconscious cues can affect privacy behaviors on SNSs. Implicit cues or primes are stimuli that can affect a person's social behavior outside of their conscious awareness (for review see Bargh, 2006). One type of implicit cue that may be of particular relevance to disclosure behaviors on an SNS is the surveillance prime. Surveillance primes are images of eyes that promote prosocial behavior (for review see Nettle, Harper, Kidson, Stone, Penton-Voak, & Bateson, 2013). Surveillance primes are thought to unconsciously trigger the perception of being observed, which in turn promotes prosocial (i.e., kind, generous, honest, moral, etc.) behavior that is oriented towards gaining the acceptance and approval of the observer(s). For example, Bateson, Nettle,

and Roberts (2006) embedded surveillance primes or images of flowers in an office break room where employees would go for coffee or tea. Any employee who took milk from the break room for their coffee/tea were expected to contribute money to an "honesty box." Bateson and colleagues (2006) found that surveillance primes increased the amount of money that employees contributed to the honesty box, suggesting that the surveillance prime triggered the unconscious perception of being observed, which in turn prompted employees to never "forget" to pay for their milk.

Although research on surveillance primes suggests that they have the capacity to affect social behaviors, it is unclear whether they promote prosocial behavior or behavior that is consistent with a relevant, explicated social norm. For example, Bateson and colleagues (2006) displayed the prime on a sign that explicitly instructed employees to pay for milk directly above the "honesty box." Since the primes appeared on a sign that explicates a norm, it is not clear whether the effect is prosocial (i.e., kind, generous, moral behavior) or merely normative (i.e., the expected behavior in that context).

The present study seeks to clarify whether surveillance primes promote prosocial versus normative behavior by uncovering whether they nudge a normative behavior that is indicated by an explicit cue. For example, if being open and honest on an SNS profile is perceived as prosocial communication (Lampe, Ellison, & Steinfield, 2007), and surveillance primes implicitly promote prosocial behavior, then we should observe a main effect of surveillance primes such that disclosure frequency and accuracy will increase when surveillance primes are present:

H2A: Surveillance primes increase disclosure frequency and accuracy relative to when they are absent.

In contrast, if surveillance primes encourage adherence to norms, then we should expect an interaction between the explicit cues and surveillance primes such that when surveillance primes are present, participants will disclose more when the explicit cue indicates *most* users disclose a lot of information on their profiles and disclose less when the explicit cue indicates *few* users disclose a lot of information on their profile:

H2B: Explicit cues and surveillance primes will interact such that the effect of the explicit cues on disclosure frequency will be larger when surveillance primes are present than when they are absent.

Methods

Participants

This study used a convenience sample of undergraduate students from a university in the northeastern United States, recruited through research participation websites, paper flyers, and quarter cards distributed from 20 March to 30 April 2014. One hundred and fifty-eight participants took the survey, but 14 of these were removed for either entering jocular information or for mentioning the surveillance primes in their survey answers (the latter would undermine the unconscious process we were testing with the surveillance primes; Nettle et al., 2013). This resulted in a final N of 144. Participants' ages ranged from 19 to 26, with an average age of 21.25. Women were overrepresented in the data (68.9%), and participants mostly identified as either White (46.7%) or Asian (29.5%).

Procedure

Undergraduate students to participate in a "usability test" for a new SNS being developed by their university. After consenting to participate and reviewing the home page, participants were asked to fill out

a profile, select privacy settings, and complete a survey. The exit survey asked participants to share their thoughts about the site, rate the accuracy of their disclosures, and answer some demographic and debriefing questions. Finally, they were thanked for their time and given compensation.

Information Sensitivity

We conducted a pilot study to ensure that the information being asked for on the profile would be perceived as more or less sensitive in this context. Forty-five different undergraduate students from a university in the northeastern United States rated how appropriate it would be to post 42 different kinds of information according to 1) how appropriate it would be to share that piece of information in a profile for a university-affiliated SNS; 2) how comfortable they would be sharing that piece of information in a profile for a university-affiliated SNS; 3) how private they perceive that piece of information to be; and 4) how sensitive they perceive that piece of information to be. Each item ranged on a scale from 1 (extremely sensitive, etc.) to 5 (not at all sensitive, etc.). The items had high reliability (Cronbach's *alphas* all higher than 0.76) and factor analysis revealed that the profile items were tapping one underlying factor [X^2 (1, $N = 45$) = 120.57, $p = 0.10$], so we averaged the items into a single perceived sensitivity score.

