

***Modeling Self-Disclosure in Social Networking Sites:
Understanding the Causes and Consequences of Self-Disclosure
through Automatic Language Analysis***

Yi-Chia Wang

CMU-LTI-15-012

Language Technologies Institute
School of Computer Science
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213
www.lti.cs.cmu.edu

Thesis Committee

Robert E. Kraut (Chair), Carnegie Mellon University
William W. Cohen, Carnegie Mellon University
Eduard Hovy, Carnegie Mellon University
Moira Burke, Facebook Inc.

*Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
In Language and Information Technologies*

© 2015, Yi-Chia Wang

Keywords: natural language analysis; natural language processing; machine learning; computer-supported cooperative work; social computing; analysis of social media; social network analysis; social network sites; social networking sites; online social networks; Facebook; computer-mediated communication; self-disclosure; self-presentation; impression management; personality; audience; context collapse; life events; relationship formation; topics; emotional valence; social distance.

ABSTRACT

Social networking sites (SNSs) offer users a unified platform to build and maintain social connections. Quality user experience relies on understanding when people feel comfortable sharing information about themselves on SNSs because self-disclosure helps to maintain friendships, increases relationship closeness, and benefits the discloser's health and well-being. This thesis develops a new machine learning model to measure self-disclosure in SNS communication at scale to better understand the contexts in which users of Facebook, the world's largest social networking site, disclose a higher or lower level of personal information about themselves. The machine learning model was built using four key features, including emotional valence, social distance between the poster and people mentioned in the post, similarity of language in the post to what others are discussing, and post topics. The model performs moderately and in line with the judgments of trained coders ($r=.60$). I then apply the model automatically to de-identified, aggregated samples of Facebook users' status updates and examine factors at three levels that might influence their self-disclosure: their stable, personal characteristics, the structure of their Facebook networks, and events in their lives.

Results from this study confirm and extend earlier psychological research on the conditions associated with self-disclosure. Specifically, this study shows that women self-disclose more than men, and users who score higher on an Impression Management scale, indicating a stronger desire to manage the impressions others have of them, self-disclose less. At the level of audience structures, results indicate that social network size negatively correlates with self-disclosure, while network density and average tie strength with friends positively correlate. However, the analysis results of product feature tests designed to make users more aware of the audience's existence were ambiguous.

Longitudinal analyses examining self-disclosure among Facebook users who experienced major life events indicate that positive events increase self-disclosure, whereas negative events constrain disclosure. In particular, users self-disclosed more during periods when they were experiencing the start of a new romantic relationship and self-disclosed less when experiencing a break-up. In addition, students disclosed more about themselves at

the start of their academic term; this peak was larger for college freshman than for college sophomores. Further, increased self-disclosure correlates with a smaller increase in users' friend count, which indicates potential tension between audience size and disclosure.

This thesis has both theoretical and practical implications. Theoretically, it advances our understanding of the conditions associated with variation in online self-disclosure. Practically, it provides methods for measuring self-disclosure at scale in social networking sites and guidance for SNS designers to improve their services by providing better affordances to users.

ACKNOWLEDGMENTS

Carnegie Mellon has been my dream school since I studied computer science at college in Taiwan. I have been lucky to be able to pursue a doctoral degree in this great education and research environment, surrounded by many brilliant people. I would not have been able to finish my degree and dissertation without the encouragement and support of my advisor, colleagues, friends, and family.

The person I want to acknowledge most is my advisor, Bob Kraut. Before I met him, my understanding of research was limited to building models that perform better than previous ones. He taught me social sciences and psychological research methods and introduced me to a new level of research. When I felt frustrated or lost about my research direction, he could always reignite my enthusiasm and guide me out of the maze with just a meeting. He has provided me invaluable intellectual guidance and financial support over the years. In addition to research, I also feel safe to discuss anything in my life with him. He is both my academia parent and life mentor. I am deeply indebted to him.

I am also extremely grateful to my committee members. Moira Burke has shown me how to conduct research that not only makes theoretically important contributions but that also has practical applications for designers. Without her help and support, I would not have had the opportunity to collaborate with Facebook on this dissertation. She is also my role model for time management and work-life balance. William Cohen and Eduard Hovy have provided me constructive feedback from different angles for my research. Their sharp insights have made this thesis stronger. My committee members have helped me to grow as an independent researcher.

I am also thankful to the organizations that have funded or provided data for my research throughout these years, including National Science Foundation, National Institute of Mental Health, and Facebook.

I want to thank my co-authors and colleagues, especially the Social Computing Lab, fellow LTI and HCII faculty and students, and the Facebook Data Science team. My conversations with them have inspired me and shaped my research ideas. Many thanks in particular to Jaime Arguello, Jason S. Chang, Sourish Chaudhuri, Mahesh Joshi, Rohit

Kumar, Tyne Liang, John Levine, Carolyn Penstein Rosé, Jian-Cheng Wu, Haiyi Zhu, Paul André, Bo Reum Choi, Laura Dabbish, Rosta Farzan, Gahgene Gweon, Sara Kiesler, Jiwei Li, Niki Kittur, Jennifer Marlow, Elijah Mayfield, Jeffrey Rzeszotarski, Yla Tausczik, Di Wang, Miaomiao Wen, Eytan Bakshy, Dean Eckles, and Brian Karrer.

My journey of studying abroad would not have been possible without the company of friends. A special thanks to Shih-Ping Kerry Chang, Yu-Chia Richard Chang, Chung-Yu Chen, Hong-Chien Doris Chen, Ming-Yu Robert Chen, Yun-Nung Vivian Chen, Justin Chiu, Ruogu Kang, Yueh-Ying Monica Han, Ting-Hao Kenneth Huang, Yi-Fen Huang, Szu-Chen Stan Jou, Frank Lin, Jung-Yuan Paisley Lin, Leon Tseng, Hao-Chuan Wang, Richard Wang, Mu-Li Wu, Genie Yang, Wayne Yang, Chien-Heng Yu, and Shoou-i Yu.

Finally, I would like to thank my beloved parents and husband, Chun-Chih Wang, Bao-Yu Chen, and Kai-Min Kevin Chang. Unlike traditional Asian parents who emphasize grades and push their children to study, my parents spent time with me and encouraged me to explore and develop my own interests since I was a child. They are the best parents ever. I have also been blessed with a wonderful husband. Kai-Min has supported me with great patience and helped me to break through barriers in life. He is always there for me when I need someone to lean on.

TABLE OF CONTENTS

LIST OF TABLES.....	10
LIST OF FIGURES	12
1 Introduction.....	13
1.1 Background Overview	15
1.2 Thesis Overview	17
1.2.1 Machine learning model of self-disclosure	17
1.2.2 The relationship of poster and audience factors to self-disclosure	17
1.2.3 Patterns of self-disclosure and network growth around life events	17
1.2.4 The effect of context collapse on self-disclosure	18
1.3 Approach and Impact	18
2 Measuring Self-Disclosure in Social Media Content.....	20
2.1 Introduction.....	20
2.1.1 Machine learning approach for self-disclosure measurement.....	21
2.2 Data Collection and Agreement Analysis of Coded Data.....	22
2.2.1 Self-disclosure instrument.....	22
2.2.2 Collecting Facebook status updates and ratings from posters.....	23
2.2.3 Agreement analysis between posters and researchers.....	24
2.3 Machine Learning Model of Self-Disclosure.....	25
2.3.1 Text processing and feature extraction	26
2.3.2 Model construction and evaluation	33
2.4 Discussion	36
2.4.1 Limitations and future directions	36
3 Personal and Audience Factors Related to Self-Disclosure.....	37
3.1 Poster Characteristics Influencing Self-Disclosure.....	37
3.1.1 Personality: impression management.....	37

3.1.2	Gender.....	38
3.2	Audience Factors Influencing Self-Disclosure	38
3.2.1	Public versus private communication	38
3.2.2	Closeness to communication targets	39
3.2.3	Context collapse on social networking sites	39
3.3	Predicting Self-Disclosure in Status Updates from Poster and Audience Characteristics	39
3.3.1	Poster characteristics and self-disclosure.....	40
3.3.2	Audience factors and self-disclosure	40
3.3.3	Who self-discloses more?	42
3.4	Discussion	43
3.4.1	Limitations and future directions	45
4	A Longitudinal Study of Online Self-Disclosure around Life Events Related to Social Network Disruption	46
4.1	Introduction.....	46
4.1.1	Life events and self-disclosure.....	48
4.1.2	The social consequences of self-disclosure.....	49
4.2	Study 1: Self-Disclosure of Users Who Changed Relationship Status	51
4.2.1	Dataset.....	51
4.2.2	Visualizing self-disclosure around relationship changes	51
4.2.3	Statistical analysis and results.....	52
4.3	Study 2: Self-Disclosure of New College Students	54
4.3.1	Dataset.....	54
4.3.2	Visualizing self-disclosure among students.....	55
4.3.3	Self-disclosure around the start of school	56
4.3.4	Network growth over time	58
4.3.5	Self-disclosure and network growth.....	60
4.4	Discussion	63
4.4.1	Limitations and Future Directions	65
5	Causal Relationships between Context Collapse and Self-Disclosure.....	67
5.1	Introduction.....	67
5.1.1	Context collapse, self-presentation, and self-disclosure	68

5.2	Audience Counter Study	70
5.2.1	Experimental design.....	70
5.2.2	Analysis and results	71
5.3	Privacy Checkup Study	74
5.3.1	Experimental design.....	74
5.3.2	Analysis and results	74
5.3.3	Propensity score matching	77
5.3.3.1	Motivation.....	77
5.3.3.2	Analysis and results of PSM	78
5.4	Discussion	81
6	Conclusion and Future Directions.....	83
6.1	Summary of Findings and Contributions	83
6.2	Discussion of General Limitations and Future Research.....	87
6.3	Closing Remarks	91
	References.....	92
	Appendix A: Informed Consent Used in the Amazon Mechanical Turk Task.....	101
	Appendix B: Screenshot of the MTurk Task	104
	Appendix C: Familiar Nickname Dictionary	105

LIST OF TABLES

Table 1. Self-disclosure measurement items for Facebook status updates as rated by Turkers.....	23
Table 2. Examples of status updates and composite ratings by Turkers and RAs.....	24
Table 3. Lists of positive and negative emoticons, adopted and modified based on the list of western emoticons on Wikipedia (Wikipedia, 2015).....	27
Table 4. Examples of status updates with counts of each type of people references, social distance features, and composite self-disclosure ratings by RAs.....	31
Table 5. Evaluation results with alternative feature sets.....	35
Table 6. Top 10 ranked topic features and their corresponding sample vocabulary in the model trained with 80 topic features.....	35
Table 7. Descriptive statistics for the variables in the regression analyses of the effects of audience factors on self-disclosure.....	41
Table 8. Results of the regression analyses predicting self-disclosure from audience factors.....	43
Table 9. Random-effects linear regression model at the user level predicting machine-coded self-disclosure of users who changed relationship status.....	53
Table 10. Descriptive statistics of the college student dataset.....	54
Table 11. Random-effects linear regression model at the user level predicting machine-coded self-disclosure.....	57
Table 12. Random-effects linear regression model at the user level predicting logarithms of friend counts with respect to base 10.....	59
Table 13. Descriptive statistics for the numerical variables in the analysis of the relationship between self-disclosure and network growth.....	62
Table 14. Distributed lag linear regression models at the user level with random-effects predicting changes in friend count. Model 1 contains controls, and Model 2 adds self-disclosure, student type, and their interaction.....	62
Table 15. Descriptive statistics for the variables in the regression analysis of the audience counter study.....	73

Table 16. Linear regression result of the audience counter study.....	73
Table 17. Descriptive statistics for the variables in the regression analysis of the privacy checkup study.	76
Table 18. Linear regression result of the privacy checkup study.....	76
Table 19. Estimate the propensity score (the probability of accepting the introduction dialogue of the privacy checkup) using logistic regression.	78
Table 20. Covariate comparison between users who accepted the intro dialogue (treated) and those who rejected it (control) before and after propensity score matching.	80

LIST OF FIGURES

Figure 1. Research scope regarding the causes and consequences of self-disclosure in this dissertation. .	15
Figure 2. Average self-disclosure before and after relationship changes (Day 0).	52
Figure 3. Average self-disclosure scores against absolute dates for the two kinds of students.	56
Figure 4. Average self-disclosure scores against relative dates for the two kinds of students using the school start date as the reference point.	56
Figure 5. Predictive marginal means of self-disclosure of the two student groups during different time periods.....	58
Figure 6. Predictive marginal means of friend counts for the two student types at different time points...	60
Figure 7. Mock-up interface of the audience counter study.....	71
Figure 8. Screenshot of the introduction dialogue of the privacy checkup study.	75
Figure 9. Screenshot of the main dialogue of the privacy checkup study.....	75
Figure 10. Density graphs of propensity score for the treated group (i.e., accept the intro dialogue) and control group (i.e., not accept the intro dialogue) before matching (left) and after matching (right).	79

Chapter

1 Introduction

When people communicate with others in person or online they share information about themselves that helps others understand who they really are. This “act of revealing personal information to others” is known as *self-disclosure* ([Archer, 1980, p. 183](#)). With the growth of online social networking sites (SNSs), many traditional social interactions have shifted toward online environments. These sites offer new paradigms for interaction, particularly the ability to broadcast personal stories to friends (e.g., Twitter tweets or Facebook status updates). This concept of one-to-many sharing is called *broadcasting self-disclosure* ([Jourard, 1971](#); [Bazarova & Choi, 2014](#)). Several theories of computer-mediated communication suggest that verbal self-disclosure will be more important and common online than offline because of online anonymity and the lack of non-verbal cues to signal thoughts or feelings ([see Table 1 in Nguyen et al., 2012](#)). Empirically, people disclose significantly more in computer-mediated communication interactions than in offline communication ([Joinson, 2001](#); [Tidwell & Walther, 2002](#)). However, a recent review suggests that the difference between online and offline self-disclosure depends on factors such as the disclosure’s personality, context and the relationship between communication partners ([Nguyen et al., 2012](#)).

