

To Post or Not to Post: The Effects of Persuasive Cues and Group Targeting Mechanisms on Posting Behavior

Bruce Ferwerda
Computational Perception
Johannes Kepler University
bruce.ferwerda@jku.at

Markus Schedl
Computational Perception
Johannes Kepler University
markus.schedl@jku.at

Marko Tkalcić
Computational Perception
Johannes Kepler University
marko.tkalcić@jku.at

ABSTRACT

When social network sites (SNSs) users intend to share content, they need to estimate the appropriateness of the content for their audience. Wrongly made estimations can result in regret about the posted content. A common strategy for users to minimize regret is to self-censor content. However, this also means that content that would have been safe to share may be left unshared. To solve sharing problems, SNSs have been focusing on improving group targeting mechanisms to give users more control over their content. As users still need to estimate the content appropriateness themselves, we asked whether improving these mechanisms is really the solution. We hypothesized that users' posting decisions consist of uncertainty and therefore providing guidance on whether it is safe to post would be more beneficial. To answer this we conducted two studies. In Study A we identified what kind of content users are self-censoring and what the reasons are. Study B was used to test and compare different solutions to limit the self-censored content found in Study A. We created a persuasive cue that predicted how the user's audience would possibly respond to the content and compared this with the effects of a group targeting mechanism. Among 215 participants we found that posting decisions consist of uncertainty and that persuasive cues are a more effective means to limit self-censorship, but can also warn users of content that is not safe to post. Making use of such cue can improve SNSs' sociability and reduce regret of wrong posting decisions.

I INTRODUCTION

Social network sites (SNSs) are rapidly growing with the adoption spreading out across a wide audience [8]. This continuous adoption have been changing online social practices and experiences [12]. The encouragement of sociability resulted in a growing social diversity in the user's social network. SNSs like Facebook have become an "all-friends-in-one-place" solution, meaning a larger number of social connections with a mix of strong and weak ties [9].

The trend of becoming an all-friends-in-one-place environment can create a "privacy dilemma," i.e., a conflict between the privacy needs of individuals and the need for sociability and content sharing. When privacy is protected, sociability and content sharing will be compromised and vice versa. In both cases the outcome is undesirable [9]. Previous research argued that these problems are exacerbated because users are experiencing difficulties to share content with specific groups in their social network [9, 17, 22]. As a result, users self-censor to anticipate regret of wrongly posted content [24]; they adjust their posting or eventually decide not posting at all. Sleeper et al. [22] found that reasons for self-censoring behavior are mainly concentrated around the anticipated feelings or opinions of the user's audience (e.g., not wanting to start an argument, worried to offend or hurt someone, feeling that the content would be boring or redundant, or feeling that the content went against the way users wanted to present themselves). This suggests that the reason for users to start censoring their content may not be the difficulties to target specific groups, but the perception users create on how their audience will respond to the posted content. Therefore, users may not seek a better way to target groups in their network to share content with when considering to post content, but a way to assure how their audience will respond to the content they want to post instead. Being able to target specific groups in an ever growing social network may rather be a strategy of users in order to better predict how their audience will respond to the content by shrinking the audience to a manageable size.

We created an online experiment to test this hypothesis in which we set up a persuasive cue that provide users with possible responses on the content against a group targeting mechanism. By comparing these two methods against each other we investigated the behavioral change these methods can induce in posting decisions and gain deeper insight in the underlying mechanisms of posting behavior. These insights provide knowledge on how to help users to make better posting decisions. This is important as SNSs are intended for people to create and share content about

themselves as a result of voluntary disclosure among multiple users [16]. Wrongly made decisions in the posting process affects the functionality of SNSs: not posted content that should have been posted jeopardizes SNSs functionality as it limits sociability, while posted content that should not be posted contributes to regret of the poster.

II RELATED WORK

In this section we discuss users' behavior when considering sharing content on SNSs. We continue the related work with how users are judging the appropriateness of the content followed by their strategy to reduce regret of posted content. Finally, we discuss recent work that tries to overcome sharing difficulties.

Behavior on SNSs Users have difficulties in defining their social connections. The "friend" category in SNSs is very broad and ambiguous. Most users tend to list anyone they have met during the course of their lives and do not actively dislike [6]. This kind of behavior results in that the user's social network include a mix of strong and weak ties [13]. SNSs have created extensive sets of privacy controls that allow users to interact at different levels of sociability. Despite these privacy controls, users are not utilizing the privacy settings provided [5]. Although users are not utilizing the privacy settings, they still actively share their content [11]. The content considered for sharing undergoes the evaluation of possible regret. Shared content becomes regret because of unforeseen or ignored consequences, such as they want to be perceived in favorable ways, they do not think about the consequences of their post, or misjudge the culture and norms within their network [24].

The imagined audience When considering content to share, users have a sense of audience. Strater and Lipford [23] found that the user's perceived audience shrinks over time; as users interact more with a certain group they start to perceive them as their primary audience and pay less attention to others. This behavior suggests that connections other than the user's perceived audience become less significant over time, thus also less important to take into consideration to share content with.