A frequency analysis of the perceived sensitivity score revealed that the data were normally distributed between the minimum (1) and maximum (5) values, ($M = 2.82$, $SD = 1.13$). Also, 10 of the 42 information types were below the bottom quartile (1.86) and 9 of the 42 information types were above the top quartile (3.91). As expected, the sensitive (e.g., "medical information") ($M = 1.44$, $SD = 0.07$), moderately sensitive (e.g., "relationship status") ($M = 2.91$, $SD = 0.10$), and nonsensitive (e.g., "graduation year") items ($M = 4.41$, $SD = 0.12$) were significantly different from each other, $F(1, 42) = 1700.00$, $p < 0.001$ $\eta 2 = 0.78$. These results suggest that users perceive information with sensitive attributes as inappropriate to disclose in this context.

The Context

The SNS used in this study was designed to look and feel like an actual SNS hosted by the students' university. This context featured pages that are common on other SNSs in order to enhance its realism. For example, when participants opened up the site, they were first taken to a "homepage" that featured mock updates as well as images taken from the university's websites. This was followed by a profile page that contained 18 fields (presented in random order) where participants were asked to enter their information (Figure 1). Data from the pilot study was used to ensure that the 18 information fields ranged in sensitivity. We recorded participants' hometown, major, graduation year, relationship status, academic accomplishments, and college affiliation in order to assess whether participants were entering credible responses or jocular information (e.g., entering "Tatooine," a primary setting of the *Star* Wars film franchise, as their hometown would be considered a jocular response). These recorded responses were manually coded as either being filled out or left blank. Participants who entered jocular information were deleted from the data. All of the other disclosures (i.e., information participants posted onto their profile) were not recorded but automatically coded as having been entered (1) or left blank (0).

After the profile page came a privacy settings page where participants were asked to make selections for five privacy settings. The privacy settings included: "Who can see my profile?"; "Who can see my updates and posts?"; "Who can see the posts that you are tagged in?"; "Who can look you up using the email you provided?"; and "Who can look you up using the phone number you provided?" The response options were (1) – "Everyone," (2) – "Students, faculty, alumni, and staff," (3) – "Students only," or (4) - "Only me" (Figure 1).

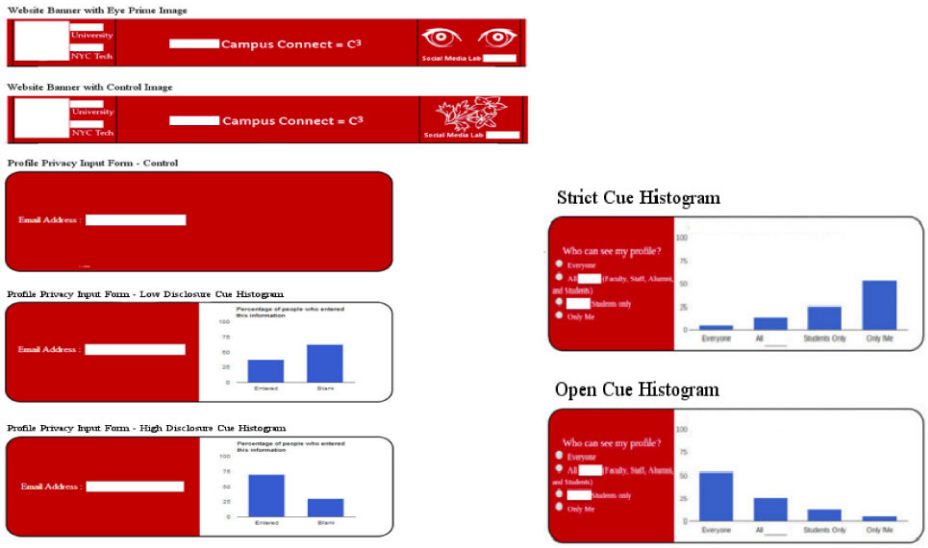

Figure 1 Surveillance Primes and High Versus Low Explicit Cues

Explicit Cues

Following prior work (Acquisti et al., 2012), explicit cues depicting how frequently other uses disclose information were presented via histograms (Figure 1). The high-disclosure condition featured histograms that indicated most users had disclosed that piece of information on their profiles, while the low-disclosure condition featured histograms that indicated few users had disclosed that information on their profiles. The control condition did not display any histograms.

Surveillance Primes

We embedded a surveillance prime or a flower (i.e., the control image) in the SNS's logo on the right-hand side of the banner at the top of each SNS page (Figure 1). The banner was fixed so that even when participants scrolled down the profile page or privacy settings page, the prime or control image was still visible.

Dependent Measures

Participants' disclosure frequency was calculated by tallying how often participants entered information onto the profile page. During the follow-up survey, participants were shown their previous profile disclosures and asked to indicate if their entry was accurate or inaccurate. These accuracy ratings were used to calculate how often participants accurately entered information onto the profile page. The other dependent measure was how private the user made their privacy settings. The data from this measure are described in Study 2.