Greater levels of online self-disclosure can be important for individuals and the sites that host their communication. Substantial offline and online communication research suggests that self-disclosure significantly helps to form and maintain personal relationships. Sharing important parts of our lives improves relationships ([Oswald et al., 2004](#)) and causes others to like us ([Collins & Miller, 1994](#)); greater self-disclosure leads to greater liking of a conversational partner, increased feelings of closeness and deeper enjoyment of the conversation ([Sprecher et al., 2013](#)); and online self-disclosure increases intimacy among Facebook friends ([Park et al., 2011](#)). Furthermore, scholars have shown that self-disclosure benefits physiological health and psychological well-being ([Pennebaker, 1997](#)). For instance, disclosing traumatic experiences or stressful life events has been shown to improve one’s immune system or overall disease activity ([Pennebaker et al., 1988](#); [Smyth et al., 1999](#)). Writing about thoughts and emotions also helps people to cope with job loss and thus find a new job more quickly ([Spera et al., 1994](#)).

Self-disclosure also has significant implications for the success of social networking sites. Because relationship maintenance motivates many people to use SNSs, and because self-disclosure both reflects and enhances social relationships, people are more likely to be satisfied with sites that encourage self-disclosure ([Special & Li-Barber, 2012](#)). Interface elements on these sites influence how much people reveal about themselves. For example, between 2005 and 2014, Facebook increased the number of fields included in users' profiles, which allowed for greater self-disclosure ([Acquisti et al., 2015](#)). They also introduced interface elements such as the privacy checkup to encourage users to undergo a privacy checkup and become aware of the audiences who could see their information they share ([Albergotti, 2014](#)).

However, online self-disclosure can have effects beyond social relationships, most commonly when people share information to a wider audience than they had intended. For example, many companies use social networking sites to recruit new employees ([Segal, 2014](#)) and an estimated 37% use social networking sites to research job applicants ([CareerBuilder, 2012](#)). Therefore, individuals who disclose too much personal information or too many details in SNSs might damage their professional image and consequently influence how recruiters perceive them. Friend networks that encompass multiple social circles can also make self-disclosure challenging since it is difficult to manage separate impressions to different audiences ([Marwick & boyd, 2010](#)).

Despite its significance, some limitations exist in the self-disclosure literature (see reviews in [Cozby, 1973](#); [Nguyen et al., 2012](#)). First, many studies are based on analyses of interview data, self-reported questionnaires, or daily dairies gathered from participants. These approaches lack exhaustive coverage, since they include only a limited, predefined list of disclosure acts; it is not possible to list every aspect of self-disclosure in questionnaires or interviews. Second, because of the burden these approaches might place on interviewers and subjects, it is difficult to collect large samples and conduct longitudinal analyses. Third, these studies are mostly retrospective, so the results reflect subjects' selective recall, which are prone to distortions. Furthermore, much of the work on the link between self-disclosure and interpersonal relationships has focused on investigating self-disclosing behavior among reciprocal relationships and its effect on increasing tie strength at the dyadic level. Few studies explore self-disclosure broadcasted to one's entire social network and its effect on expanding that network.

In order to address these gaps and extend psychological research on self-disclosure, we studied people's self-disclosure behavior in SNSs. We focus on SNSs for two reasons. First, many social interactions and communications increasingly take place online. The explosive growth of the Internet and online social environments has opened many exciting opportunities for researchers to study the nature of self-

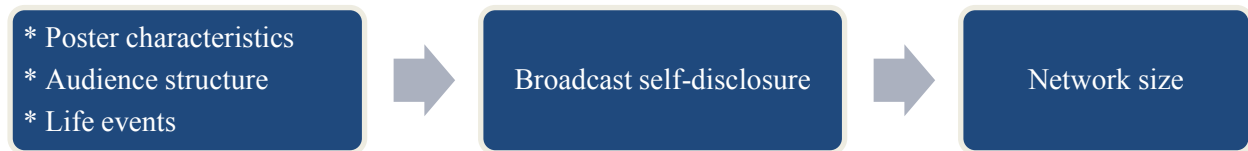


Figure 1. Research scope regarding the causes and consequences of self-disclosure in this dissertation.

disclosure. Specifically, these environments create a long-term archive of conversations and social network information, which can be used to quantitatively investigate self-disclosure. Second, although SNSs offer users a unified platform to present themselves and build social connections, these sites also introduces new challenges of self-disclosure because they allow users to share with multiple audiences at once. However, online self-disclosure and its connection with social outcomes are not yet well understood.

Given the importance of online self-disclosure and the research opportunities available on SNSs, the goal of this thesis is to understand the underlying mechanisms of online self-disclosure. In particular, I study the causes and consequences of broadcast self-disclosure through automatic language analysis. I develop a novel machine learning model to automatically measure self-disclosure in SNS communication at scale and use it to examine factors that might influence people’s broadcast self-disclosure at three levels: their personal characteristics, the structure of their online social networks, and the events in their lives. I also study the relationship between broadcast self-disclosure and social network growth. Figure 1 illustrates the research scope of this thesis.

1.1 Background Overview

Self-disclosure is the revelation of personal information ([Archer, 1980](#)). Self-disclosure can vary on several dimensions, including depth, breadth, and amount ([Altman & Taylor, 1973](#); [Derlega & Chaikin, 1977](#)). Depth of self-disclosure is the level of intimacy of information being shared. Disclosure can involve revealing vulnerable information, such as one’s worst fears or sexual orientation; on the other hand, it can also be shallow (e.g., “I like Chinese food”). Breadth dimension refers to the number of different topics or areas that are disclosed, whereas amount means the volume or frequency of self-disclosure. In this study, we focus on depth of self-disclosure and consider more intimate information to be more disclosure.

Self-disclosure can be classified into several types: descriptive, evaluative, affective, and topical ([Morton, 1978](#)). Descriptive self-disclosure describes facts about self; evaluative disclosure is about one’s attitudes

or opinions towards some objects or events; affective revelation is related to emotions and moods; topical disclosure refers to discussions of sensitive topics, such as political affiliation and sex life. This classification scheme inspired the development of some features in our machine learning model, including emotional valence and topic features, which we will describe in details in Chapter 2.

Although self-disclosure is good for relationship building and well-being and is intrinsically rewarding ([Tamir & Mitchell, 2012](#)), it may also endanger a discloser's privacy because of sharing information with others ([Altman, 1975](#)). To maximize rewards and minimize personal risks, a common disclosure strategy is to disclose within a dyadic boundary and share information with a trustworthy target ([Pearce & Sharp, 1973](#)). However, broadcast self-disclosure in SNSs contradicts the traditional understanding of self-disclosure based on dyadic interactions, since information is broadcast to one's entire network. There is a need to uncover the mechanisms and motivations behind this new type of self-disclosure behavior.

The Disclosure Decision Model was proposed to explain variations in self-disclosure across different social settings ([Derlega & Grzelak, 1979](#); [Omarzu, 2000](#)). According to the Disclosure Decision Model ([Derlega & Grzelak, 1979](#); [Omarzu, 2000](#)), people have five distinct motives for self-disclosure: social approval, intimacy, relief of distress, identity clarification and social control. Social approval is the default motive, for increasing social acceptance and liking ([Baumeister, 1982](#)). Intimacy disclosure is an attempt to develop closer relationships with others. Self-disclosure can also be used to relieve distress by talking about negative emotions or issues ([see also Rimé et al., 1998](#)). Identity clarification helps us define our identity to ourselves and others by talking about ourselves. Finally, disclosure for social control is often used in the situations where individual would like to regulate what others think about them in order to acquire rewards or benefits from the targets. Although some studies have been conducted to extend the decision model to explain self-disclosure on SNSs, the results are still unclear. This study untangles some aspects of these motivations by analyzing self-disclosure patterns of life events that evoke different and multiple motivations for self-disclosure.

Furthermore, since the main difference between broadcast self-disclosure and the conventional notion of self-disclosure is audience, a large portion of this study focuses on the effect of audience characteristics on self-disclosure. Several studies have investigated the relationship between audience structure and disclosure in SNSs ([Facebook, 2010](#); [Park et al., 2012](#); [Kivran-Swaine & Naaman, 2011](#); [Lin et al., 2014](#); [Choi & Bazarova, 2015](#)), but their findings are not consistent. For example, [Facebook \(2010\)](#) reported that users with more Facebook friends used less positive emotion words while [Lin et al. \(2014\)](#) found the opposite pattern. [Kivran-Swaine and Naaman \(2011\)](#) identified a negative correlation between network density and emotion words in Twitter tweets while [Lin et al. \(2014\)](#) discovered a positive relation in

Facebook status updates. These inconsistent findings suggest that more research is needed to understand the underlying mechanisms of audience structure in influencing self-disclosure.

1.2 Thesis Overview

This section presents a brief overview of the studies in this thesis. The details of each study can be found in Chapters 2 to 5.

1.2.1 Machine learning model of self-disclosure

Chapter 2 presents a novel machine learning model to measure self-disclosure in Facebook status updates. The model performs moderately and agrees with the judgments of trained coders (*Pearson* $r=.60$). Features of the model were derived from theories about the nature of personal self-disclosure. These features include message length, use of positive and negative emotional vocabulary, mentions of close social ties, use of non-normative language and discussion of topics differing in intimacy. I then apply the model to detect self-disclosure in all Facebook status updates in the subsequent analyses.

1.2.2 The relationship of poster and audience factors to self-disclosure

To validate the machine learning model of self-disclosure and to better understand online broadcast self-disclosure, in Chapter 3 we apply the model to two data sets containing approximately nine million de-identified Facebook status updates. We show that the results are consistent with those found or suggested by prior literature. Specifically, we focus on the relationship of personal characteristics and audience factors to self-disclosure. We find that status updates exhibited higher self-disclosure if the posters (i.e., the authors of the status updates) score lower on a self-reported scale measuring the trait of impression management, are women, and maintain networks of Facebook friends that are smaller, denser and of higher average tie strength.

1.2.3 Patterns of self-disclosure and network growth around life events

Chapter 4 further explores the relationship between social network factors and self-disclosure, especially whether changes in social network really influence self-disclosure. It examines changes in the degree to which people self-disclose after beginning or ending a romantic relationship or entering college—events that can significantly affect social networks. It demonstrates that users disclose more intimate contents when starting a new relationship and disclose less when breaking up with someone. It also shows that college students, especially freshmen, self-disclose more at the beginning of the school year. The results imply that social approval is the primary motivating factor for broadcasting self-disclosure. While seeking social approval motivates people in new relationships or entering college to self-disclose more, it makes people who break up disclose less because of the negative interpretation of breakup.

In Chapter 4, I also explore the social consequences of self-disclosure. Although self-disclosure is known to help people develop and maintain personal relationships at the dyadic level (e.g., [Collins & Miller, 1994](#); [Oswald et al., 2004](#)), we do not know whether broadcast self-disclosure (i.e., one-to-many) supports relationship formation as does private offline self-disclosure. Since broadcast self-disclosure is intended to be seen simultaneously by many different people in one's social network, I focus on its effect on expanding one's entire social network. Surprisingly, the results show that a higher level of broadcast self-disclosure correlates with a significantly lower increase of friend count in SNSs.

1.2.4 The effect of context collapse on self-disclosure

Social networking sites create a new problem for online self-presentation called *context collapse*, or the collapse of an individual's multiple audiences into one single context ([boyd, 2008](#); [Marwick & boyd, 2010](#)). Context collapse makes it more difficult for people to manage separate impressions of themselves to different audiences in online SNSs than in offline settings ([Marwick & boyd, 2010](#)). As a result, one important research question is: how does context collapse change people's self-presentation? To understand the causal relationship between context collapse and self-disclosure, Chapter 5 presents two controlled experiments. In one experiment, the existence of multiple audiences on SNSs was made salient by adding an audience counter in the status update composer. In the other experiment, a privacy checkup tool was shown to some users, which allowed them to review the privacy settings of their status update composer. We find that neither the audience counter nor privacy checkup dialogue influenced how much users self-disclosed in status updates. This implies that context collapse does not affect self-disclosure. However, a follow-up analysis of the privacy checkup dataset using propensity score matching indicates that it has a small negative effect on self-disclosure, which suggests that the treatments in the two experiments might not be strong enough to stimulate users' sense of multiple audiences.

1.3 Approach and Impact

We chose Facebook as the research site for several reasons. First, Facebook is the world's biggest and most popular social networking website, with almost 1.5 billion active users per month ([Facebook, 2015](#)). Facebook users generate roughly 55 million status updates ([Branckaute, 2010](#)) and 3.2 billion "likes" clicks and comments ([Facebook, 2012](#)) every day, and the average Facebook user spends 700 minutes and generates 90 pieces of content every month ([Branckaute, 2010](#)). Second, Facebook is a platform that includes different kinds of online social activities with different social relationships. It provides users a single environment to manage their social connections, such as family, college friends and co-workers, and various platforms for different kinds of social communication, including direct communication versus broadcast, and private versus public communication.