Users create an "imagined" audience to estimate the appropriateness of the content presentation [19]. They use cues from their social media environment to construct and augment knowledge about their audience [7]. However, as the interpretation of these cues are subjective, the created image may not be accurate and

can be deviated from the actual readers. When the environment involves higher interaction with its readers, the awareness of the imagined audience becomes more prominent and thereby the specifics becomes more important [19]. A deviated image of the audience can therefore become problematic. Especially in SNSs where sharing is part of being social, the possible misconception users create can significantly influence the way users share their content.

Self-censorship One of the strategies to minimize the risk of regret is to self-censor content [19, 22, 24]. Although self-censorship is an effective strategy to prevent regret, it also increases the chances that content that would have been safe is left unshared. In a qualitative study by Sleeper et al. [22] in the U.S., most commonly found content subject to self-censoring is external content (e.g., entertainment, politics) followed by personal opinions and updates. The most important reason connected to these contents is to control the self-presentation. To a slightly lesser extent, other reasons of censoring content were found: users did not want to start an argument or discussion, or were afraid to be boring or repetitive. These self-censoring reasons apply to almost half of the content considered to be shared [22].

Privacy settings To help users in their sharing decisions, research has focused on providing different methods to increase control and transparency of disclosure behavior. Extensively investigated are methods to improve group targeting settings, such as machine-learning solutions [3] as well as ways to give users more control about what to share and with whom [21]. A recent system providing an advanced sharing method are "circles" on Google+ [17]. Although such settings provide users with more control, they do not help prevent users posting content that they should not post. Additionally, research about privacy settings indicate that users often do not use the custom privacy settings available, but adhere to standard settings instead [22–24]. These standard settings provide an "all or nothing" situation. Users are able to set their information open to everybody or protecting their whole profile by utilizing the setting to restrict it to "friends only." Although users choose for a specific setting, they still deliberately choose to self-censor information and not share content [23].

Persuasive cues Some studies have been focusing on accommodating persuasive cues by providing a justified reason for the user to disclose information, such as giving a reason why it would be better to disclose [10], or appealing to the social norm by displaying what

others have done [1, 4, 20]. The presentation of these kind of persuasive cues have been studied in different ways. For example Patil, Page, and Kobsa [20] used a descriptive way to present the aggregated privacy choices of one's social circle, while Besmer, Watson, and Lipford [4] chose for a visual approach to display the social norm.

III THE ACTUAL SHARING PROBLEM

Based on the just discussed literature, we can conclude that improving group targeting mechanisms may not be the right solution as users do not make sufficiently use of them. Even if users choose a setting, they still rely on what they *think* is acceptable. Ajzen [2] defines this influence on the decision making as the subjective norm in which behavior is influenced by the perceptual judgments and beliefs of relevant others about the intended behavior. Results of previous research on SNSs indeed suggest that the underlying reasons for users to start censoring their content are the perceptions they create about responses of their audience on the content [19, 22–24]. What is shared is judged on its perceived safeness and appropriateness, as well as what is considered to be socially acceptable and normal within their social network [23]. As these judgments are subjectively created by the user, they are uncertain to some extent.

The work in this study is based on the assumption that posting decisions consist of uncertainty. We believe that by presenting a cue that provides possible responses of the user's audience on the content, we can minimize that uncertainty. Providing such cue will help users in their decision making process and prevent them from wrongly censoring or posting content. To the best of our knowledge, the usage of this kind of cues has not yet been explored in the context of content sharing. Our study was designed to answer the following main questions:

1. Will presenting possible responses of the audience change posting behavior in compliance?
2. Will providing a group targeting mechanism in a self-censoring state change posting behavior?
3. When in a self-censoring state, will presenting responses have a stronger effect on behavioral change than a group targeting mechanism?

Users create a subjective judgment about what is acceptable within the audience to which the content

is shown. Therefore, we expect that when users know how the audience will possibly respond on the content, they will adjust their posting behavior towards the polarity of the responses. Presenting responses of the user's audience on the content will reduce self-censoring behavior, but also warns users of posting content that they should not. In both cases, this cue reduces regret of a wrongly made decision.

We expect group targeting mechanisms to alleviate self-censoring problems to some extent. By using group targeting mechanisms, users are gaining more control over their content. That is, to whom they share the content with. By targeting specific groups in the network, users are able to shrink their ever growing social network to a size in which they can make more easily judgments about the appropriateness of the content. We expect that it only alleviates the self-censoring problem partly, hence the appropriateness estimation of the content is still subjectively created by the user. Thus, giving users more control will have a smaller effect compared to presenting possible responses of the audience.

Given that we expect both methods to have a positive effect on posting decisions, there may be an accumulation of the effect size when both methods are combined. We expect that combining the methods will provide a positive interaction effect.

IV METHODOLOGY

The experiment that would provide insights in posting behavior was conducted in two steps, which we will refer to as *Study A* and *Study B*. The goal of Study A was to identify 1) content for posting, and 2) reasons for censorship. We used these outcomes in the design of Study B, where we observed the actual posting and censorship behavior. In Study B, we put the user in a posting position for a selected content. Then the participant was subjected to a self-censorship inducing scenario and put in one of the experimental conditions. Afterwards, the participant was asked again if they wanted to post the content. The difference between the first posting position and the second were the observed behavior changes.

1 STUDY A

In order to acquire the desired data (content for posting and reasons for censorship) we followed the experiment design laid out by Sleeper et al. [22]. The workflow of the study is depicted in Fig. 1.