Results

This study employed a 3 (sensitivity: low vs. medium vs. high) X 3 (explicit cue: high-disclosure cue vs. low-disclosure cue vs. control) X 2 (prime: surveillance prime vs. control image) mixed design, with

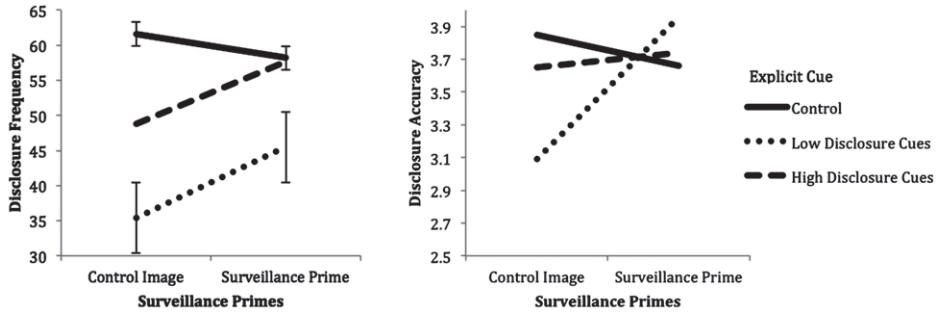

Figure 2 Effect of No, Low, and High Disclosure Explicit Cues and Surveillance Primes on Disclosure Frequency and Accuracy in Study 1

disclosure frequency and accuracy as the primary dependent variables. Given the nonindependence of the data, in which each participant responded to sensitive, moderately sensitive, and nonsensitive disclosure items, disclosure responses were nested within participants in a linear mixed model (SPSS v23). Disclosure behavior (i.e., frequency or accuracy) was entered as the dependent variable, and sensitivity (within-subjects factor), explicit cue (between-subjects factor), and prime (between-subjects factor) were entered as independent variables.

Disclosure Frequency

Explicit cues significantly affected disclosure behavior, $F(2, 348) = 20.67$, $p < 0.001$, $\eta 2 = 0.03$. Pairwise comparisons (*Bonferroni* corrected) revealed that participants disclosed less frequently in the low-disclosure ($M = 40.4$, $SE = 0.23$) compared to the high-disclosure ($M = 53.3$, $SE = 0.02$) condition ($p < 0.01$, $\eta 2 = 0.03$). Participants also disclosed less frequently in the low-disclosure ($M = 40.4$, $SE = 0.23$) compared to the control ($M = 59.9$, $SE = 0.02$) condition ($p < 0.01$, $\eta 2 = 0.01$) (Figure 2). This pattern of results is consistent with our prediction that explicit cues to disclosure frequency trigger bandwagon heuristics (H1), with participants' disclosures adhering to patterns of prior users.

We next examined how the surveillance primes affected disclosure frequency. Consistent with the prediction that surveillance primes would increase prosocial behavior (H2A), there was a main effect of surveillance primes on disclosure frequency, $F(1, 348) = 3.99$, $p < 0.05$, $\eta 2 = 0.02$. Participants disclosed more frequently when surveillance primes were present ($M = 53.8$, $SE = 0.02$) than when they were absent ($M = 48.6$, $SE = 0.02$) (Figure 2). The contrasting hypothesis (H2B) predicted an interaction effect in which surveillance primes will enhance the effect of the explicit cues on disclosure frequency. The data were inconsistent with this prediction; there was no interaction effect on disclosure frequency $F(2, 338) = 0.3$, $p = 0.86$, $\eta 2 = 0.01$. Taken together these data suggest that the surveillance primes increased prosocial disclosure frequency rather than enhancing the operation of the explicit cues.

Disclosure Accuracy

Our research question (RQ1) asked whether participants might react to explicit cues about disclosure frequency by modifying the accuracy of their disclosures. There was no main effect of the explicit cues on disclosure accuracy. The effect of surveillance primes on accuracy approached but did not achieve significance, $F(2, 338)^1 = 3.55$, $p = 0.06$, $\eta 2 = 0.0001$. Participants reported being slightly more accurate when surveillance primes were present ($M = 3.79$, $SE = 0.10$) than when they were absent ($M = 3.53$, $SE = 0.09$).

There was, however, an unexpected interaction effect on disclosure accuracy, $F(2, 338) = 4.65$, $p < 0.05$, $\eta 2 = 0.03$. Decomposing this interaction by selecting for cases in the low- and high-disclosure conditions revealed that the presence ($M = 3.93$, $SE = 0.19$) versus absence ($M = 3.06$, $SE = 0.19$) of surveillance primes increased accuracy in the low-disclosure condition, $F(1, 103) = 10.84$, $p < 0.001$, $\eta 2 = 0.07$. In contrast, the presence ($M = 3.87$, $SE = 0.18$) versus absence ($M = 3.61$, $SE = 0.15$) of surveillance primes did not affect accuracy in the high-disclosure condition $F(1, 102) = 0.14$, $p = 0.71$, $\eta 2 = 0.0001$.