The research data in this thesis was de-identified, aggregated behavioral and social network information from Facebook server logs. The social network data provided information for audience structures and social outcomes of self-disclosure. All analyses were performed automatically and in aggregate such that no text or individual data was viewed by researchers. I used research methodologies from both computer science and social psychology—I applied natural language processing and machine learning techniques to measure self-disclosure in conversation logs, which is the source of the main variables in the models, and conducted statistical analyses to test hypotheses derived from social psychology literature.

This thesis makes significant theoretical contributions to the existing literature in linguistics and social sciences. First, while most linguistics research focuses on studying language itself (such as language structure or meaning), this work considers how language is used in social context. It advances our knowledge of how people use language to disclose personal details to their social relationships in online social networking environments where their social circles collapse into one. Second, it examines the conditions under which people change the degree of their disclosure and the social consequence of doing so. Third, the results of this research can be generalized more than past research. Natural language processing techniques and machine learning approaches were used to automatically analyze de-identified posts, and the findings are based on a general sample of online populations and large-scale data analyses.

This research also has practical contributions for improving user experience in social web. By knowing how users of social networking sites present themselves to multiple audiences, SNS designers can provide better affordances to users. Furthermore, the machine learning model of self-disclosure offers solutions for measuring self-disclosure at scale in SNSs and can be used as the basis for building tools to provide users active feedback and social norms information.

Chapter

2 Measuring Self-Disclosure in Social Media Content

Social networking sites (SNSs) offer users a unified platform to build and maintain social connections. Understanding when people feel comfortable sharing information about themselves on SNSs is critical to a good user experience, because self-disclosure helps maintain friendships and increase relationship closeness. In this chapter, we introduce a novel machine learning model to measure self-disclosure at scale in SNS communication and use it to understand the contexts in which it is higher or lower in subsequent chapters. Features in the model include message length, use of positive and negative emotional vocabulary, mentions of close social ties, use of non-normative language and discussion of topics varying in intimacy. They are derived from theory about the nature of personal self-disclosure. Performance of the model is moderate, and agrees with the judgments of trained coders (*Pearson* $r=0.60$).

2.1 Introduction

In the psychological literature, researchers have usually considered self-disclosure as a stable personal trait and measured or assessed it through self-reported questionnaires ([Cozby, 1973](#)). A typical example of self-disclosure questionnaires is Miller et al.'s Self-Disclosure Index ([Miller et al., 1983](#)), which consists of 10 items in 10 topic areas (e.g., “things I have done which I feel guilty about”, “what is important to me in life”, and “my worst fears”). Subjects respond to each item by indicating the extent to which the information has been disclosed to a target person, such as parents and friends. This approach has limitations. It usually measures disclosure to a specific target, so we cannot directly apply it to measure broadcast self-disclosure. It also has a coverage issue, since it includes only a limited, predefined list of disclosure acts; it is not possible to list every aspect of self-disclosure in questionnaires. Most importantly, we are interested in the language of self-disclosure, especially how people use language to reveal intimate information about themselves when communicating with others. That is, instead of considering self-disclosure as a stable attribute of a person, which does not change much over time, we consider it as a characteristic of a communication episode, which varies from communication to communication. Recently, some scholars have started to analyze self-disclosure in both face-to-face conversations and computer mediated communication by manually coding participants' conversations or

posts (e.g., [Joinson, 2001](#); [Barak & Gluck-Ofri, 2007](#)). However, human coding is not scalable for examining large archives of conversations produced in SNSs.

Given its importance to relationship building, an automated measure of self-disclosure in SNSs that can be applied at scale could help social scientists better understand the conditions that encourage or discourage self-disclosure; offer feedback to SNS members about whether their content is revealing more or less about themselves than they desire; and enable service providers to track how changes to site design influence users' self-disclosure. Our goals are to develop an automated measure of the degree to which users self-disclose in SNSs and use this measure to better understand the conditions that encourage or discourage self-disclosure in online settings.

2.1.1 Machine learning approach for self-disclosure measurement

Although several studies have demonstrated that it is possible to construct automatic self-disclosure text analyzers ([Bak et al., 2012](#); [Bak et al., 2014](#); [Balani & Choudhury, 2015](#); [Wang et al., 2015](#)), the models and approaches proposed in these studies were either domain-specific ([Balani & Choudhury, 2015](#); [Wang et al., 2015](#)), difficult to interpret ([Bak et al., 2014](#); [Balani & Choudhury, 2015](#)), or provided no ground truth against which to evaluate their accuracy ([Bak et al., 2012](#)). For example, the highly accurate self-disclosure classifier constructed by [Balani and Choudhury \(2015\)](#) uses over a thousand features, which makes it difficult to interpret why specific features predict self-disclosure. We seek to build a supervised machine learning model that can approximate human judgments about the degree to which people are revealing personal information in their online posts. A successful model should be accurate, parsimonious (i.e., use few textual features), interpretable, and domain-independent.

Supervised machine learning algorithms use statistical procedures (analogous to multiple regression) to map a set of input features to a set of output categories or numerical values. To build successful predictive models, some researchers develop new machine learning algorithms that can better recognize patterns or mappings between input features and output labels in sample data; others explore features that can better explain an underlying problem and represent its data (a.k.a. feature engineering). In this work, we follow the later approach, since our goals are to design a self-disclosure model with good performance and to understand the language or textual cues that people use to self-disclose. Instead of blindly adding thousands of features (e.g., unigrams and bigrams) into machine learning algorithms, we seek a small set of features that can be interpreted and that capture the underlying principle of self-disclosing behaviors.

Building and validating our machine learning model involved three major steps, which we describe in more detail below. First, human judges hand-coded the degree of self-disclosure in 2,000 sample

Facebook posts provided by social media users to the researchers with informed consent. Their judgments represent the training data and the “ground truth” for evaluating the accuracy of the machine learning estimates. Second, we represented the posts as a set of input features for the machine learning algorithms. One of the input features is message length, because producing sufficient communication is a prerequisite for self-disclosure. More importantly, we introduced four linguistic features which we believe are key ingredients of self-disclosure: emotional valence, the presence of certain topics, social distance between the poster and a person mentioned in the post, and how well the content of a post aligns with social norms. The output is a numerical value representing the degree of self-disclosure in a post. Finally, we constructed statistical machine learning models from the hand-coded data and then evaluated their accuracy.

2.2 Data Collection and Agreement Analysis of Coded Data

In this section, we describe how we operationalized the judgments of self-disclosure and collected 2,000 Facebook status updates with self-disclosure annotations from the posters (Facebook users recruited from Amazon Mechanical Turk) and trained judges. After that, we analyzed the agreement between the self-disclosure scores of posters and external judges.

2.2.1 Self-disclosure instrument

Many self-report questionnaires have been used to measure self-disclosure, including Jourard’s Self-Disclosure Questionnaire ([Jourard & Lasakow, 1958](#)), Miller et al.’s Self-Disclosure Index ([Miller et al., 1983](#)), and Rust’s Impression Management scale ([Rust & Golombok, 2009](#)). However, most questionnaires conceptualize self-disclosure as a stable personal disposition to reveal personal information. Recently, [Barak and Gluck-Ofri \(2007\)](#) developed a 3-item rating scale that evaluates the degree to which *a single online forum post* exposed the author’s personal information, thoughts, and feelings. However, their 3-item rating scale was designed for trained expert judges and does not have a strong internal consistency when used by untrained judges, such as general internet users and crowdsourcing workers, so we selected several items from Miller et al. Self-Disclosure Index and added them to the scale to improve internal consistency. In other words, our self-disclosure scale is based on definitions and questions adapted from Barak and Gluck-Ofri Self-Disclosure Rating Scale and Miller et al. Self-Disclosure Index.

In particular, we conducted pilot studies on Amazon Mechanical Turk (MTurk) to determine the appropriate set of questions and modifications (described more fully below). We selected and modified questions so that they can be used to measure posters’ self-disclosure in a single post rather than a stable

To what extent does this post involve (1: not at all; 7: completely)

- A. personal information about yourself or people close to you, such as accomplishments, family, or problems you are having?

 - B. personal thoughts on past events, future plans, appearance, health, wishful ideas, etc.?

 - C. your feelings and emotions, including concerns, frustrations, happiness, sadness, anger, and so on?

 - D. what is important to you in life?

 - E. your close relationships with other people?
-

Table 1. Self-disclosure measurement items for Facebook status updates as rated by Turkers.

trait. To reduce respondent burden, we created a reliable scale with only five items. In each round of pilot studies, respondents were asked to enter the text of one of their Facebook status updates and answer several questions using a 7-point Likert scale, ranging from 1 (“not at all”) to 7 (“completely”) (e.g., “To what extent does this post involve your feelings and emotions, including concerns, frustrations, happiness, sadness, anger, and so on?”). Previous research on self-disclosure has used coarser Likert scales (e.g., a 3-point scale from (Barak & Gluck-Ofri, 2007)), but we employed a 7-point Likert scale that allows for greater variance for later model training. After several rounds of testing with larger sets, we created a situational self-disclosure scale based on the five questions listed in Table 1. The composite value of the answers to these five questions represents the self-disclosure level in a specific update. The scale is reliable, with a Cronbach’s alpha of 0.72.

2.2.2 Collecting Facebook status updates and ratings from posters

In order to construct a dataset of Facebook status updates with hand-coded self-disclosure annotations, while also honoring users’ privacy and Facebook’s terms of service, we recruited active Facebook users from Amazon Mechanical Turk (<https://www.mturk.com>). Amazon’s Mechanical Turk (MTurk) is an online marketplace for crowdsourcing that allows requesters to post jobs and workers (Turkers) to choose jobs. Jobs on MTurk are known as a Human Intelligence Task (HIT).

Participants were paid \$0.50 US to submit and rate their most recent Facebook status update in terms of the degree of self-disclosure it contained. To control the annotation quality, our HIT only accepted workers from the United States who had 98% or more of their previous submissions accepted. Workers

	Status updates provided by Turkers	Turker	RA
1	It was so warm out on Saturday... why is there snow everywhere now? :C	1.8	1.6
2	Well, I got into the University of [omitted] MA program. More than 200 applicants, 24 spots. #1 program in the country. so there's my brag and I think I've accomplished enough for this year so can I just play animal crossing or sims for a few weeks, thank you.	5.6	5.2
3	There are few things I cherish more on this Earth more than leftover spaghetti.	5.2	1.7

Table 2. Examples of status updates and composite ratings by Turkers and RAs.

were shown an informed consent document in which they were notified that research assistants would be reading their submitted status updates (see Appendix A). This task was approved by Carnegie Mellon University’s Institutional Review Board (IRB). To encourage workers to take the task seriously, we first asked about their Facebook experience, including “How many days in the past week did you use Facebook?”, “How many friends do you have on Facebook?”, and “How many photos do you have on Facebook?” Participants were then asked to copy and paste their most recent English pure-text status update and to rate their post regarding its degree of intimacy for each of the five questions in Table 1. Appendix B contains a screenshot of the HIT.

Table 2 shows example status updates contributed by study participants and their composite ratings of self-disclosure. Although some ratings seemed appropriate, others seemed highly idiosyncratic. For example, the Turkers who contributed examples 2 and 3 evaluated them as having substantial self-disclosure (greater than 5 on the 7-point scale). Most experts, however, would consider pride about admission to a competitive graduate program (Example 2) to be more self-disclosing than a light-hearted statement about leftover spaghetti (Example 3). This observation led me to consider to collect ratings from trained expert judges.

2.2.3 Agreement analysis between posters and researchers

Our goal was to build an accurate machine learning model that could be used for examining self-disclosure on social networking sites. However, the accuracy of this model could be tempered by noisy training data as a result of individual differences in Turkers’ diligence in the judgment task or interpretation of the self-disclosure questions and the 7-point Likert scale. This problem of noisy data is

compounded because only a single, unique poster evaluated each of the 2,000 status updates. Furthermore, posters can only describe their intent, but not how an external audience would interpret and evaluate their posts. Indeed, people are poor at judging how others will interpret their online communication (Kruger et al., 2005). To mitigate this problem, external judges can act as proxies for intended readers or audiences for posts. Thus to reduce noise in the training data and to capture audience judgments, we supplemented posters' judgments of self-disclosure with those of trained external judges.

We recruited four research assistants (RAs, 1 male and 3 females) with diverse backgrounds from a research-oriented university. RAs were instructed to rate each update from an audience's point of view using the same five items as the posters (Table 1 above), wherein the word "you" was replaced with phrase "the poster." For example, the RAs' version of Question A in Table 1 was, "To what extent does this post involve personal information about the poster or people close to him/her, such as accomplishments, family, or problems the poster is having?" In other words, RAs were asked to imagine the poster's intent.

The four RAs initially coded a common set of 50 posts, and met to resolve any disagreements and reach consensus for each of the 50 posts. The average correlation of their ratings was 0.79 before discussion, which increased to 0.82 after discussion. After the training, the four RAs annotated the remainder of the 2,000 posts. Each status update was evaluated by at least two RAs. The expert judgment of a post was then computed by averaging the scores of the RAs who rated the post. The mean and median of the annotations were 2.52 and 2.12, respectively, and the standard deviation was 1.28.

The right-most column in Table 2 presents the average RAs' ratings for the three examples. Across the 2,000 messages, posters and RAs agreed moderately on the degree of self-disclosure displayed in messages ($r=.60$, $N=2000$, $p<.001$). This suggests that RAs or audiences could perceive posters' self-disclosure intent to a reasonable extent. Given this finding and the observation of noisy posters' annotations, we decided to build our machine predictive model based on RAs' annotations to ensure the consistency and quality of the model.