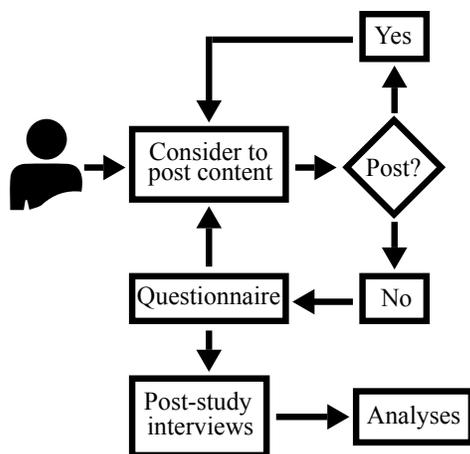


Figure 1: Study A: Workflow

We recruited 21 university students (8 males, aged between 20-26 years, median age 22 years) for a one week diary study and a concluding interview. Participants were asked to fill-in a questionnaire every time they self-censored a content posting on Facebook. The questionnaire consisted of three questions: 1) context of the content, 2) content type (e.g., photo, video etc.), and 3) reason for censoring. After a week of reporting posting censorship, participants were invited for a post-study semi-structured interview where they discussed in more details the censored content and the reasons for censorship.

1.1 DATA CODING AND ANALYSIS

The goal of data coding was to map the raw data from the diary study and interviews into a set of self-censored contents and reasons. The results of coding are reported in Fig. 2 and Fig. 3.

In total, participants reported 88 self-censored contents in the diary study. Following the coding scheme of Sleeper et al. [22], coding took place in four steps: 1) two researchers coded half of the content each, 2) based on these codes, the researchers created a set of higher level codes, 3) these were used to code the remaining halves (for each researcher), and 4) disagreements and inconsistencies were discussed jointly.

1.2 RESULTS

Participants self-censored various kind of content (see Fig. 2). Most censored contents in our sample were about personal opinions and updates. Personal opin-

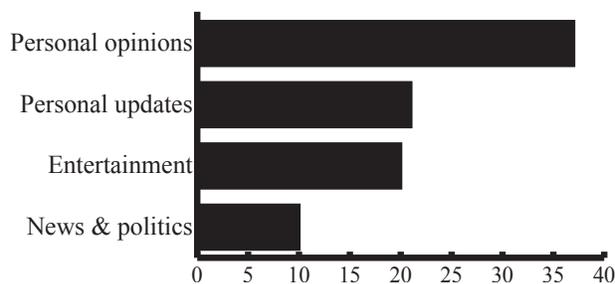


Figure 2: Number of self-censored content by type.

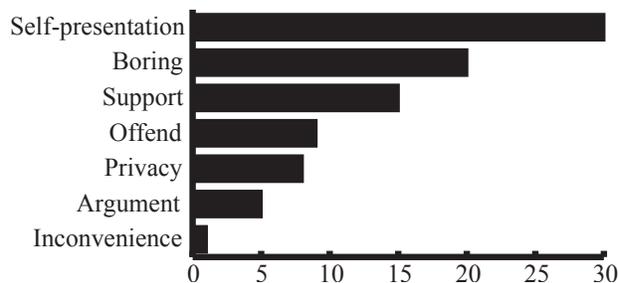


Figure 3: Number of self-censored reasons by type.

ions were mostly about how participants felt about specific things happening in their life whereas personal updates were in general about happenings of participants throughout the day. These were mainly expressed as status updates on Facebook. Next to personal content, entertainment content (i.e., music, humoristic, sports, and fashion items) are most commonly censored. Finally, to a much lesser degree participants self-censored political and news items.

Responses of participants about the reason of self-censoring could be placed in one of the following categories (see Fig. 3):

- **Self-presentation:** worried that the content would hurt their own presentation.
- **Boring/repetitive:** concerned that the content would be perceived as boring or repetitive.
- **Support/sympathy:** worried that friends would not respond on the content (comment or like).
- **Argument:** did not want to get involved into an argument about the topic with anybody.
- **Offend:** felt that the content may hurt somebody.
- **Privacy:** felt that the content would violate privacy of others.

- **Inconvenience:** too much effort to post the content.

The main reason for participants to censor their content is because of self-presentation concerns. This is for a big part related to the personal opinions and updates categories. Participants were mainly worried that posting something about themselves would possibly create a negative image toward others. Other significant reasons is the believe that the content would be perceived as boring, or that they would not get the appropriate support and sympathy from others. To a lesser degree participants were worried about privacy, creating an argument, or offending someone.

In the interviews we additionally asked participants if they would post the self-censored content under different circumstances, such as being able to target specific groups in their network or if they would know how their social network would possibly respond. Striking was that all the participants said that for personal opinions and updates that there was nothing to change their decision. For all other content, participants seemed to be more open minded. Furthermore, we asked participants about disclosure concerns with certain groups in their social network. Depending on the content type, most participants expressed concerns about sharing content with some of the groups in their network except for "close friends." Participants did not show concerns about disclosing any content with this group.