Discussion

The findings from Study 1 revealed that both explicit and surveillance primes on the site influenced disclosure frequency in an SNS profile. Not only does this replicate and extend Acquisti and colleagues' (2012) prior work to the context of SNS disclosures, these findings are also consistent with the MAIN model's bandwagon heuristic regarding how people process explicit cues online (Sundar, 2008).

The effects of the surveillance primes were more subtle, but when they were present they increased disclosure frequency overall, and they increased disclosure accuracy when low-disclosure explicit cues were also present on the profile page. Together these findings suggest that they encourage prosocial communication on SNSs, although consistent with prior work the effects were small (Nettle et al., 2013).

In addition to disclosure behaviors, privacy settings are another way that users can regulate access to their information in an SNS. Study 2 explores how explicit cues and surveillance primes affect privacy setting decisions and subsequent disclosure behavior.

Study 2

How do people decide whether to select strict or open privacy settings in SNSs? One possible framework that may shed light on this question is Nissenbaum's (2010) framework of contextual integrity, which posits that transmission principles affect disclosure behavior both offline and online. Transmission principles are implicit norms that place constraints on how information should be shared, prescribing when the transfer of various kinds of information is appropriate. For example, when close friends disclose sensitive information to each other face-to-face, they usually adhere to a transmission principle that discourages either from sharing the other's sensitive information outside of the original disclosure context. However, when close friends interact in an SNS, there is less certainty that the same transmission principles will be upheld because access to that information depends on privacy settings (Acquisti & Grossklags, 2005).

Selecting appropriate privacy settings can help users adhere to relevant transmission principles given that they help users regulate who can access their disclosures in an SNS context. For example, users often select strict privacy settings in order to ensure their parents or employers will be restricted from seeing their posts on Facebook (Debatin et al., 2009). However, stricter privacy settings are certainly not a guarantee for preventing unauthorized individuals from seeing a sensitive disclosure (Brandimarte, Acquisti, & Loewenstein, 2013).

Following Study 1, we expect explicit cues about privacy-setting norms to trigger bandwagon heuristics that prompt users to adhere to the norms described in the explicit cues. Indeed, prior work establishes that people base their privacy-setting decisions on perceived privacy-setting norms (Lewis et al., 2008; 2011). We therefore predict that explicit cues that indicate how others set their privacy settings will influence how users set their privacy settings on an SNS:

H1: Explicit cues that indicate other users have adopted strict privacy settings will increase the selection of strict privacy settings compared to explicit cues that indicate other users have adopted open privacy settings.

Surveillance Primes and Privacy-Setting Decisions

Several social psychology studies, and the results from Study 1, suggest that surveillance primes enhance prosocial behaviors. Previous studies have found that selecting open privacy settings implies a user is open, sociable, and eager to interact with others (Lampe et al., 2007), suggesting that selecting open privacy settings could be considered prosocial in an SNS context. If that is the case, then users should select more open privacy settings when surveillance primes are present in this context.

H2: Surveillance primes increase open privacy-setting selection relative to when they are absent.

Effects of Privacy Settings on Disclosure

The control paradox posits that users sometimes disclose a great deal of sensitive information online because they assume their privacy settings will protect who can access their information, when in fact unauthorized third parties can still access and share their information (Brandimarte et al., 2013). Moreover, most SNSs require users to fill out a profile before selecting privacy settings, encouraging users to consider their self-presentation goals before their privacy needs. In contrast, requiring users to select privacy settings first may highlight transmission principles rather than social goals, which in turn should affect how much information they disclose in this context. We compare the disclosure behavior of participants in Study 1, who filled out a profile before deciding on their privacy settings, with that of participants in Study 2, who decided on their privacy settings before filling out a profile:

H3: Participants disclose more information if they select privacy settings before they disclose information on a profile than if they select privacy settings after disclosing information.

Methods

Participants

A different set of undergraduate students from the same northeastern U.S. university as Study 1 were recruited via research participation websites, paper flyers, and quarter cards. Students who participated in the pilot study or Study 1 were barred from participating in this study. Recruitment lasted from 2 May to 21 May 2014. One hundred and twelve participants took the survey, but six participants who mentioned the surveillance primes in their survey were removed, resulting in a final N of 106. Participants' ages ranged from 19 to 26, with an average age of 21.33. There were slightly more women (56%) than men, and participants mostly identified as either White (43.0%) or Asian (36.4%).