2.3 Machine Learning Model of Self-Disclosure

To automatically measure self-disclosure we introduced five key linguistic features, including post length, emotional valence, the presence of certain topics, social distance between the poster and a person mentioned in the post, and how well the content of a post fits into social norms. The output was a numerical value representing the degree of self-disclosure in a post. In the following section we explain the rationale and extraction process for each feature.

When computing topic, social distance, and social normativity features, we needed a large sample of status updates to serve as baseline text to understand the prevalence of topics, named entities, and phrases across Facebook. So, we collected a random sample of 8,011,980 English Facebook status updates posted between November 2013 and October 2014—a full year to capture all regular events and holidays. This dataset is referred to as the “one-year dataset” throughout this text. This dataset was de-identified and analyzed in aggregate on Facebook’s servers in accordance with the company’s data use policy; models were built from counts of terms. No text was viewed by researchers, except for the researchers’ own status updates in order to validate the data processing procedures described below. No Facebook user’s experience was changed by this data analysis.

2.3.1 Text processing and feature extraction

Post length is the number of words in a post. One component of self-disclosure is the amount of detailed personal information revealed. Revealing more details about oneself often requires writing more text rather than less. Thus, we expected that longer posts would be more revealing than shorter posts.

Positive emotion and negative emotion: According to the self-disclosure instrument in Table 1, self-disclosing behavior includes revealing emotions and feelings. [Wang et al. \(2015\)](#) have also demonstrated that emotion words are important predictors for emotional self-disclosure in health-support groups. Therefore, to assess the emotional valence of a post, we considered the frequency of positive and negative tokens. A post token was considered “positive” or “negative” if it was found in the positive / negative emotion dictionaries of the Linguistic Inquiry and Word Count program (LIWC) or matched positive / negative emoticons listed in Table 3. The Linguistic Inquiry and Word Count program (LIWC) is a popular tool that calculates the frequency with which words in a text match each of 68 dictionaries representing linguistic dimensions (e.g., pronouns, tense), psychological constructs (e.g., positive emotion), and personal concerns (e.g., leisure, death) ([Pennebaker et al., 2007](#)). In addition, Internet users frequently use punctuation-composited icons (emoticons) that resemble facial expressions to convey their feelings and emotions, so we selected and modified several common types of emoticons from the list of western emoticons on [Wikipedia \(2015\)](#) and categorized them into positive or negative emotions (Table 3). The lists of positive emotion emoticons include smiley (e.g., :-) and :}), laugh (e.g., :D and =D), playful (e.g., :P and xp), and wink (e.g., *) and ;)]; negative emoticons are sadness (e.g., :(and :c), crying (:’-(and :’(), angry (e.g., :-|| and :@), and disgust (e.g., D8 and v.v).

Emotion	Type of emoticon	List of emoticons
Positive	Happy	:~))
	Laugh	>:D :-D :D 8-D 8D x-D xD X-D XD =-D =D =-3 =3
	Playful	>:P :-P :P X-P x-p xp XP :-p :p =p :-P :P :-b :b
	Smiley	>:] :-) :) :o) :] :c) :>=] 8) =) :} :^)
	Wink	>:] ;-) ;) *-) *) ;-] ;] ;D ;^)
Negative	Angry	:- :@
	Crying	QQ
	Disgust	D:< D: D8 D; D= DX v.v D-':
	Sad	>:[:-(:(:-c :c :-< :< :-[:[:{ >.> <.< >.<

Table 3. Lists of positive and negative emoticons, adopted and modified based on the list of western emoticons on Wikipedia ([Wikipedia, 2015](#)).

Topic features: Different topics are often associated with different degrees of self-disclosure. Some topics, like physical appearance or work, potentially contain more personal details than other topics, such as weather and sports. Consider the following examples from the 2,000 Facebook status updates collected from MTurk. :

a1) *Can't wait for warm weather.*

a2) *People need to realize that telling me that I'm "too skinny" is offensive to me.*

The author of Example (a2) revealed more about herself / himself in the post than the author of Example (a1). In particular, Example (a1) is only about weather but nothing personal, while Example (a2) is about the author's feelings on others' judgments towards her / his body.

To identify common topics in status updates, we used the statistical generative method, Latent Dirichlet Allocation (LDA), which can be used to discover hidden topics in documents as well as the words associated with each topic ([Blei et al., 2003](#)). It analyzes large numbers of unlabeled documents by clustering words that frequently co-occur and have similar meaning into "topics."

Before training our LDA model, we went through several steps to pre-process and clean the one-year status update sample. Our experience suggests that this pre-processing and pruning results in far superior topic models than those from un-pruned data. Status updates were segmented into sentences and then

tokenized with the Apache OpenNLP library ([OpenSource, 2010](#)), stemmed with the Porter stemmer ([Porter, 2006](#)), and lowercased. We removed punctuations and replaced URLs, email addresses, and numbers with tags. Updates were then represented as an unordered set of unigrams (single words) and bigrams (word pairs).

Across all terms in the de-identified eight million status updates, 83.24% of unigrams only appeared once (4,724,533 / 5,675,743), and 1000 unigrams accounted for 29.17% of all text (60,463,634/207,275,054). For example, 10.24% of updates contain “love,” the most frequent unigram appearing in over 10% of status updates (820,380/8,011,980). Though “love” is a meaningful word, its sheer popularity makes it unhelpful in topic modeling, because it co-occurs with so many different terms in so many contexts. Similarly, very low frequency terms (e.g., “vaguebook” or “Jawar”) are not helpful, as they do not co-occur often enough with other terms to distinguish clear topics. This skew of words is a well-known phenomenon in natural language known as Zipf’s law ([Zipf, 1949](#)). Therefore, we pruned high and low frequency unigrams (those that occurred in more than 0.5% or less than 0.01% of the updates) and bigrams (those that occurred in less than 0.015% of the updates) to reduce noise and vocabulary size. In addition, we excluded all unigrams from a 500-word stopword list (e.g., “the” and “in”); bigrams were filtered if both words were stopwords. After pruning, 63.31% of the status updates had fewer than eight n-grams (5,072,623/8,011,980); these documents were too short for successful model training. Therefore, we built topic models from the remaining status updates ($N= 2,939,357$).

To identify topics in status updates, we built an LDA model that treats each status update as a document. The model was set to derive 80 latent topics; this setting produced topics with greater interpretability to human judges than models deriving 50, 60, 70, 75, 100, or 120 topics. Topic dictionaries were generated from the 500 terms most strongly associated with each topic, and two experts familiar with SNS content manually named each dictionary. Examples of topics derived from the LDA analysis include *Sports* (e.g., “football”, “player”, “score”), *Medical* (e.g., “doctor”, “hospital”, “blood”), *Food* (e.g., “cook”, “coffee”, “chicken”) and *Christianity* (e.g., “heaven”, “christ”, “the lord”). See Table 6 for additional examples. Each LDA topical feature calculates the frequency of words in a message that match its corresponding dictionary. We used this dictionary-based approach, rather than topic distributions of updates from the LDA model, because it can be applied quickly and at scale in real time environment.

Social distance: Dunbar’s circles of intimacy categorize people’s social connections into circles with different levels of intimacy ([Dunbar, 1992](#)). The first circle is the smallest but has the highest level of intimacy of all, mainly including family members and best friends. The number of people in this circle does not exceed five. The second circle consists of close friends that we can turn to for support when we

need it, which contains no more than fifteen people. The next two circles are general friends and acquaintances, such as people we know at school or at work. Inspired by Dunbar's circles of intimacy, we argue that post authors have an imaginary distance between themselves and each of the people referenced in the post, an estimate of the degree to which they participate in each other's lives. Moreover, Dunbar claimed that there is a cognitive limit to the number of stable friendships that one can maintain, suggesting intimate others have more psychological overlap with self and thus talking about close others is a signal of a higher level of self-disclosure. The fifth item in Table 1 also supports the idea. Consider the following examples (which are not from the corpus, but are used to capture the essence of our approach):

b1) My husband can't give up cigarettes.

b2) President Obama can't give up cigarettes.

Both examples have the same topic (someone's bad habits). However, it is obvious that (b1) discloses more personal information about the author and her circumstances than does (b2), since it refers to the author's husband with whom she presumably has a closer relationship than she does to the President. This example shows that the social distance between a poster and people mentioned in a post is an important self-disclosure indicator. Thus, we propose a novel feature measuring the average social distance between posters and the target(s) they refer to in the post. Prior studies have shown that count of first-person words (e.g., "I," "my," and "myself") can be an effective indicator of self-disclosure in both offline and online communication ([Derlaga & Berg, 1987](#); [Joinson, 2001](#)). In contrast to those studies, which only count first-person words, we considered all types of person references.

The social distance feature extraction process involved three steps. First, we identified and extracted all people mentioned in each of the 2,000 labeled status updates. Person references include singular and plural first-person pronouns (e.g., "I," "me," "our"), familiar nicknames (e.g., "babe," "darling," "honey"), family relations (e.g., "husband," "daughter") and friends (e.g., "buddy," "friend"), as well as named entities (e.g., "Harry Potter," "Michael Jackson," "Barack Obama"). Second-person and third-person words were not included because it was impossible to infer the social distance between a post author and second-person or third-person word without knowing its antecedent. Moreover, the person-nouns for which they were substituted would have been taken into account when we searched for all people mentions. Except for named entities, all other people words were extracted using a dictionary-based approach, since they comprised a limited set of words. Specifically, we utilized the first-person singular, first-person plural, family, and friend dictionaries in LIWC, and manually created a familiar nickname dictionary. The full list of familiar nicknames can be found in Appendix C.

Second, we identified named entities (NEs) and distinguished private ones from public ones. While a private name was defined as a person known to the post author, a public name referred to a celebrity (e.g., a singer or politician) whom the poster was unlikely to know personally. The person-name entity recognizer in the OpenNLP toolkit was applied to find all named entities in status updates. In order to differentiate private names from public ones, we introduced a semi-automatic approach to construct a celebrity name list from the one-year dataset. We first used the person-name recognizer to extract all named entities in the dataset, and then discarded those that occurred fewer than five times. This automatic process resulted in 9,629 unique entities. However, since the name recognizer was not 100% accurate, there were incorrectly identified entities in the list, such as “Be Safe,” “Merry Christmas,” and “God Bless.” To correct this error, we manually pruned the name list, which resulted in 8,434 unique person names. This final list was our celebrity dictionary (examples include “Robin Williams,” “Peter Pan,” and “Steve Jobs”). A named entity was categorized as public if it was included in the celebrity dictionary; otherwise, it was classified as private.

Third, we calculated social distance—the average distance between a poster and other people referred to in the post—for each of the 2,000 status updates. To calculate social distance, we placed people references into one of four categories, and assigned each a relative social distance score of 0, 1, 2, or 3 based on the likelihood the person reference participated in the poster’s life. Those who were more likely to be involved in the poster’s life would be assigned a shorter distance score, with 0 representing the poster and 3 representing members of the public. Although we used weights of 0 to 3 to represent social distance, any monotonic coding would produce similar results, as long as psychologically closer people were assigned lower weights. Formally, the social distance of a status update s , $social_distance(s)$, was defined as following:

$$social_distance(s) = \begin{cases} \frac{1}{N} \sum_{i=1}^N distance(p_i), & N > 0 \\ 3, & N \leq 0 \end{cases}$$

$$distance(p) = \begin{cases} 0, & p \in \{LIWC_I\} \\ 1, & p \in \{LIWC_we, LIWC_family, DIC_nickname\} \\ 2, & p \in \{LIWC_friend, NE_private\} \\ 3, & p \in \{NE_public\} \end{cases}$$

where $P = \{p_1, p_2, \dots, p_N\}$ denoted the set of people referenced in s ; $social_distance(s)$ was the arithmetic mean of $distance(p_i) \forall p_i \in P$ when P was a non-empty set, otherwise it was set to 3. $distance(p)$ was a case function that returned a value indicating the pseudo social distance between the author of s and the people mentioned, p , according to its category.

The function returned a distance of 0 when p belonged to the LIWC “I” dictionary, since first-person singular words referred to the author herself / himself. It assigned a distance of 1 when p was family (LIWC_family), or someone close to the author so that she / he used first-person plural words (LIWC_we) to indicate they did something together or used a familiar nickname (DIC_nickname) to refer to the person. Though personal pronouns may indicate other psychological phenomena (such as distancing with the “royal we”) (Pennebaker et al., 2003), we expect these uses to contribute an insignificant amount of noise and wash out at scale. Moving a bit further away from the social circle of the author were people whom the poster knew but was not so familiar with, including general friends (LIWC_friend) and private named entities (NE_private), which were assigned at 2. The final type of people references was celebrities (NE_public). We assumed most posters do not know celebrities personally, so the function returned a distance of 3 when p was found in the celebrity list. Table 4

Status updates provided by Turkers	Type of people references				Social distance	RA
	0	1	2	3		
Praise be to God! <u>My daughter</u> was finally released from the hospital after a week stint there and in ICU. Thanks for everyone's prayers!	1	1	0	0	0.5	7.0
Just completed <u>my</u> bachelors degree! It's been an amazing past 4 years, thank you to everyone who has been a part of it!	1	0	0	0	0.0	5.7
To celebrate the arrival of Spring break, going to be waking at 4AM to work a shift at Hannaford. Then <u>we</u> get to drive an hour and a half for a gallery opening <u>Sarah</u> is included in. <u>I</u> hope lots of people come and see her work and maybe someone will buy the piece.	1	1	1	0	1.0	4.3
As soon as <u>Oliver</u> woke up he asked to call <u>Lynda Davis</u> , watch frozen, and call <u>daddy</u> to talk about frozen - in that order.	0	1	2	0	1.7	3.1
Congrats to <u>Ben Walker</u> . He won a Genetics Society of America Undergraduate Travel Award to present his research findings at the 11th International Conference on Zebrafish Development and Genetics in June.	0	0	1	0	2.0	2.2
Gotta give it to <u>Clint Eastwood</u> , this looks like a musical that would actually be worth watching!! And <u>Christopher Walken</u> is it!!	0	0	0	2	3.0	2.2
It's a <u>John Legend</u> kind of night :)	0	0	0	1	3.0	1.8
Happy Friday. Hope everyone has a good weekend.	0	0	0	0	3.0	1.3

Table 4. Examples of status updates with counts of each type of people references, social distance features, and composite self-disclosure ratings by RAs.