1.3 DISCUSSION

Our goal in this study was to better understand the contents and reasons for self-censoring behavior. Additionally, as our study took place with South Korean participants, previous findings of Sleeper et al. [22] gave us the opportunity to explore possible cultural differences. Most of our findings show agreements with the results of previous research. Although we found less varied content, the content types we found are in line with the findings of Sleeper et al. [22] among U.S. participants. Furthermore, a resemblance can be seen in the order of the major categories. The less variation in our study may be explained by the age range and occupation variation of our sample. We focused on university students between 20 and 30 only, while the sample of Sleeper et al. [22] varied in their occupation and age (20-51). Still, similar self-censoring reasons were found.

A compelling additional reason we found among South

Koreans is the concern for support and sympathy. Our participants expressed a substantial concern of getting no responses on their postings. This longing for support and sympathy may be connected to cultural dimensions. As most Asian cultures, Koreans tend to have a collectivistic nature that may explain the importance of this reason [15]. Despite this difference, our findings show in general similar trends between Americans and Koreans indicating that cultural differences may not be so prominently present as thought.

2 STUDY B

Study B represents the core experiment with which we wanted to get insights into how a group targeting mechanism and a response prediction on the content can influence the posting behavior. We developed an online experiment where the users 1) first chose a content, then 2) decided whether to post it or not, then 3) were subjects to a manipulation (we had manipulations dealing with user groups and with predicted response; see §2.1 for details), and finally 4) decided again whether to post the content or not. This experiment design, outlined in Fig. 4, allowed us to gather the data needed to make the conclusions to our initial research questions.

We recruited 215 participants (104 males, age ranging from 20 to 30 years, median age 24 years) among university students in South Korea. Participants were asked to post content through a Facebook-like web application. The application was a visual clone of Facebook running on our server.

In a pre-step we chose five groups that were to be used in some of the manipulations (see §2.1). Participants were asked to define two groups of friends out of their social network. In addition to these two groups we added three more: *friends*, *public* (default Facebook groups), and the group *close friends* (participants in Study A suggested that they would have had a different posting behavior if there was the possibility to target such a group).

In the next step we put the user in a self-censoring situation. Self-censoring is one of the possible results of a decision made about content that users are considering to post. In other words; when a user has a positive attitude to post the content but feels hesitant to continue the posting. Based on the outcomes of Study A (see §1.2), we selected the set of contents that were offered for posting to the participants. This set was composed of 9 different video clips of roughly 1 minute of length: movie (2x), music (2x), humor

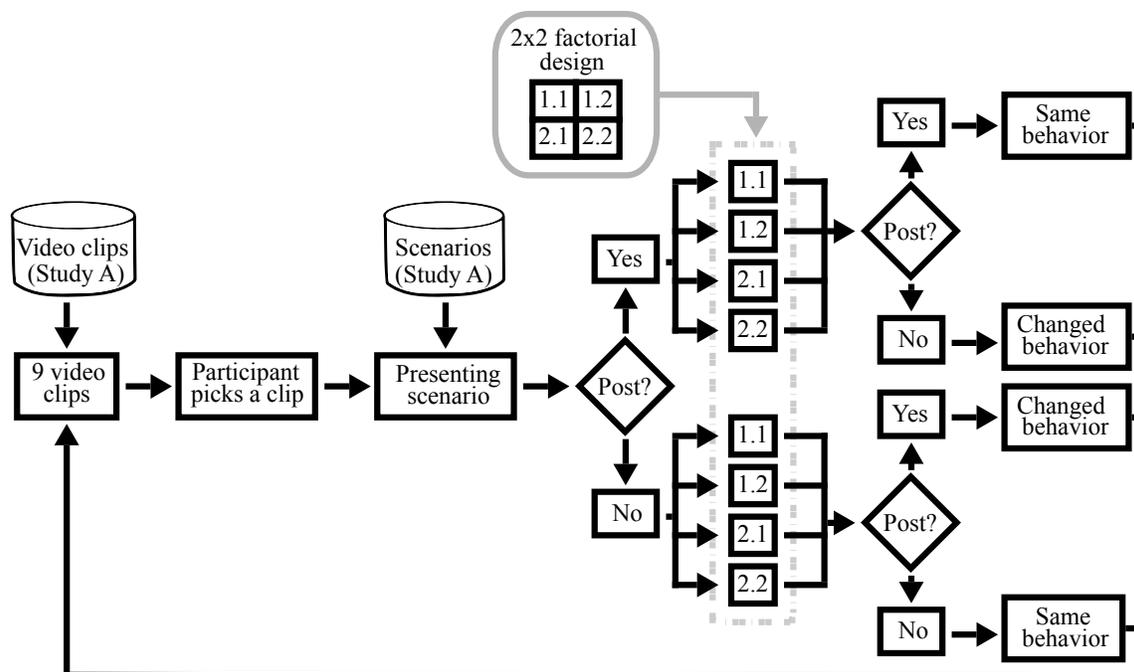


Figure 4: Study B: Workflow

(2x), TV-series (1x), politics (1x), and news (1x). Participants chose one of the content items that they would consider to share. This ensured the necessary positive attitude towards posting the content. To ensure the possibility to hesitate on posting the content, participants were then presented with a reason for not posting (chosen from the set of reasons collected in Study A; see §1.2) and were asked whether they wanted to post the content. If the participant decided not to post, self-censorship has occurred as the participant had a positive attitude to post by choosing their own content.