System

The SNS used in Study 2 was similar to the SNS used in Study 1 except for three key changes. First, the order of the profile and privacy setting pages was reversed. Second, the explicit cues and surveillance primes were featured on the privacy settings page rather than the profile page. Third, the explicit cues indicated privacy setting norms rather than disclosure frequency norms.

Procedure

The procedure for Study 2 was the same as the procedure used in Study 1, except participants were asked to select privacy settings before filling out a profile.

Explicit Cues

The strict privacy condition featured histograms that indicated most users had selected strict privacy settings (e.g., 54% had selected "Only Me," 26% has selected "Students Only," etc.), whereas the open-privacy condition featured histograms that indicated most users had selected more open privacy settings (e.g., 54% had selected "Everyone", and 26% had selected "Students, faculty, alumni, and staff," etc., Figure 1). The control condition did not include any histograms.

Surveillance Primes

The same surveillance prime and control images used in Study 1 were used in Study 2.

Dependent Measures

Following the procedure for Study 1, privacy settings were calculated by tallying how often participants selected open versus strict privacy settings for the five items. Participants' decisions were reliable (Cronbach's $\alpha = 0.90$, $M = 2.32$, $SD = 0.83$) and were averaged to yield a single privacy setting factor, from open = 1 to strict = 4. Disclosure frequency and accuracy was calculated by computing on average how frequently and accurately participants disclosed overall.

Results and Discussion

This study employed a 3 (explicit cues: strict vs. open vs. control) X 2 (prime: surveillance prime vs. control image) between subjects design with privacy-setting decisions as the primary dependent variable. Recall that each participant's privacy-setting decisions were averaged into a single privacy-setting score. We entered privacy settings as the dependent variable, and explicit cue (between-subjects factor) and prime (between-subjects factor) as independent variables in a general linear model (SPSSv23).

Privacy settings

Explicit cues significantly affected how strict or open participants set their privacy settings on the SNS, $F(2, 98) = 3.11$, $p < 0.05$, $\eta 2 = 0.06$ (Figure 3). Participants selected stricter settings in the strict ($M = 2.51$, $SE = 0.13$) relative to the open ($M = 2.01$, $SE = 0.15$) conditions. This pattern of results is consistent with our prediction that explicit cues about privacy settings trigger bandwagon heuristics (H1) and can prompt users to follow transmission principles (Nissenbaum, 2010) in an SNS context.

The surveillance primes had no effect on privacy setting behavior $F(2, 98) = 0.19$, $p = 0.66$, $\eta 2 = 0.0001$, failing to support H2.

Effects of Privacy Settings on Disclosure Behavior

Given that the samples from Studies 1 and 2 were drawn from the same university student population, we compared the effect of privacy settings on disclosure behavior across the two samples, which allowed us to examine the effect of presentation order on disclosures. Recall that in Study 1 participants first filled out their disclosure information and then their privacy settings. In contrast, in Study 2 participants decided on their privacy settings before completing their disclosures.

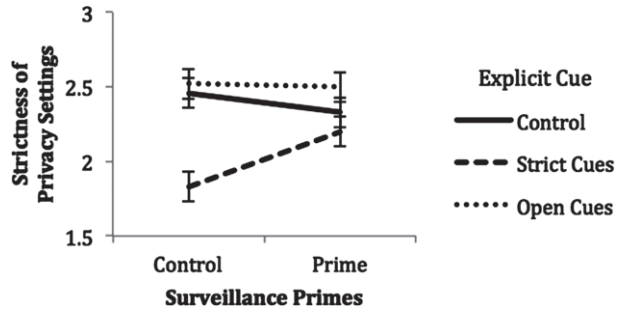

Figure 3 Effect of No, Strict, and Open Privacy Setting Explicit Cues and Surveillance Primes on Privacy Setting Decisions in Study 2

The analysis of disclosure frequency by presentation order (selecting privacy settings first vs. filling out a profile first) revealed that participants disclosed more information when they selected privacy settings ($M = 53.8$, $SE = 0.02$) before completing a profile than when they completed a profile ($M = 48.4$, $SE = 0.02$) before selecting privacy settings, $F(1, 682) = 3.95$, $p < 0.05$, $\eta 2 = 0.01$. A second analysis with disclosure accuracy as the dependent variable revealed that accuracy was unaffected by the placement of the privacy settings page, $F(1,226) = 0.12$, $p = 0.73$, $\eta 2 = 0.0001$. These results replicate and extend the control paradox, and are consistent with H3. However, the paradox did not extend to the accuracy of the disclosures.