(Note that these updates were provided by Turkers under informed consent.)

presents some examples of status updates illustrating how self-disclosure scores vary with social distance features. All the people references in the updates are underlined. The “type of people references” column shows count of people references in each of the four categories. The “social distance” and “RA” columns are the social distance features calculated based on the equations described above and self-disclosure scores assigned by RAs, respectively.

Social normativity: Text is less self-revealing when people are saying what everyone else is saying than saying something unique. For instance,

c1) I love my family.

c2) I hate my family.

Example (c) presents another case, where both status updates share a topic and target (“my family”). Although at first glance it seems that the levels of self-disclosure of (c1) and (c2) are the equal, after taking a deeper look most people would agree that (c2) contains a bit more private information about the author than (c1). This is likely because expressing love for family fits most people’s expectation. On the other hand, expressing hatred for family deviates from social norms. This observation suggests that self-disclosure relates to the appropriateness of post content with respect to social expectations and norms. Therefore we hypothesized that degrees of self-disclosure increase as content becomes more surprising, i.e., deviates from social norms or expectations.

We quantified social normativity as the difference between the language of a status update and the language of the Facebook community as a whole. Specifically, we built a statistical language model representing the linguistic usage of the community, and then calculated the cross-entropy of the update using the Facebook language model. A statistical language model is a probability distribution trained over word sequences (i.e., a corpus) which can be used to assess the probability of an order of words occurring in the corpus ([Chen & Goodman, 1996](#)). Cross-entropy is a measurement often used in natural language processing applications for evaluating how well a language model predicts a test word sequence. In other words, it can be used to gauge whether one’s post fits into a corpus. For instance, Danescu-Niculescu-Mizil and his colleagues compared users’ posts in an online community with all the posts in the community to argue that users adapt to communal linguistic norms overtime ([Danescu-Niculescu-Mizil et al., 2013](#)). We adopted a similar approach.

We first used the de-identified one-year corpus (i.e., the random sample of the eight million Facebook status updates) to construct a bigram language model with Good-Turing smoothing ([Good, 1953](#)). The model was built from the CMU-Cambridge Statistical Language Modeling Toolkit ([Clarkson &](#)

[Rosenfeld, 1997](#)) (Refer to [Chen and Goodman \(1996\)](#) for more details about n-gram language models and smoothing techniques.) This language model represented the social norms of the Facebook community, that is, it characterized how the general Facebook community would expect Facebook users to represent themselves. Given a status update s , we computed its social normativity based on the bigram language model $LM_{Facebook}$ as shown below:

$$social_normality(s) = -H(s, LM_{Facebook}) = \frac{1}{N} \sum_{i=1}^N \log P_{LM_{Facebook}}(b_i)$$

where $H(s, LM_{Facebook})$ was the cross-entropy of s under the $LM_{Facebook}$; s was composed of bigrams (b_1, b_2, \dots, b_N) ; $P_{LM_{Facebook}}(b_i)$ denoted the probability of the bigram b_i in $LM_{Facebook}$. A status update with a smaller social normativity value suggested its language looked less similar to the language on Facebook, which we believed indicated that the update would contain more self-disclosure.

2.3.2 Model construction and evaluation

The purpose of our evaluation was to contrast the performance of the machine learning models built using our proposed features with a feature set consisting of unigrams and bigrams. Feature sets of unigrams and bigrams are frequently used as a baseline for model evaluation. In order to assess the contribution of each proposed feature, we evaluated them separately and in combination. Our dataset was the 2,000 status updates collected from MTurk workers and annotated by RAs. Details of our results are below.

Given the input feature representation of a status update, we built a machine learning regression model that output a numerical value of the degree of self-disclosure. We used the sequential minimal optimization (SMO) algorithm for support vector machine regression ([Shevade et al., 2000](#)) implemented as the SMOreg procedure in Weka ([Witten et al., 2011](#)), a machine learning toolkit, to build the regression models. We used the default linear kernel with all other parameters also set to defaults. The dataset was randomly split into partitions for 10-fold cross-validation. We chose 10-fold cross-validation over leave-one-out cross-validation because they are similar in terms of the size of data points used for training (1,900 versus 1,999), and 10-fold cross-validation is more time-effective than leave-one-out validation. We reported accuracy in terms of the average Pearson correlation across the 10 folds between the RA-coded ratings and predicted self-disclosure.

Table 5 presents the accuracy results. For a baseline model, we stemmed the raw text, removed stop words, and kept unigrams and bigrams occurring five or more times as features. It had a correlation of .47 (Model 1), but required 814 features. Interesting results were achieved by more parsimonious models that

used post length (Model 2), positive/negative emotion (Model 3), social distance (Model 4) or the social normativity feature (Model 5). Although Models 2-5 performed worse than the baseline model (as indicated by the correlations of .37, .39, .31 and .17, respectively), our model performed better than the baseline when using the five features together (Model 6; correlation of .48).

Surprisingly, and contrary to our hypothesis, the social normativity feature can predict self-disclosure with a correlation of .17, though it is a positive predictor rather than a negative one. This suggests that a status update using language similar to the Facebook community was considered slightly higher in self-disclosure, perhaps because there is a small positive norm of self-disclosure on the site. Furthermore, the model built with 80 topic features (Model 7) achieved a correlation of .57, substantially better than the baseline. These moderate accuracy correlations support our assumption that post length, positive/negative emotion, social distance, social normativity, and topics are essential components and indicators of self-disclosure.

To understand the topics most relevant to self-disclosure, we further examined the top 10 ranked topics in Model 7 (see Table 6). We found that, for example, topics like *Politics* and *Memorial* were positive indicators of self-disclosure, while *Christianity* and *Deep Thoughts* were negative signals. This finding is consistent with our expectation: *Politics* typically relates to one's personal attitude or opinion toward an issue and *Memorial* is about missing someone, whereas *Christianity* and *Deep Thoughts* posts are often quotes about religion or the universe, which reveal less personal information. It is noteworthy that two high-ranked topics (*Family Relationships* and *Names*) overlap with the information in the social distance feature. This might be why adding social distance and social normativity features with the 80 topic features only results in minor accuracy gains over the topic model by itself (Model 8). Lastly, we built a model combining all the proposed features (Model 9) that achieved the highest correlation among all the experiments (.60). Given its adequate validity, we then applied Model 9 to detect self-disclosure for all the status updates in the later analyses.

	Feature set	Num. of features	Corr.
1	Baseline (unigrams + bigrams)	814	0.47
2	Post length	1	0.37
3	Positive/negative emotion	2	0.39
4	Social distance	1	0.31
5	Social normativity	1	0.17
6	Post length + positive/negative emotion + social distance + social normativity	5	0.48
7	Topics	80	0.57
8	Topics + social distance + social normativity	82	0.59
9	Post length + positive/negative emotion + topics + social distance + social normativity	85	0.60

Table 5. Evaluation results with alternative feature sets.

Topic	Sample vocabulary	Regression weight
Christianity	shall, christ, spirit, the lord, of god	-0.70
Birthday	love you, happy birthday, my baby	0.51
Family Relationship	husband, wife, my mom, marry, my dad, the best, my daughter, in law	0.50
Politics	country, nation, american, govern	0.33
Deep Thoughts	the world, human, earth, create, key, purpose, soul, inspire, life	-0.32
School	student, write, teacher, test, grade	0.29
Weekend plan	wait for, n't wait, relax, spent, time with, so excited, this weekend, yay	0.26
Memorial	miss, angel, rip, heaven, pass away	0.24
Names	mary, smith, jack, jame, johnson	0.23
Medical	doctor, hospital, blood, leg, surgery	0.23

Table 6. Top 10 ranked topic features and their corresponding sample vocabulary in the model trained with 80 topic features.

2.4 Discussion

In this chapter, we developed a supervised machine learning model to detect the degree of self-disclosure in status updates. Through the process of building the model, we demonstrated that message length, emotional valence, the presence of certain topics, social distance between a poster and people mentioned in a post, and content alignment with social norms were important constituents of self-disclosure. Our model not only performs better but also uses many fewer features than an n-gram model.

The model evaluation results show that topic features better explain the variance in predicting human judgments of self-disclosure than the other four feature types. We interpret this difference in two ways. First, the 80 topic features are substantially more than the number of features for post length, emotional valence, and social distance or social normativity. Second, the topic features overlap with some aspects of the other four features. For example, we noted that some topic features are similar to the emotion valence features and the components used in the social distance feature, such as family and first-person words. Model 9 in Table 5 provides evidence for this interpretation: it shows that adding the other four features with the 80 topic features marginally improved model accuracy over Model 7, which only used the topic features.

2.4.1 Limitations and future directions

Selection bias in the Amazon Mechanical Turk sample may have weakened the model. We know little about workers who did not choose to participate in the study or how representative our sample is. By virtue of their online employment, these workers may be more technologically savvy or spend more time on Facebook, and thus their self-disclosure behaviors and perceptions may be different from people who use Facebook less often. Furthermore, workers were asked to select their most recent post but may not have, and the sample was limited to Facebook English posts collected at a particular time. Future work should gather ratings from a more representative sample and test the generalizability of the model on other SNS platforms.

Although our self-disclosure machine learning model performs reasonably well, there is still room for improvement, given that the average annotation correlation among RAs is 0.7, which can be considered the upper bound for model performance. Our current approach utilizes a linear kernel to train the model, which assumes features are independent. A next step would be to consider combinations or interactions among features. Moreover, as we pointed out earlier, some topic features captured concepts or information similar to those in the social distance feature. Thus, another improvement might be to remove redundant features or disentangle the relationships among features.

Chapter

3 Personal and Audience Factors Related to Self-Disclosure

To demonstrate the validity of the machine learning model of self-disclosure introduced in Chapter 2, as well as advance our understanding of online self-disclosure, we used the model to replicate empirical patterns found in experimental and survey research on self-disclosure or suggested by network structure theory. We applied the model to two data sets containing almost nine million de-identified Facebook status updates. In particular, we focused on individual differences among poster and audience factors that might affect self-disclosure. Results show that status updates exhibited higher self-disclosure if the authors score lower on a self-reported scale measuring the trait of impression management, are women rather than men, and maintain networks of Facebook friends that are smaller, denser and of higher average tie strength.

3.1 Poster Characteristics Influencing Self-Disclosure

3.1.1 Personality: impression management

Goffman says in *The Presentation of Self in Everyday Life*, “When an individual appears in the presence of others, there will usually be some reason for him to mobilize his activity so that it will convey an impression to others which it is in his interests to convey” ([Goffman, 1959](#)). This phenomenon is known as *self-presentation* and sometimes called *impression management* ([Goffman, 1959](#); [Schlenker, 1980](#)), which refers to the process through which people try to control the images others form about them. Impression management is generally considered the inverse of self-disclosure, by controlling the personal information one reveals. To assess individuals’ desire to manage the impressions they make on others and appear socially acceptable, researchers have developed self-report impression management scales to measure this concept as a stable personality trait, such as the Self-Monitoring scale ([Snyder, 1974](#)), the Balanced Inventory of Desirable Responding ([Paulhus, 1991](#)), and Rust’s Impression Management scale ([Rust & Golombok, 2009](#)). Example items in these scales include “I have never been dishonest,” “I have never dropped litter on the street,” and “I have some pretty awful habits” (reversed). These items suggest that people with a stronger desire to manage impressions would tend to hide the truth about themselves from others if they believe it would damage their images in others’ eyes. The impression management model proposed by [Leary and Kowalski \(1990\)](#) suggests that self-disclosure can endanger people’s

impressions of the discloser, since it involves the revelation of one's internal world, which usually consists of personal information or emotions that are socially awkward or morally questionable. We expect this tendency would influence how much individuals are willing to disclose to others, especially in a wide-audience environment such as Facebook status updates. Thus, we hypothesize:

H1: Individuals with a stronger desire for impression management will self-disclose less.

3.1.2 Gender

In the United States it is both a cultural stereotype and an empirical reality that women self-disclose more than men. A meta-analysis involving over 23,000 people in 205 studies found that women on average were more self-revealing than men (the average weighted effect size $d=.18$) ([Dindia & Allen, 1992](#)). Women self-disclosed more when demands for positive self-presentations were lower, including when talking to other women ($d=.35$) rather than men ($d=.00$) and when talking to friends ($d=.28$), spouses ($d=.22$) or parents ($d=.25$) rather than strangers ($d=.07$). In this study, we reexamine the following hypothesis:

H2: Women will self-disclose more than men.