In the next step the participants were exposed to one of the four conditions that could have an effect on their posting behavior. After being exposed to the conditions the participants were asked again if they wanted to post the content. For each participant, the experiment flow was designed in such a way that they were exposed to all four conditions in a random order.

2.1 MANIPULATIONS

We wanted to observe the influence of two factors on the posting behavior: a) group targeting mechanism, and b) predicted audience's response. Hence we used a 2x2 within-subject factorial design consisting of the following manipulation conditions (see Tab. 1): (1.1)

with group targeting mechanism and with predicted audience's response, (1.2) with group targeting mechanism only, (2.1) with predicted audience's response only, and (2.2) without manipulation (control condition). Using a within-subject design the participants went four times through the procedure. For each round the content choices, as well as the reasons for self-censoring, differed. To cancel out order effects, participants were randomly assigned to one of the conditions until all four were met.

Condition	Group targeting mechanism	Predicted audience response
(1.1)	Yes	Yes
(1.2)	Yes	No
(2.1)	No	Yes
(2.2)	No	No

Table 1: Manipulations

Audience's response prediction This condition was designed to present possible responses of the user's social network members. In order to effectively capture the effect of the manipulation, we designed it to oppose the participant's choice of posting. Since participants could choose whether or not to follow the self-censoring scenario, we created two different textual messages based on work of Patil et al. [20]: 1)

positive/comforting message, and 2) negative/warning message. When participants decided to censor the content, a comforting message was presented trying to change their decision (e.g., "We analyzed your social network and based on their [participant's audience] responses on similar content they would like this posting."). A warning message was presented when participants did not censor their content. That is, when they decided to post the content regardless the self-censoring scenario provided (e.g., "We analyzed your social network and based on their [participant's audience] responses on similar content they will not like this posting."). Next to the text messages, we also included a visual method (based on work of Besmer et al. [4]) where we used a bar indicating the percentage of the user's social network that would like the posting. For the comforting message we set the bar on 95%, and for the warning message we set the bar on 5%. This condition alone was present in the manipulation ((2.1); see Tab. 1).

Group targeting mechanism In this condition we displayed the five groups of users (two subject generated groups, *public*, *friends*, and *close friends*) and required the subject to choose one of the groups when posting the content. We did not define a default group in order to avoid situations in which participants would just click through or get influenced by the default option. This condition alone was present in the manipulation ((1.2); see Tab. 1).

Combined audience's response prediction and group targeting mechanism We combined the audience's response prediction messages and the group targeting mechanism condition in order to investigate the existence of an accumulative effect. Both conditions were presented in the manipulation ((1.1); see Tab. 1).

Control To ensure that the observed effects are not due to the fact that participants could think their previous decision over, we included a control condition. This condition gave participants the opportunity to change their previous posting decision without the addition of a guiding message or added functionality. To this end, we provided just a message asking if they wanted to change their previously made decision.

2.2 MEASUREMENTS

In order to investigate the behavioral change the conditions induce in posting behavior, we captured three variables: 1) the posting behavior of participants before the condition presented (post or not post), 2) the

experimental condition presented to the participants (i.e., (1.1), (1.2), (2.1) or (2.2)), and 3) the posting decision after the condition presented (post or not post). The influence of the condition was measured by whether participants changed their posting decision after the condition was presented. Additionally, to gain a better understanding about the final posting decision made by the participants, we asked them to write down the reasons for their behavior.

2.3 RESULTS

As we gave participants the choice to follow the self-censoring scenario we created, or to ignore it, we obtained different starting posting behaviors. We divided our analyses in order to investigate the behavioral change of the manipulations on not posting (self-censoring) and posting behavior. In this section we first discuss the results of participants that followed the self-censoring scenario. That is, whether the manipulations could induce a behavioral change from self-censoring (not posting) to posting. We continue with the results of our analysis where we discuss the results of participants that ignored the self-censoring scenario. In other words, whether the manipulations could induce a behavioral change from posting to not posting (self-censoring).

In general, our interest was the effect of the manipulations on posting behavior and the effect size of each. We performed a repeated measures logistic regression to test this, by using a generalized linear model (GENLIN) with a binomial distribution and a logit link function.

Not posting (self-censoring) In this section we present our results with the self-censoring cases ($n = 90$). The GENLIN's goodness-of-fit is reflected in the quasi likelihood criterion ($QIC = 188.434$) and the corrected quasi likelihood under independence model criterion ($QICC = 118.631$). A low difference between QIC and QICC indicates that the model has a good correlation structure and the predictors obtained fit the model well.

To investigate the effects of the logistic regression, odds ratios (OR) are reported. We first assessed the baseline odds (control condition) in order to see the effects of the manipulation conditions. When participants did not receive the audience's responses on the content nor able to target specific user groups in their social network, we found the odds that they changed their initial posting decision (censoring) to be 0.207 ($CI = 0.076$ to 0.564 , Wald Chi-Square = 9.479 ,

$p = 0.002$). In other words, when participants were just given the opportunity to readjust their decision, 0.207 participants decided to change their posting decision for every participant that did not. The lack of a higher odds ratio in a certain direction supports our assumption that posting decisions in a self-censoring state is uncertain as participants' posting decision fluctuates.