General Discussion

The goal of these studies was to examine how explicit cues and surveillance primes can influence how users share and protect their information on SNSs. Study 1 revealed that explicit cues can affect disclosure frequency[2], and that surveillance primes can affect how accurately users disclose information on an SNS. Study 2 revealed that explicit cues, but not surveillance primes, can modify privacy setting decisions on an SNS. These results make several important contributions, including 1) testing how explicit cues can trigger bandwagon heuristic processing that affects users' disclosure and privacy setting behavior, 2) providing an important replication of Acquisti and colleagues (2012) comparative privacy concept and extending the effects to the SNS context, 3) introducing the subtle role that surveillance primes can have on privacy behaviors, and 4) providing important implications for the contextual integrity framework (Nissenbaum, 2010) and the control paradox (Brandimarte et al., 2013).

Explicit Cue Effects on Privacy in SNS Contexts

Disclosure and privacy-setting norms have been found to influence user behavior on Facebook (Lewis et al., 2008; 2011; Utz & Krämer, 2009), but it is still unclear how these norms are learned. The findings from Studies 1 and 2 suggest that explicit cues displaying disclosure and privacy-setting trends can influence how newcomers learn about an SNS's privacy norms. These data are consistent with observations by Sundar (2008) that "the prominence of others as sources and the bandwagon heuristic they cue with their implicit endorsement" can affect user behavior in a variety of computer-mediated contexts (p. 84). Seeing that most users have disclosed their phone number on an SNS profile, for example, may trigger thoughts such as "if most users think it is a good idea to disclose their phone number on their profile, then I should too." The primary reason to follow social norms is to fit in and gain others'

approval, and online users even sanction other users who do not adhere to SNS norms (McLaughlin & Vitak, 2012). Future research should explore how not following explicit cues that indicate disclosure and privacy setting trends affect the way current users view those who violate these expectations in an SNS context.

This replication and extension of Acquisti et al. (2012) paints a clearer picture about how people make privacy decisions online, pointing to the importance of replication in communication and technology research. Recent research has drawn renewed attention to the importance of replication and extension in the social sciences. Nosek collaborated with 269 colleagues to replicate 100 published psychology studies (Open Science Collaboration, 2015). Only a fraction of the replications yielded significant results and the replications that did achieve significance still had very small effect sizes. This suggests that the tendency to prioritize novelty undermines researchers' ability to uncover the fundamental processes that guide online disclosure and privacy behavior. As more people incorporate SNSs into their lives, we need more research to understand what fundamentally influences their disclosure and privacy-setting behaviors. Replicating Acquisti and colleagues' (2012) methods in an SNS context shows how powerful explicit cues can be on users' disclosure and privacy setting behavior on these sites. We hope that future work will continue to use explicit cue methodology to uncover its effects on how people disclose, withhold, and protect their information online.

Priming Privacy

To the best of our knowledge, this is the first research to examine how surveillance primes can affect privacy decisions in SNSs. The effects of the surveillance primes suggest that cues that are outside of a person's conscious awareness can subtly affect privacy decisions. Recent research has begun to show that privacy decisions are not always rational; they can be affected by a person's feelings or immediate desires (Acquisti & Grossklags, 2005). The present study takes this one step further and demonstrates that they can be influenced by cues outside of a person's conscious awareness, although the effects are weak. The findings from Study 1 demonstrated that unconscious cues such as surveillance primes increase how frequently users disclose information in an SNS profile. This suggests that a variety of unconscious factors may affect how users disclose information in SNSs, potentially affecting their ability to adequately protect their sensitive information on these sites. Future research is required to explore how unconscious cues affect SNS disclosure behavior in order to uncover the myriad ways privacy decisions may be influenced by primes.

The findings from Study 1 also suggest that surveillance primes unconsciously trigger behavior that is prosocial rather than normative. Other work suggests that disclosing frequently and accurately on a profile is considered prosocial in an SNS context (Gosling, Gaddis, & Vazire, 2007; Lampe et al., 2007), which may explain why surveillance primes increased disclosure accuracy in the low-disclosure frequency cue condition. Participants may have compensated for adhering to the low-disclosure frequency cue, which is not prosocial, by disclosing more accurately. Future research needs to replicate this finding and explore whether surveillance primes not only promote prosocial communication but also potentially decrease antisocial communication (e.g., flaming or trolling) in SNS and other online contexts.

One important question is why the surveillance primes influenced the accuracy of participants' disclosures in Study 1, but failed to affect privacy setting decisions in Study 2. Given that the means were in the expected direction (participants selected more open settings when eye primes were present than when they were absent in the strict cues condition, see Figure 3), one possibility for this failure may have been a lack of power. Another possibility is that making decisions for privacy settings falls outside the scope of behaviors relevant for surveillance primes. Privacy-setting decisions are more abstractly related to communication behaviors, such as disclosures, and may therefore not be susceptible to the influence

of behavioral primes. The results here, nonetheless, suggest that surveillance primes do not affect privacy setting decisions to the same degree as disclosure behavior.