3.2 Audience Factors Influencing Self-Disclosure

Audience structure can affect language use during social interactions ([Herring, 2007](#)). SNS offers users a unified platform to build and maintain various kinds of social connections ([Marwick & boyd, 2010](#); [Parks, 2010](#)), which can help researchers understand and compare how individuals adjust their self-disclosure according to various audience factors.

3.2.1 Public versus private communication

Communication through SNSs can be distinguished based on how directed and public the interaction is ([Burke et al., 2010](#); [Bazarova et al., 2012](#)). Whereas *directedness* measures whether the target of the communication is a particular friend, *publicness* measures the possibility that others might see an individual's communication and the number of others who might see it. Facebook status updates are undirected, since they are typically published to the entire social network of a poster and not targeted at any specific person. However, the degree of publicness of status updates depends on the number of friends in an individual's online network. In other words, having more friends implies that updates are more "public." Since self-disclosure means revealing private, personal details and people have less control of who sees their status updates when they have more friends, we hypothesize:

H3: Network size negatively correlates with self-disclosure.

3.2.2 Closeness to communication targets

Empirical studies of dyadic relationships show that people reveal fewer personal details to acquaintances than to close friends ([Collins & Miller, 1994](#)). We expect a similar result when considering one's social network as a whole. People with stronger ties in their networks should be more comfortable disclosing:

H4: Average tie strength positively correlates with self-disclosure.

3.2.3 Context collapse on social networking sites

Much of what we know about self-disclosure comes from studies of dyads (e.g., [Collins & Miller, 1994](#); [Oswald et al., 2004](#)); we know less about when people self-disclose to wider audiences, such as on social network sites. These sites allow people to share with others from many parts of their lives at once, a phenomenon known as *context collapse* ([boyd, 2008](#); [Marwick & boyd, 2010](#)). Context collapse may cause people to self-disclose fewer details, because they would feel uncomfortable sharing intimate information that may be appropriate for family and friends with relative strangers in their networks. That is, they might self-censor and only present information appropriate to the lowest common denominator ([Hogan, 2010](#)).

Context collapse online makes impression management challenging ([Marwick & boyd, 2010](#)). People have to meet the expectations and interests of many different audiences. Given that people's networks are comprised of both weak and strong ties, they may self-disclose less as their networks become more diverse. The degree of context collapse is likely signaled by *network density*, or the interconnections among the ties in one's social network. In other words, higher network density suggests that a user's friends are more connected and thus the user has fewer disconnected clusters. Therefore we hypothesize:

H5: Network density positively correlates with self-disclosure.

3.3 Predicting Self-Disclosure in Status Updates from Poster and Audience Characteristics

In this section, we examine the relationships of poster characteristics and audience factors with self-disclosure. Unlike prior research that considers these relationships at the dyadic or message level, we examine them at the personal network level by, for instance, averaging self-disclosure of a person's status updates in total. We also consider the average tie strength that person has with all of her / his Facebook

friends because status updates are not targeted at any specific person and may be visible to all Facebook friends of a user.

3.3.1 Poster characteristics and self-disclosure

To test the relationship of the self-reported trait of impression management (H1) and gender (H2) with self-disclosure, we utilized the dataset from the myPersonality project (<http://mypersonality.org/>). The project, founded by [Kosinski et al. \(2013\)](#), uses a Facebook app to collect data from Facebook users, such as their profile information and social network statistics, which it de-identifies and combines with personality scores measured by questionnaires that participants chose to complete. Participants explicitly consented to sharing this data for research purposes; the data does not include any information about those participants' friends. Specifically, the dataset contains users' status updates as well as their demographic information and self-reported impression management scores based on Rust's Impression Management scale ([Rust & Golombok, 2009](#)). We applied our machine learning model to measure self-disclosure in the users' updates, computed an average self-disclosure value for each user based on all her / his updates, and compared the average values with the corresponding self-report impression management scores. Analyzing the data from 2,878 users in the myPersonality dataset, we found a correlation of -0.19 ($n=2,878$, $p<0.0001$), which shows a negative relationship between self-reported desire for impression management and self-disclosure. This finding confirms Hypothesis 1. We also calculated the correlation of users' gender (1 for male and 0 for female) with their average self-disclosure values. We found that men self-disclose less than women, with a correlation of -0.23 ($n=153,726$, $p<0.0001$). This finding confirms Hypothesis 2.

3.3.2 Audience factors and self-disclosure

The automatic self-disclosure model was applied to a new dataset of all posts written by a random sample of 412,470 active, English language Facebook users for approximately one month in late 2014. All data was again de-identified and analyzed in aggregate on Facebook's servers; texts were not viewed by researchers and user experience on the site was unchanged. We included demographic information as control variables, including gender, age and the number of days users logged into Facebook in the past month. While gender was a binary variable with one (1) indicating male and zero (0) for female, age and log-in days were continuous, numeric variables. We also included a snapshot of users' social network size and structure at the beginning of the data collection period.

Dependent Variable

- Self-disclosure: The average of the machine-coded self-disclosure values of all of a user’s status updates.

Independent Variables

- Social network size: The number of friends a user had in the beginning of the data collection period.
- Social tie strength: We estimated tie strength between each user in the sample and all of his or her friends, using counts of communication frequency and other dyad-level variables, substantively identical to the techniques described in [Gilbert and Karahalios \(2009\)](#) and [Burke and Kraut \(2014\)](#).
- Social network density: The number of friendship connections among a user’s friends. We normalized this measure by the total number of possible links among a user’s friends to correspond to the portion of possible connections within a user’s friend network that were actually connected.

Except for the binary variable *Male*, numerical control and independent variables were standardized and centered, with a mean of 0 and standard deviation of 1. Additionally, we took the log of the variable *Network size* before it was standardized, since it had a skewed distribution. Table 7 reports the descriptive statistics for the variables entered into regression models before standardization.

Variable	Mean	Median	Std. Dev.	Min	Max
Age	35.65	32	14.24	14	114
Number of logins	26.40	28	4.26	0	28
Network size	492.76	329	558.19	0	4,968
Tie strength	0.32	0.31	0.05	0.07	1
Network density	0.09	0.07	0.06	0	1
Self-disclosure	2.50	2.31	.83	1	7

Table 7. Descriptive statistics for the variables in the regression analyses of the effects of audience factors on self-disclosure.

3.3.3 Who self-discloses more?

Table 8 presents five linear regression models predicting self-disclosure. Model 1 reports the effects of the control variables. In the rest of the models, we tested hypotheses regarding social network features and self-disclosure. Because *Network size* correlates with *Network density* ($r=-0.32$) and *Tie strength* ($r=-0.53$), we first tested the effects of the three network variables separately in Models 2, 3, and 4. We then analyzed their effects together in a single model (Model 5). The intercept in the models represents a woman with all numerical variables at their means who would disclose at a level of 2.595 on a 1 to 7 scale. Betas represent the effect on self-disclosure from a binary variable having a value of 1, or a one standard deviation increase in continuous independent variables. We also reported R-squared values in Table 8. Although the values are small, the outcome we were predicting (i.e., self-disclosure in one's language) is relatively subtle. Many things can affect one's language, such as culture, education level, and job. Our goal was to see if there was a reliable relationship.

Model 1 shows that males revealed significantly less about themselves in their status updates than females (2.319 versus 2.595), and that older posters disclosed more than younger posters. However, the significant negative beta for *Number of logins* suggests that the more active someone is on Facebook, the less he or she self-discloses.

In Model 2, we found that when controlling for demographic information and activity level, users' social network size negatively predicted their self-disclosing behavior. Self-disclosure levels decreased 0.01 point for users who had one standard deviation more friends. These findings confirm Hypothesis 3.

In Model 3, we investigated the effect of *Tie strength*. The result demonstrates that the closer users were to their friends, the more they self-disclose in status updates. This finding confirms Hypothesis 4.

Model 4 shows a positive correlation between *Network density* and *Self-disclosure*. That is, when a user has more friends who are also friends with each other, she / he would be more willing to share personal details. This finding confirms Hypothesis 5.

In Model 5, we examined the simultaneous effects of the three network variables on self-disclosure. While the effects of *Tie strength* and *Network density* were similar to those in Model 3 and 4, the effect direction of *Network size* changed from negative to positive. Possible explanations for this surprising result will be discussed in the next section.

Control/Independent Variable	Model 1		Model 2		Model 3		Model 4		Model 5	
	Beta	SE	Beta	SE	Beta	SE	Beta	SE	Beta	SE
Male	-.276***	.003	-.275***	.003	-.267***	.003	-.276***	.003	-.267***	.003
Age ¹	.100***	.001	.097***	.001	.091***	.001	.100***	.001	.093***	.001
Number of logins ¹	-.043***	.001	-.042***	.001	-.045***	.001	-.043***	.001	-.046***	.001
Network size ²	-	-	-.010***	.001	-	-	-	-	.007***	.002
Average tie strength ¹	-	-	-	-	.030***	.001	-	-	.033***	.002
Network density ¹	-	-	-	-	-	-	.004**	.001	.003*	.001
(Intercept)	2.595***	.002	2.595***	.002	2.592***	.002	2.596***	.002	2.592***	.002
R ²	0.0429		0.0430		0.0440		0.0429		0.0441	
Number of observations						412,398				

1: Standardized and centered. 2: Logged (base 10), standardized, centered.

*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

Table 8. Results of the regression analyses predicting self-disclosure from audience factors.

In order to test whether adding the three audience factors as predictor variables significantly improves Model 1, we conducted the likelihood ratio test to evaluate the difference between Model 1 and each of the four audience models. The tests show that Model 2 ($p < 0.0001$), Model 3 ($p < 0.0001$), Model 4 ($p < 0.001$), and Model 5 ($p < 0.0001$) fit the data significantly better than Model 1.

3.4 Discussion

In this study, we used the machine learning model of self-disclosure developed in the last chapter to replicate patterns from previous empirical work and theory. We showed that women self-disclose more than men, and people who more strongly desire to manage the impressions they make on others self-disclose less. We then demonstrated that social network size negatively correlates with self-disclosure, while network density and average tie strength positively correlate with self-disclosure. Most of the results are consistent with those found or suggested by prior literature and thus validate the effectiveness of the machine learning model we proposed.

One unexpected result in our analyses is that the estimates of the effects of network size in Model 2 and 5 have different signs. Although network size correlates with tie strength and network density, we

confirmed that multi-collinearity is not a problem, with all the variance inflation factors less than 1.7. The result may be substantive, rather than methodological, challenging our assumptions about the meaning of the network variables and how they affect self-disclosure. Although we hypothesized that a larger network size would lead to less self-disclosure because it makes communication more public, it may be that people believe that posting to larger networks exposes messages to weaker ties. Even though network size was designed to measure publicness, it grows by disproportionately adding weaker ties into the network ([Meo et al., 2014](#)). Thus size and average tie strength are intrinsically linked. As a result, when average tie strength is held constant (Model 5), adding more people to the network seems to lead to an increase in self-disclosure.

This study not only replicates empirical patterns found in previous research but also extends existing literature in social sciences and linguistics. It advances our knowledge of how people self-disclose and maintain relationships in SNS by utilizing machine learning to analyze a large archive of online communication text. Most early research on self-presentation or self-disclosure in online environments focuses on dyadic contexts and online dating sites. For example, scholars have investigated how online daters manage their profile presentations to attract potential dates ([Ellison et al., 2006](#); [Hancock et al., 2007](#)). In recent years, studies have begun to explore how people perform to their entire social network, not just potential dates (e.g., [Mehdizadeh, 2010](#); [Bazarova et al., 2012](#); [Barash et al., 2010](#); [Park et al., 2011](#)). Self-presentation to one's social network differs from self-presentation to potential dates. Online dating services target the development of romantic relationships, typically among dyads of roughly the same age. In contrast, online social networking services support people as they present themselves to a variety of partners with various types of social relationships ([Parks, 2010](#)).

Moreover, our research may be more generalizable than past research on online self-disclosure (e.g., [Park et al., 2011](#); [Bazarova et al., 2012](#); [Special & Li-Barber, 2012](#)), since it was based on a diverse, large sample of online communication. As a sensitivity test, we replicated the analyses reported here on de-identified, aggregated posts from Facebook users in Australia and Singapore, and found similar results. Using the automatic self-disclosure model introduced in Chapter 2, we will also be able to develop and test more theories regarding online self-disclosure.

The findings in this study also have practical implications for improving user experience in the social web. If web designers know how users of social networking sites navigate multiple audiences to manage impressions, they can improve their services by providing better affordances. For example, when network size and diversity become large enough that a person might not feel comfortable sharing personal news with friends, the site might nudge that person to share to a smaller group or custom list of friends.

3.4.1 Limitations and future directions

The findings reported here are based on a static view of the relationship between audience network structure and self-disclosure. We can only make correlational claims, not causal ones. One future research direction would be to perform a controlled experiment, in which participants' online network size and diversity would be made more or less salient in order to examine its effects on self-disclosure. Another direction would be to analyze audiences' responses to posters, so that we will have a better understanding of how audiences perceive and react to self-disclosure and whether they interpret the presenter's messages in the same way that the presenter intended. Are posts that are higher in self-disclosure perceived as higher quality by friends of the poster? Or are other post features more important? The answers to these questions would help site designers understand the degree to which context collapse affects the quality of post inventory.