Exploring the effect of presenting participants with possible responses of their audience, we found an increase of the baseline odds. When participants knew how their audience would respond on the to be posted content, the baseline odds increased by a factor 9.302 ($CI = 2.548$ to 33.955 , Wald Chi-Square = 11.397 , $p = 0.001$). This means that the odds of participants changing their initial posting decision in this condition is 1.93. We also found an increase of the baseline odds when participants were only given a group targeting mechanism. The baseline odds increased by a factor 5.375 ($CI = 1.748$ to 16.529 , Wald Chi-Square = 8.609 , $p = 0.003$). In other words, the odds of participants posting their content when they could target specific groups is 1.11. Given these results, we can conclude that both conditions induce a posting decision change. However, the odds are higher when participants knew how their audience would respond.

The interaction effect of combining the two manipulation conditions show a minimal change in the odds ratio ($OR = 0.015$, $CI = 0.042$ to 0.563 , Wald Chi-Square = 7.968 , $p = 0.005$). This suggest that there is an effect between knowing how the audience would respond and being able to target user groups. However, the odds increase is negligible small.

The choice that participants could make between different content types could have played a role in the effect of the conditions. To investigate this, we dummy coded the content type in order to add them as a covariate. The only (marginally) significant results were found for political ($OR = -0.131$, $CI = 0.016$ to 1.061 , Wald Chi-Square = 3.626 , $p = 0.057$) and news items ($OR = -0.241$, $CI = 0.056$ to 1.033 , Wald Chi-Square = 3.671 , $p = 0.055$). Surprising is the negative odds ratio of both items, indicating a decrease in the odds of changing posting behavior. Analyzing the quantitative data obtained, revealed that participants took this content more serious. They found this content *heavy loaded*; the topics discussed were serious and having an *wrong* opinion about this could significantly influence the way others would look at them.

Posting When asking participants ($n = 125$) why they did not follow the self-censoring scenario, they responded that they found the reasons given not severe enough for them to start censoring the content they really liked. As participants indicated to have a high positive attitude toward this content, it gave us the change to explore to which extent uncertainty consist in posting behavior. Prior to the analysis, we developed additional expectations. As we believe that posting decisions consist of uncertainty, we expected that the conditions involving the audience's response would be able to change posting behavior. Furthermore, as the group targeting mechanism condition did not have any guiding message, we expected this to not have any effect on posting decisions. That is, we expected this condition to have the same effect as the control condition.

The two quasi-likelihood criteria to assess the goodness-of-fit indicate a fairly good model fit ($QIC = 125.996$ and $QICC = 126.317$). Results show a baseline odds of 5.716 ($CI = 2.035$ to 16.054 , Wald Chi-Square = 10.9464 , $p = 0.001$). Meaning that 5.716 participants stayed with their initial decision of posting the content for every participant that did not. This suggest that participants were more confident about their posting as the odds of staying with their prior decision is much higher than changing it. The main effect of the audience's response condition show a decrease of the baseline odds ($OR = -0.325$, $CI = 0.092$ to 1.140 , Wald Chi-Square = 3.081 , $p = 0.079$). Meaning that 1.85 participants changed their posting to not post for every participant that kept on posting. As expected, a non significant main effect was found for the group targeting condition ($OR = 5.030$, $CI = 0.604$ to 41.893 , Wald Chi-Square = 2.231 , *ns*). Additionally, a non significant interaction effect was found ($OR = .171$, $CI = .017$ to 1.684 , Wald Chi-Square = 2.291 , *ns*). Meaning that there is no difference between the audience's response condition with or without the option to target groups.

As with the previous section, the political content revealed a marginally significant effect ($OR = -0.172$, $CI = 0.024$ to 1.249 , Wald Chi-Square = 3.626 , $p = 0.057$). The negative relation indicates an increase of the odds ratio of changing posting behavior. That is, participants were more likely to change their decision from posting to not posting for this kind of content.

2.4 DISCUSSION

In Study B we investigated three different methods to influence users' posting decisions on SNSs by presenting: possible responses of the user's audience, group targeting mechanism, and a combination of the two methods. Our results show that all three methods have an effect on users' posting decisions. By comparing these different methods, we were able to get a deeper understanding about the needs of users when posting content. Looking at the effect of each method, knowing whether their audience would like the posting seem to be more beneficial for users than able to target specific groups in their social network. It gives support to our notion that users are uncertain in their posting decisions as they were easily influenced. This uncertainty is highlighted in the results of the control condition in the self-censoring condition. Results showed that participants frequently reconsidered their self-censoring behavior when they were given the opportunity. When participants already decided to post the content, they were less prone to reconsider and were more likely to stick with their posting. They seem to be confident and sure about the content. However, this confidence is still affected when presented with how their audience would respond to the content. When participants knew that the content would not be perceived well, they started to come back at their posting decision.

Despite the fact that all three methods provided significant results on behavioral change, we found no support for our expectation of an accumulative effect for the combined condition. This may be explained by the fact that we guided participants to share content with the close friends group. The close friends group was chosen because the results of Study A indicated that users did not perceive disclosure problems with this group. Strater and Lipfort [23] noted that users are starting to see the group with which they interact most with as their primary audience. As users do not feel any disclosure problems with the close friends group, they may interact most with this group already and therefore seeing this group as their primary audience. Presenting general responses of the user's audience (audience's response condition) or more specific as in the combined condition, would not make a difference for the user as they depict the same primary audience.