Contextual Integrity and Privacy Behaviors

This is the first study to provide an empirical test of behavioral data that are consistent with the Nissenbaum's (2010) transmission principle. Recall that transmission principles place constraints on how information can or should be shared, prescribing when the transfer of information is appropriate. We reasoned that if privacy settings are used to help ensure that relevant transmission principles are upheld in an SNS, it is likely that explicit cues prescribing strict or open privacy-setting norms should affect privacy-setting behavior in an SNS. We found an asymmetry in the effect of the explicit cues for privacy settings, with explicit cues for selecting stricter settings producing a larger effect than cues for more open settings (Figure 3). That is, explicit cues were more effective at encouraging users to share less information than they were at encouraging users to share more. This suggests that users may adopt disclosure and privacy-setting norms if they perceive following these norms will help them uphold the site's transmission principles.

Control (Paradox) by Design

According to Norberg and colleagues (2007), trust plays an important role when deciding it is safe to disclose personal or sensitive information in a given context. The control paradox posits that the more control users feel that they have over their information, the more willing they are to share sensitive information in an SNS context (Brandimarte et al., 2013). This is likely because users trust their privacy settings will adequately bar unauthorized third parties from accessing their information. By merging the data from Studies 1 and 2, we explored how selecting privacy settings before filling out a profile influences disclosure behavior, and found that participants who selected privacy settings before filling out a profile disclosed more information than participants who filled out a profile before selecting privacy settings. It is possible that participants who selected privacy settings first felt that their settings would effectively protect their information, and therefore felt comfortable disclosing more information on their profile. Future research should explore if indeed a sense of control is the mechanism that drives why stricter privacy settings lead to more disclosures in SNS profiles.

Limitations

One important limitation is that these studies used an SNS design that, while custom-made and novel, presented as a university-affiliated tool. This may have caused participants to only disclose information that seemed relevant to the university context (e.g., disclosing their major and which clubs they belonged to rather than their past relationships). Also, the number of participants in this study resulted in rather low power for detecting effects given our experimental design and future studies should recruit larger samples. Relatedly, the effect sizes were small, especially for the surveillance primes. While the small but reliable effects for the explicit cues represent an important replication, the surveillance prime effects are novel and warrant replication. Finally, although this study used undergraduate participants, SNS user populations tend to be diverse (Duggan & Brenner, 2013). Future research should explore how explicit cues and primes affect disclosure and privacy setting behavior of diverse populations of actual SNS users.

Conclusion

Subtle changes to an SNS's design can increase the amount of sensitive information that users disclose on SNSs. Findings from Studies 1 & 2 suggest that explicit cues can affect disclosure and privacy-setting

behaviors on SNSs. The findings from Study 1 also suggest that surveillance primes, under limited circumstances, can influence disclosure decisions on SNSs. Overall, these results suggest that people strive to adhere to privacy norms in SNSs, and that explicit cues, primes, and even the sequence of information disclosure can significantly affect privacy-related behavior on these sites.

Acknowledgements

The authors wish to thank Scott Cambo, Jessie Taft, and the Cornell Social Media Lab for their generous support of this research.

Notes

1 There were 10 participants who did not rate the accuracy of their profile disclosures, hence the 10-point difference in degrees of freedom between disclosure rate and accuracy.
2 Although the high disclosure explicit cue increased disclosure frequency relative to the low disclosure frequency cue in Study 1, participants disclosed the most information in the control condition. This may be because even though the high disclosure explicit cues indicated that most users had disclosed a lot of information, they may have also hinted at a minority who feel it is unsafe to disclose a lot of information in this context. This in turn may have prompted participants to be somewhat wary of how much they should disclose in this context. However, the finding that participants disclosed most frequently in the control condition denotes a tendency to disclose frequently on an SNS profile when no cues suggest users should do otherwise. Future research should continue to explore how users can be reminded of their privacy concerns via disclosure frequency cues as well as policy reminders (Metzger, 2006).

References

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy*, (1), 26–33. doi: 10.1109/MSP.2005.22

Acquisti, A., John, L. K., & Loewenstein, G. (2012). The impact of relative standards on the propensity to disclose. *Journal of Marketing Research*, 49(2), 160–174. doi: 10.1509/jmr.09.0215

Bargh, J. A. (2006). What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior. *European Journal of Social Psychology*, 36(2), 147–168. doi: 10.1002/ejsp.336

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real world setting. *Biology Letters,* 2(3), 412–414. doi: 10.1098/rsbl.2006.0509

Baym, N. K., & boyd, d. (2012). Socially mediated publicness: An introduction. *Journal of Broadcasting & Electronic Media*, 56(3), 320–329. doi: 10.1080/08838151.2012.705200