Chapter

4 A Longitudinal Study of Online Self-Disclosure around Life Events Related to Social Network Disruption

Social networking sites (SNSs) offer new ways for people to share personal stories with their friends. In particular, people going through major life changes, such as starting college or going through a romantic breakup, can reach out to a network of friends for support. In a pair of longitudinal studies using de-identified, aggregate behavioral data and a novel machine learning model for self-disclosure, we demonstrate that these major life events are related to changes in how people present themselves online. In particular, people increase self-disclosure to their networks at the start of a new romantic relationship and decrease it when going through a breakup. Students self-disclose more at the start of the academic year, and incoming college freshmen (who may be trying to make new friends) do so at higher levels than college sophomores. We also show that self-disclosure and network growth negatively correlate among college freshmen, indicating potential tension between audience size and disclosure.

4.1 Introduction

Self-disclosure is a well-studied mechanism in human communication that refers to the “act of revealing personal information to others” ([Archer, 1980, p. 183](#)). It maintains and develops personal relationships (e.g., [Collins & Miller, 1994](#); [Oswald et al., 2004](#)), and creates trust, which improves relationships beyond the content exchanged. Self-disclosure benefits health and psychological well-being, by helping a person cope with emotions and major life events (e.g., [Pennebaker et al., 1988](#); [Pennebaker, 1997](#); [Smyth et al., 1999](#); [Spera et al., 1994](#)). The importance of self-disclosure has led scholars to attempt to better understand the contexts in which people reveal a higher or lower level of intimate information about themselves.

With the growth of online social networking sites (SNSs) many traditional social interactions have shifted toward online environments. These sites offer new paradigms for interaction, especially the capability to broadcast personal stories to friends (e.g., tweets on Twitter or status updates on Facebook.) This concept of one-to-many sharing is called “broadcasting self-disclosure” or “public self-disclosure” ([Jourard, 1971](#); [Bazarova & Choi, 2014](#)). Several theories of computer-mediated communication suggest that the

frequency and importance of written self-disclosure increase in online environments increase because non-verbal cues are absent. However, the effects of broadcasting self-disclosure and its connection with social outcomes are not yet well investigated.

Factors at multiple levels influence the circumstances in which people choose to self-disclose ([Greene, 2006](#)), including stable individual differences such as demographics, personality and culture ([Dindia & Allen, 1992](#); [Zhao et al., 2012](#)), and social network factors such as social relationships between disclosers and disclosure targets ([Collins & Miller, 1994](#)). [Wang, Burke, et al. \(2016\)](#) show that individual differences and social network characteristics correlate with the degree to which people verbally self-disclose online. Women, for example, and those with denser social networks disclose more intimate personal information than do men or people with loose social networks. However, like most research on the determinants of self-disclosure, their findings are based on cross-sectional data, which limits accurate assessment of causal relationships among variables (e.g., does network structure influence self-disclosure or vice versa?).

The current research further investigates the relationship between social network factors and self-disclosure, specifically whether changes in social network really influence self-disclosure. Using a longitudinal design, we apply an automated machine learning classifier to millions of de-identified Facebook status updates to determine levels of self-disclosure before and after events that are likely to disrupt a person's social network: beginning or ending a romantic relationship, or entering college. Changes in a romantic relationship, for example, do more than alter the connection with a boyfriend or girlfriend: they impact people's connections to their partner's social network. Similarly, when students enter college, they often reduce contacts with their high school mates and make new friends in their new environment. By analyzing the same people longitudinally as their networks change, this design holds unobserved stable individual differences constant and can establish a temporal order of events (though the design is not as strong as a random-assignment experiment in establishing causal relationships). Furthermore, because prior research has not established whether positive or negative experience prompts more self-disclosure ([Rimé et al., 1998](#); [Pasupathi et al., 2009](#)), this research compares self-disclosure after beginning and ending a relationship. The contrast can increase our understanding of the role of emotional valence in causing self-disclosure.

This research also considers the social consequence of broadcast self-disclosure in SNSs. Although self-disclosure is good for relationship development and maintenance, most prior research has examined the link between self-disclosure and relationships at the dyadic level, i.e., among pairs who can reciprocate both information exchanged and relationship closeness. However, it is unclear whether research on

reciprocating dyadic self-disclosure generalizes to the effects of self-disclosure broadcast to an entire social network and its effect on expanding that network. This study addresses this issue by observing new college students who often want to form new friendships. Examining the association between broadcast self-disclosure and change in network size provides some evidence about whether this strategy for making new friends works.

4.1.1 Life events and self-disclosure

People are selective when it comes to disclosing their life experiences ([Pasupathi et al., 2009](#)). The Disclosure Decision Model explains variations in self-disclosure across different social settings ([Derlega & Grzelak, 1979](#); [Omarzu, 2000](#)). According to the model, people have five distinct motives for self-disclosure: social approval, intimacy, relief of distress, identity clarification and social control; social approval is the default motive ([Baumeister, 1982](#)). That is, people self-disclose to increase social acceptance and liking, develop closer relationships with others, relieve distress by talking about negative emotions or issues ([see also Rimé et al., 1998](#)), define our identities to the self and others by talking about ourselves, and regulate what others think about them in order to acquire rewards or benefits from the targets. Although these seem to be major motives for self-disclosure, researchers do not agree about the types of life experiences that are associated with higher degrees of self-disclosure; different events might evoke different or multiple motivations for self-disclosure. Therefore, we present an exploratory study to identify event characteristics that correlate with changes in the degree of self-disclosure. The three events—beginning or ending a romantic relationship, or entering college—can help us untangle these motivations.

Numerous diary studies show that people are more likely to disclose life experiences with higher emotional intensity ([Rimé et al., 1998](#); [Pasupathi et al., 2009](#); [Garrison et al., 2012](#); [Ryan & Kahn, 2015](#)). For instance, [Pasupathi et al. \(2009\)](#) reached this conclusion in a study in which participants wrote down the most memorable events in their day for seven days, rated the intensity of the emotions associated with each event, and indicated whether they had disclosed the events to others. Similarly, [Rimé et al. \(1998\)](#) showed that people disclosed emotional events more often, regardless of gender, culture, or age. A theory of narrative identity helps explain this finding: it posits that people make meaning of their lives and construct an identity by integrating life experiences into evolving story of self ([McAdams, 2011](#)). In doing so, they frequently share their experiences with others, helping to define themselves and cope with their emotions ([McLean et al., 2007](#); [Pasupathi, 2001](#); [Rimé et al., 1998](#)). According to this account, emotional recovery relies on self-disclosure.

We expect that people who change relationship status or enter college will disclose more, as these events are emotionally intense yet also typically involve a significant change in social network. When people

break up, they might be motivated to disclose more for identity clarification and stress relief; conversely, they might disclose less because of perceived negative impact on social approval. On the other hand, individuals who start a new relationship might disclose more in order to gain social approval, clarify an individual identity, or enhance intimacy. Finally, people might be more likely to increase self-disclosure when moving to college in order to gain social approval from new friends, clarify an individual identity, relieve stress, or increase intimacy with classmates.

Although people talk more about emotional events, the significance of emotional valence to self-disclosure is unclear. While [Rimé et al. \(1998\)](#) imply that people disclose more about important events regardless of valence, [Pasupathi et al. \(2009\)](#) report that people are especially likely to disclose emotionally negative events. Recently, researchers have begun to consider emotional valence in broadcast self-disclosure on SNSs (e.g., [Gross & Acquisti, 2005](#); [Thelwall, 2008](#)). In particular, [Bazarova et al. \(2012\)](#) examined posters' language in Facebook status updates, wall posts, and private messages through automated content analysis. They reported that posters used significantly fewer negative emotion words in status updates than in wall posts and private messages, which suggests that people are more likely to publically disclose emotionally positive events yet reserve emotionally negative events for private communications. Our research contrasts emotionally positive events (i.e., beginning a relationship) with emotionally negative events (i.e., breaking up) to better understand the role of emotional valence in prompting self-disclosure. Therefore, we ask:

RQ1: What types of changes in people's lives prompt them to broadcast self-disclose?

4.1.2 The social consequences of self-disclosure

A main focus of self-disclosure research has been the effect of self-disclosure on various relational outcomes. Social penetration theory posits that self-disclosure helps to develop interpersonal relationships ([Altman & Taylor, 1973](#)), and its propositions have been empirically supported. For example, research shows a positive relationship between self-disclosure and dyadic trust ([Larzelere & Huston, 1980](#)), friendship maintenance ([Oswald et al., 2004](#)) and interpersonal liking ([Cozby, 1972](#)). In a meta-analysis of self-disclosure literature [Collins and Miller \(1994\)](#), identified three aspects of self-disclosure and liking: (1) people disclose more to those they like, (2) people like those who disclose more, and (3) disclosers tend to increase their liking of those to whom they have disclosed. Other researchers have extended these findings to computer-mediated communication (e.g., [Joinson, 2001](#); [Jiang et al., 2011](#); [Utz, 2015](#)).

Although self-disclosure has been shown to develop and maintain relationships (e.g., [Collins & Miller, 1994](#); [Oswald et al., 2004](#); [Park et al., 2011](#)), much of the work in this research line has focused on self-

disclosing behavior among reciprocal relationships and its effect on increasing tie strength at the dyadic level. It is not clear, however, how these findings about one-on-one self-disclosure in dyads generalizes to the types of self-disclosure frequently seen on SNSs, in which personal information is broadcast to a large and undifferentiated audience. For instance, according to [Collins and Miller \(1994\)](#), individuals disclose more to those they like and like those more to whom they disclose more. However, when people broadcast their personal information on SNSs, there is no single individual who is the disclosure target. Since broadcast self-disclosure is meant to be seen by many different people in one's social network, we are interested in the effect that broadcast self-disclosure has on expanding one's entire social network. This leads to our second research question:

RQ2: What is the relationship between broadcast self-disclosure and social network growth?

When considering the effects of self-disclosure on relational outcomes, most scholars measured feelings of connection. The use of SNSs and status updates can produce a feeling of connectedness among users ([Leimeister et al., 2010](#); [Grieve et al., 2013](#); [Burke & Kraut, 2014](#)). [Utz \(2015\)](#) further demonstrated that the typical connection between self-disclosure and the feeling of connection also exists on SNSs, that is, self-disclosure in status updates can increase people's feeling of connection. Researchers have also examined SNS content and its link with audience commitment and responsiveness as the relational outcomes. For example [Wang et al. \(2013\)](#) found that Facebook status update topics with varying degrees of self-disclosure elicited different amounts of comments and "like" clicks.

In contrast to the research reviewed above, the current study examines the relationship between self-disclosure and changes in network size over time. Understanding how broadcast self-disclosure relates to fluctuations in friend counts is as important as understanding feelings of connection or audience engagement. People's networks grow over time ([Dunbar, 1992](#)), but self-disclosure might affect, or be affected by, the rate of growth. For example, self-disclosure can foster liking and increase tie strength among dyads ([Cozby, 1972](#); [Collins & Miller, 1994](#)), and we assume disclosure in public has the similar effect. That is, broadcast self-disclosure on SNSs should increase the average liking and thus tie strength between the discloser and her / his audience, which would result in more people in the audience becoming friends with the discloser. On the other hand, a person going through a major life change, such as entering college, may be making many new friends at once, and be hesitant to share very personal stories among all of these new friends. So, it's not clear whether self-disclosure would increase network size, or if it would decrease as network size increases.

To answer our research questions, we automatically analyzed self-disclosure behavior in two de-identified, aggregate Facebook datasets. We collected samples from users who, based on their Facebook profiles, changed relationship status or entered college, and used automated techniques to estimate the degree of self-disclosure shown in their status updates around the time of these events. We measured the degree of self-disclosure in each status update by applying the machine learning model introduced in Chapter 2. To visualize how self-disclosure of these two kinds of users changes with time, we plotted average self-disclosure level against dates by aggregating status updates posted on the same day. We then built statistical models to test the differences in self-disclosure between users who ended a relationship and users who started a relationship. We also statistically compared self-disclosure of new college students with college sophomores during the beginning of school. In order to test the connection between self-disclosure and changes in network size, we took three snapshots of social network data for the new college students and explored its effect on building new ties.

4.2 Study 1: Self-Disclosure of Users Who Changed Relationship Status

4.2.1 Dataset

The first study examines self-disclosure patterns in a random de-identified sample of Facebook users who changed their relationship status from "in a relationship" to "single" or vice versa on their Facebook profile in July 2013. To ensure that users were not changing their relationship status at random, we only included those who had not changed their relationship status for at least 100 days prior. We limited the analysis to English-speaking users located in the United States who had at least three posts during the month prior to the relationship status change and at least three posts during the month after. For each person in the sample, we analyzed one month of de-identified status updates prior to the relationship status change, and one month after. There were 36,122 users in the breakup sample who posted approximately 1.5 million status updates, and 2,032 users in the new relationship sample, who posted approximately 150,000 status updates. All data was observational, de-identified, and analyzed in aggregate; no text was visible to researchers, and no one's Facebook experience was changed in connection with this study.