We also found differences among content type. These differences can be explained by the "load." We used different content genres that can be divided in two groups: 1) heavy loaded content, and 2) light loaded

content. The heavy loaded content are more about serious items and comprised political and news content types. Movie, music, humoristic, and TV-series fall in the light loaded content category as they consist of frivolous and entertaining items. Participants tend to respond differently to possible responses of their audience depending on this content load. Our results indicate that a comforting message has less effect on heavy loaded content than a warning message. This suggests that for this kind of content, users are more cautious and prefer to err on the safe side when having the slightest doubt. The quantitative data we obtained about the reasons for participant's final posting decision indicate that the content load indeed plays a prominent role. Especially for the political content, participants were very careful by expressing their opinion as it indicates their political side. For light loaded content, these concerns were less present. Although this kind of content can also express a certain preference, participants felt that they could "laugh it away" when it would play against them. For this type of content participants seemed to be more sensitive for a comforting message that gives them that little push to post the content. However, when they are feeling confident about the content, the effect of the warning message decreases. This behavior is seen when participants did not follow the self-censoring scenario we presented to them. Additionally, it can also explain why a part of the participants did not follow the scenario. As they felt they could easily repair damage made to their self-presentation, they found the reasons given not severe enough to start censoring the content. This changeable behavior that we could induce by presenting possible responses accentuate the uncertainty that exists when users are trying to post their content on SNSs.

V CONCLUSION

Results suggest that posting behavior consist of uncertainty. Especially when users are self-censoring their content, they tend to change their decision quite frequently when given the chance. Although users seem to be more sure and confident about posted content, results indicate that some degree of uncertainty still exist. Group targeting mechanisms seem to contribute to influence posting decisions to some extent. Improving group targeting mechanisms can only partly take away uncertainty by allowing users to shrink their social network to a size that they can easily estimate the content appropriateness. However, users can still make wrong estimations. Results show that by providing predictions about how the audience

would respond to the content significantly help users to make better posting decisions; when to self-censor the content, or when it is safe to post. With regards to our research questions (see §III) this means that: 1) presenting users with possible responses of their audience do result in an adjusted posting towards the polarity of the responses, 2) group targeting mechanisms influence posting behavior in a self-censoring state, and 3) audience's response predictions do have a stronger effect on posting behavior than group targeting mechanisms.

Additionally, the posting guidance provided by the audience's response predictions depends on the content type. Users seem to be more cautious and suspicious of posting content that is serious and heavy loaded. Posting this content can damage their self-presentation which may be difficult to repair. Therefore, users are less sensitive for positive guidances for posting, and more sensitive for negative guidances (when there is any doubt of hurting their self-presentation, they will err on the safe side of their posting behavior; not posting the content). For light loaded content, users can use a little push to persuade them to continue posting. Even if this kind of content will not be received well by their audience, the consequences for their self-presentation are less severe and will be more easy to repair as the content is light and not serious. Therefore, participants seem to care less about what kind of responses the content will produce.

The effects found of giving guidance in user's posting decisions give raise to the question whether research should continue focusing on improving group targeting mechanisms. As SNSs are made for sociability, greater advantage should be taken of the information that can be derived from a user's social network instead of trying to limit sociability by providing restrictions.

VI LIMITATIONS AND FUTURE WORK

Study A relied on qualitative, self-reported data, and a relatively small sample was used. Although results seem to be consistent with findings of Sleeper et al. [22], we still lack the ability to generalize our findings. Furthermore, both studies (A and B) used a specific subset of actual SNS users; university students between 20s and 30s. Brandtzaeg et al. [9] argue that variation can occur among different age ranges and occupations. For example, older and younger users have other privacy perceptions and therefore could respond differently to the manipulations we provided. To alleviate these limitations, a bigger and wider sam-

ple need to be obtained in future studies in order to properly map the social diversity on SNSs.

In Study B, we made use of a combined condition where we merged the audience's response condition with the group targeting condition. To measure the influence of the message effectively, we tried to guide sharing behavior toward one specific group; close friends. An interesting direction would be to investigate the effects when given posting guidance for each user group separately to see if users have separate posting thresholds for different user groups. For example, users may desire a higher percentage of approval of the content when considering sharing content with close friends than with everybody.

Our findings could have been influenced by cultural dimensions. As we already stressed out in the discussion of Study A, the additional reason found for self-censoring may be explained by the collectivistic nature of South Koreans. The results of Study B need to be interpreted with some precaution as they could have been influenced by cultural dimensions too. One explanation of our findings is that self-esteem in collectivistic cultures is not derived through idiosyncrasy [14], but rather through harmony with the group [25]. This makes people in these cultures tend to fit in rather than to stand out [18]. Therefore, our results could have been compromised by participants trying to fit in by following the cue we provided (audience's response and combined conditions). Future studies should try to answer this question by using a more individualistic society.

Lastly, as we focused specifically on content sharing, our findings might be well applicable to other areas involving information disclosure.