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2013). Misplaced confidences: Privacy and the control paradox. *Social Psychological and Personality Science*, 4(3), 340–347. doi: 10.1177/1948550612455931

Debatin, B., Lovejoy, J. P., Horn, A. K., & Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication*, 15(1), 83–108. doi: 10.1111/j.1083-6101.2009.01494.x

DePaulo, B. M., Wetzel, C., Weylin Sternglanz, R., & Wilson, M. J. W. (2003). Verbal and nonverbal dynamics of privacy, secrecy, and deceit. *Journal of Social Issues*, 59(2), 391–410. doi: 10.1111/1540-4560.00070

Duggan, M., & Brenner, J. (2013). *The demographics of social media users.* Pew Research Center. Retrieved from http://www.pewinternet.org/2013/02/14/the-demographics-of-social-media-users-2012/

Fuster, G. G. (2010). Inaccuracy as a privacy-enhancing tool. *Ethics and Information Technology*, 12(1), 87–95. doi: 10.1007/s10676-009-9212-z

Gosling, S. D., Gaddis, S., & Vazire, S. (2007). Personality impressions based on Facebook profiles. *ICWSM*, 7, 1–4.

Hollenbaugh, E. E., & Ferris, A. L. (2014). Facebook self-disclosure: Examining the role of traits, social cohesion, and motives. *Computers in Human Behavior*, 30, 50–58. doi: 10.1016/j.chb.2013.07.055

Lampe, C., Ellison, N., & Steinfield, C. (2007). A familiar Face(book): Profile elements as signals in an online social network. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 435–444. doi: 10.1145/1240624.1240695

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. doi: 10.1111/j.1083-6101.2008.01432.x

Lewis, K. (2011). The co-evolution of social network ties and online privacy behavior. In S. Trepte & L. Reinecke (Eds.), *Privacy online* (pp. 91–109). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-21521-6

Litt, E. (2012). Knock, knock. Who's there? The imagined audience. *Journal of Broadcasting & Electronic Media*, 56(3), 330–345. doi: 10.1080/08838151.2012.705195

Marwick, A. E. & boyd, d. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, 16(7), 1051–1067. doi: 10.1177/1461444814543995

McLaughlin, C., & Vitak, J. (2012). Norm evolution and violation on Facebook. *New Media & Society*, 14(2), 299–315. doi:10.1177/1461444811412712

Metzger, M. J. (2006). Effects of site, vendor, and consumer characteristics on web site trust and disclosure. *Communication Research*, 33(3), 155–179. doi: 10.1177/0093650206287076

Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I. S., & Bateson, M. (2013). The watching eyes effect in the dictator game: It's not how much you give, it's being seen to give something. *Evolution and Human Behavior*, 34(1), 35–40. doi: 10.1016/j.evolhumanbehav.2012.08.00

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life.* Stanford, CA: Stanford University Press.

Norberg, P. A., Horne, D. R., & Horne, D. R. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *The Journal of Consumer Affairs,* 41(1), 100–126. doi: 10.1111/j.1745-66606.2006.00070.x

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science,* 349(6251), aac4716. doi: 10.1126/science.aac4716

Petronio, S. (2002*). Boundaries of privacy: Dialectics of disclosure.* Albany, NY: State University of New York.

Sundar, S. S. (2008) The MAIN Model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), Digital media, youth, and credibility *(pp. 73–100)*. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press.

Stutzman, F., & Kramer-Duffield, J. (2010, April). Friends only: Examining a privacy-enhancing behavior in Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1553–1562). ACM.

Utz, S., & Krämer, N. (2009). The privacy paradox on social network sites revisited: The role of individual characteristics and group norms. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 3(2), 2. Retrieved from: http://cyberpsychology.eu/view.php?cisloclanku=2009111001&article=1

Waters, S., & Ackerman, J. (2011). Exploring privacy management on Facebook: Motivations and perceived consequences of voluntary disclosure. *Journal of Computer-Mediated Communication*, 17(1), 101–115. doi: 10.1111/j.1083-6101.2011.01559.x

About the Authors

Erin L. Spottswood is an assistant professor in the Department of Communication at Portland State University where she examines how subtle cues embedded in online environments influence interpersonal communication dynamics. Her work has appeared in *Human Communication Research*, *Computers in Human Behavior*, and *Media Psychology*.
Address: University Center Building 520 SW Harrison Street, Suite 440 Portland, OR 97201. E-mail: spot@pdx.edu

Jeffrey T. Hancock is Professor in the Department of Communication at Stanford University and the Director of the Stanford Center for Computational Social Sciences. Hancock studies the psychological dynamics of social media.
Address: Building 120, Stanford University, Stanford, CA 94305. E-mail: jeff.hancock@stanford.edu