VII ACKNOWLEDGMENTS

This research is supported by the Austrian Science Fund (FWF): P25655; and by the EU FP7/2007-2013 through the PHENICX project under grant agreement no.601166. Furthermore, we want to thank Sunwha Lee and Yeunju Kim for their assistance.

References

- [1] A. Acquisti, L. K. John, and G. Loewenstein, "The impact of relative standards on the propensity to disclose," *Journal of Marketing Research*, vol. 49, no. 2, pp. 160–174, 2012.
- [2] I. Ajzen, "The theory of planned behavior," *Organizational behavior and human decision processes*, vol. 50, no. 2, pp. 179–211, 1991.
- [3] S. Amershi, J. Fogarty, and D. Weld, "Re-group: Interactive machine learning for on-demand group creation in social networks," in *Proc. CHI*. ACM, 2012, pp. 21–30.
- [4] A. Besmer, J. Watson, and H. R. Lipford, "The impact of social navigation on privacy policy configuration," in *Proc. SOUPS*. ACM, 2010, p. 7.
- [5] J. Bonneau and S. Preibusch, "The privacy jungle: On the market for data protection in social networks," in *Economics of information security and privacy*. Springer, 2010, pp. 121–167.
- [6] D. Boyd, "Friends, friendsters, and myspace top 8: Writing community into being on social network sites," *First Monday*, vol. 11, no. 12, 2006.
- [7] —, *Why youth (heart) social network sites: The role of networked publics in teenage social life*. MIT Press, Cambridge, MA, 2007.
- [8] P. B. Brandtzæg and J. Heim, "Why people use social networking sites," in *Online communities and social computing*. Springer, 2009, pp. 143–152.
- [9] P. B. Brandtzæg, M. Lüders, and J. H. Skjetne, "Too many facebook "friends"? content sharing and sociability versus the need for privacy in social network sites," *Intl. Journal of Human-Computer Interaction*, vol. 26, no. 11-12, pp. 1006–1030, 2010.
- [10] S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge, "Location disclosure to social relations: why, when, & what people want to share," in *Proc. CHI*. ACM, 2005, pp. 81–90.
- [11] B. Debatin, J. P. Lovejoy, A.-K. Horn, and B. N. Hughes, "Facebook and online privacy: Attitudes, behaviors, and unintended consequences," *Journal of Computer-Mediated Communication*, vol. 15, no. 1, pp. 83–108, 2009.
- [12] N. B. Ellison, C. Lampe, and C. Steinfield, "Feature social network sites and society: current trends and future possibilities," *interactions*, vol. 16, no. 1, pp. 6–9, 2009.
- [13] N. B. Ellison, C. Steinfield, and C. Lampe, "The benefits of facebook "friends": social capital and college students' use of online social network sites," *Journal of Computer-Mediated Communication*, vol. 12, no. 4, pp. 1143–1168, 2007.
- [14] J. A. Goncalo and B. M. Staw, "Individualism–collectivism and group creativity," *Organizational Behavior and Human Decision Processes*, vol. 100, no. 1, pp. 96–109, 2006.
- [15] W. B. Gudykunst and Y. Matsumoto, "Cross-cultural variability of communication in personal relationships," *Communication in personal relationships across cultures*, pp. 19–56, 1996.
- [16] A. N. Joinson, "Looking at, looking up or keeping up with people?: motives and use of facebook," in *Proc. CHI*. ACM, 2008, pp. 1027–1036.
- [17] S. Kairam, M. Brzozowski, D. Huffaker, and E. Chi, "Talking in circles: selective sharing in google+," in *Proc. CHI*. ACM, 2012, pp. 1065–1074.
- [18] S. Kitayama, H. R. Markus, and C. Lieberman, *The collective construction of self esteem: Implications for culture, self, and emotion*. Kluwer Academic/Plenum Publishers, 1995.
- [19] A. E. Marwick, "I tweet honestly, i tweet passionately: Twitter users, context collapse, and the imagined audience," *New Media & Society*, vol. 13, no. 1, pp. 114–133, 2011.
- [20] S. Patil, X. Page, and A. Kobsa, "With a little help from my friends: can social navigation inform interpersonal privacy preferences?" in *Proc. CSCW*. ACM, 2011, pp. 391–394.
- [21] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao, "Understanding and capturing people's privacy policies in a mobile social networking application," *Personal and Ubiquitous Computing*, vol. 13, no. 6, pp. 401–412, 2009.
- [22] M. Sleeper, R. Balebako, S. Das, A. L. McConahy, J. Wiese, and L. F. Cranor, "The post that wasn't: exploring self-censorship on facebook," in *Proc. CSCW*. ACM, 2013, pp. 793–802.

- [23] K. Strater and H. R. Lipford, “Strategies and struggles with privacy in an online social networking community,” in *Proc. BCS HCI*. British Computer Society, 2008, pp. 111–119.
- [24] Y. Wang, G. Norcie, S. Komanduri, A. Acquisti, P. G. Leon, and L. F. Cranor, “I regretted the minute i pressed share: A qualitative study of regrets on facebook,” in *Proc. SOUPS*. ACM, 2011, p. 10.
- [25] P. Wink, “Beyond ethnic differences: Contextualizing the influence of ethnicity on individualism and collectivism,” *Journal of Social Issues*, vol. 53, no. 2, pp. 329–349, 1997.