4.2.2 Visualizing self-disclosure around relationship changes

We applied the machine learning model of self-disclosure introduced in Chapter 2 to automatically measure the degree of self-disclosure in each status update posted between one month before and after the date of a relationship change. Figure 2 shows the average self-disclosure score per day, with dates centered at the date of the relationship change (Day 0). The y-axis shows the average daily self-disclosure level normalized by the total number of posts per day. Greater spikes in the line representing people

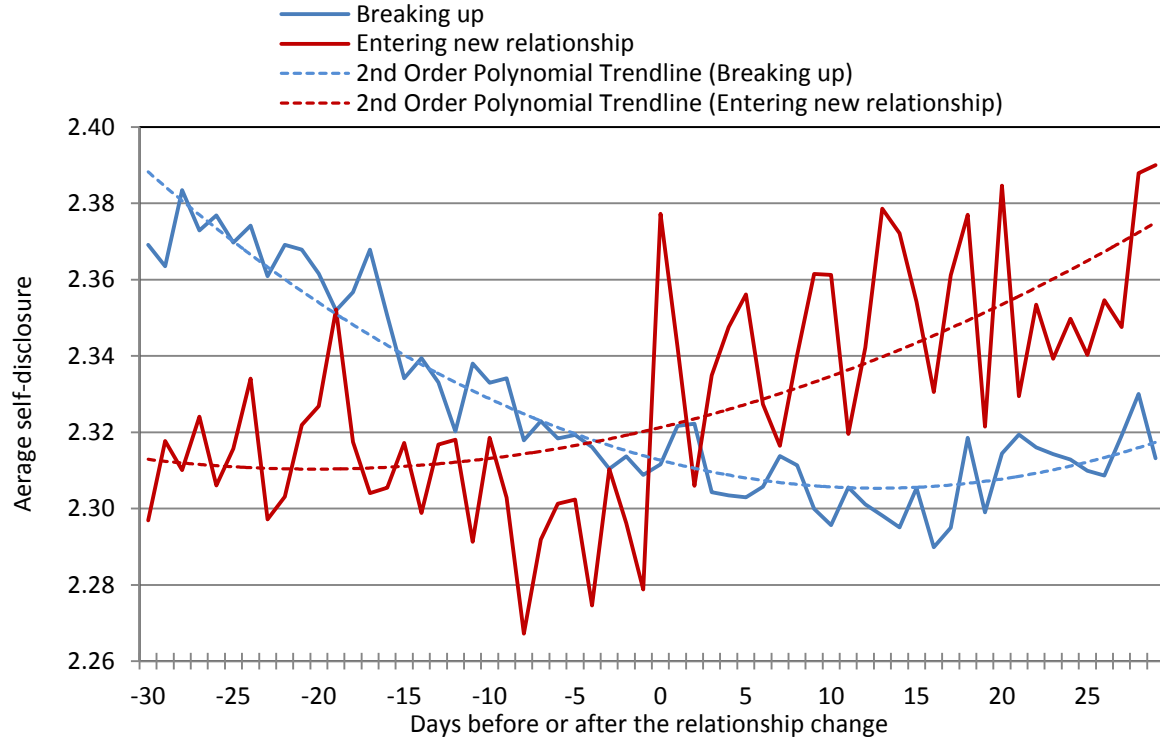


Figure 2. Average self-disclosure before and after relationship changes (Day 0).

starting a relationship (red) than in the line representing people ending a relationship (blue) reflect the discrepancy between the total persons for each group.

The two lines move in opposite directions: people entering a relationship started to self-disclose more than usual about one month prior, and continued to increase their self-disclosure during the month after they announced their new relationship. In contrast, people who were about to end a relationship began to self-disclose less about one month prior to the breakup, reached the lowest point 17 days after breakup, and then climbed up after that.

4.2.3 Statistical analysis and results

This section statistically tests the effect of changing relationship status on self-disclosure levels in status updates, moderated by posted dates. We built a random-effects linear regression model which grouped status updates at the user level (because the same person wrote multiple posts) to deal with non-independence of observations (Kennedy, 2003), with post self-disclosure as the dependent variable. Independent variables included the event type (0 for new relationship and 1 for breakup), and the day

offset (i.e., the day difference between posted date and the relationship status change date). Control variables included the poster's gender, age, tenure on Facebook, number of days in the past month the person logged in to the site, and Facebook friend count. Continuous variables were standardized, with a mean of zero and standard deviation of one; friend count was log-transformed before standardization because of its skewed distribution.

The results are shown in Table 9. The intercept in the model indicates that women who started a new relationship with *Age*, *Facebook tenure*, *Monthly login days*, and *Friend count* at their means on average disclosed at a level of 2.44 (on a 1 to 7 scale) on the date of their announced relationship change. Betas coefficients represent the effect on self-disclosure when a binary variable like *Event type* or *Male* or the integer variable *Day offset* increased by one, or a continuous control variable like *Age* or *Friend count*

Dependent variable: Self-disclosure			
Explanatory variable	Beta		Std. Err.
(Intercept)	2.439	***	.0081
Male	-.210	***	.0041
Age¹	.042	***	.0020
Facebook tenure¹	.007	***	.0021
Monthly login days¹	-.014	***	.0019
Friend count²	-.036	***	.0022
Event			
0 - New relationship			
1 - Breakup	-.019	*	.0082
Day offset	.001	***	.0001
Day offset * Event			
1 - Breakup	-.002	***	.0001
Number of observations	1,647,798 (38,891 individuals)		
R-squared	0.015		

1: Standardized and centered. 2: Logged (base 10), standardized, centered. *: p<0.05, **: p<0.01, ***: p<0.001

Table 9. Random-effects linear regression model at the user level predicting machine-coded self-disclosure of users who changed relationship status.

increased by one standard deviation. The R-squared value of 0.015 is small because the outcome we are predicting (self-disclosure in status updates) is relatively subtle and influenced by many unmeasured variables, like personality and other events. The model shows that males, frequent Facebook users, and users with more friends self-disclosed less; older users and users with longer Facebook tenure self-disclosed more. When holding all control variables constant, the negative beta of *Event* indicates that people who ended a relationship disclosed less than people who entered a relationship. The interaction between *Event* and *Day offset* is also significant: people who started a relationship disclosed more over time, while people who broke up gradually disclosed less over the following month.

4.3 Study 2: Self-Disclosure of New College Students

4.3.1 Dataset

To understand the relationship between entering college, self-disclosure, and network properties, we analyzed aggregate, de-identified status updates and friend count data from a random sample of college students first entering college in Fall 2014, as indicated in the Work or Education sections of their Facebook profiles. Non-English and non-US users were excluded from the sample, as was anyone who deleted their accounts. The resulting corpus contained samples from 228,608 users (*Freshmen* row, Table 10). In order to have a baseline of self-disclosing behavior and to control for the effects of calendar year, we included another sample of users in the dataset: students beginning their second year in college (*Sophomores*) in Fall 2014. Table 10 presents descriptive statistics for these students.

We investigated the links between student type, school start date, self-disclosure and social network size in three analyses. The first analysis shows that self-disclosure increased with the start of school; freshman displayed a stronger increase than sophomores. The second analysis shows that Facebook friend counts increased linearly with time and with the start of the school year, but more strongly for freshmen. The third analysis is a moderation analysis to test whether the increase in friend counts was strongest for those who self-disclosed the most. We measured self-disclosure in status updates for the five months prior to the start of school and two months after, covering approximately 16 million de-identified updates. All the

Type of users	Sample Size	Average updates	Average friends
Freshmen	228,608	70.53	540.44
Sophomores	441,367	55.33	561.68

Table 10. Descriptive statistics of the college student dataset.

analyses were completed automatically; no text was viewed by researchers. To understand the relationship between self-disclosure and changes in network size, we also took snapshots of their network size at three time points, including the beginning of the semester, one month and two months after the semester start date.

4.3.2 Visualizing self-disclosure among students

We again applied the self-disclosure machine learning model to the de-identified status updates in the sample. Figure 3 presents daily self-disclosure averages per student type. This figure reveals several patterns. First, the average self-disclosure levels of college freshmen and sophomores were similar. Second, there are two big peaks—May 11, 2014 (Mother’s Day) and June 15, 2014 (Father’s Day). These peaks may be an artifact of our machine learning methodology, where one important feature of self-disclosure is the inclusion of relationally-close others. Posts on these holidays are likely to contain phrases like “Happy mother’s day,” or “I love you, mom,” which may not be legitimate expressions of self-disclosure. The graph also exhibits a weekly recurring pattern of self-disclosure: students disclosed more about themselves during week days, possibly sharing what they did last weekend or their plan for next weekend. Finally, the average self-disclosure scores seem to be higher between August 15 and September 15, compared to the scores before or after that period, likely because semesters begin during that period. There is a small, positive correlation between the average self-disclosure level per post on a day and total post count that day ($r=0.26$), indicating that the number of posts people make can reflect self-disclosure to a certain degree.

Since students had different school start dates, in order to further confirm whether self-disclosure is related to the start of the school year, we plotted another graph that depicts average self-disclosure scores against relative dates using the school start date of every user as the reference point (Figure 4). From the graph it is clear that the self-disclosure levels of the two groups started to increase gradually about two months before school started, reached a local maximum on the first day of school, and then decreased slowly. The freshmen group had a bigger increase in self-disclosure than the sophomore group, which suggests that entering college can prompt people or provide them opportunities to reveal more about themselves.

Visually comparing the levels of self-disclosure of freshmen with sophomores suggests that entering college, an important life event with major implications for students’ identities, social networks and emotions, was associated with more broadcast self-disclosure than the routine starting of a new school year.

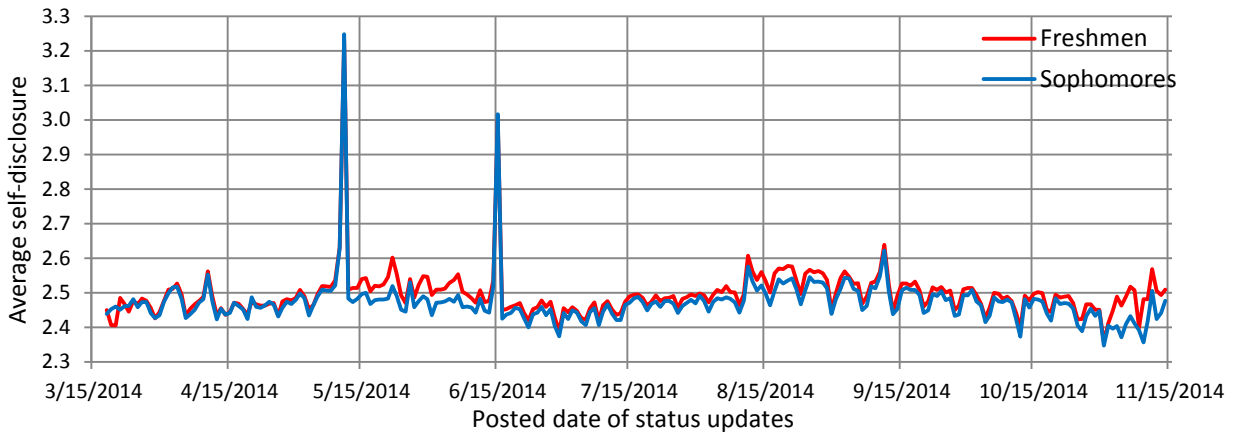


Figure 3. Average self-disclosure scores against absolute dates for the two kinds of students.

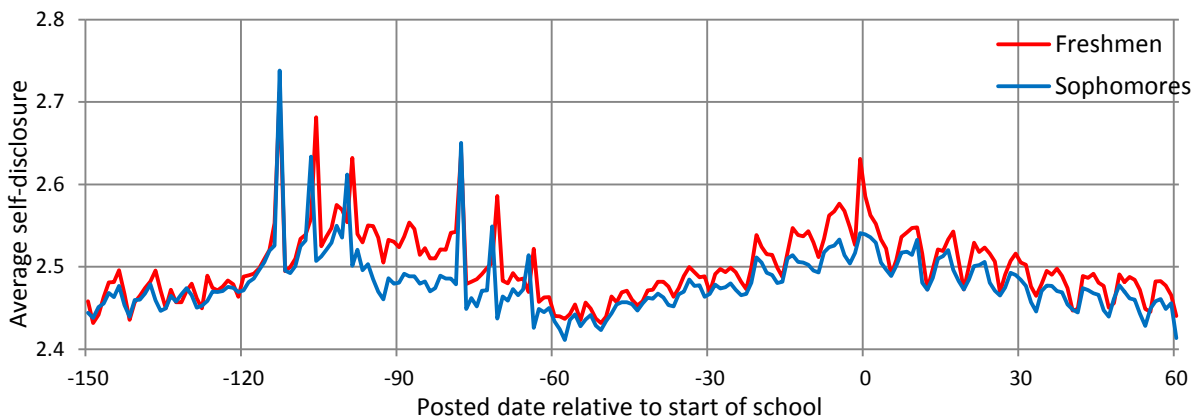


Figure 4. Average self-disclosure scores against relative dates for the two kinds of students using the school start date as the reference point.

4.3.3 Self-disclosure around the start of school

In this section, we statistically tested the link between entering college and self-disclosure. We examined whether the increase in degree of self-disclosure during the school-start period was larger for freshmen than sophomores. To do so, we examined average self-disclosure level by student types at three 30-day time periods: before-school-start (45 to 16 days prior to school), school-start (15 days before and after school start), and post-school-start (16-45 days after school start). We computed three self-disclosure scores for each user by averaging the machine-coded self-disclosure values of their status updates in each of the three time periods. We conducted a random-effects, hierarchical, linear regression analysis with

time period nested within student to account for the non-independence of self-disclosure (Kennedy, 2003). Average self-disclosure per period per user was the dependent variable. Students who did not post at least once per period were excluded.

To investigate whether self-disclosure increased linearly with time or peaked at the start of school, we included two time-related, orthogonal contrasts: *Linear time* (-1, 0, +1) and *Quadratic time* (-1, +2, -1). In particular, the interaction of *Student type* with *Quadratic time* allows us to determine whether the self-disclosure increase around the beginning of the semester was higher for freshmen than sophomores.

Table 11 presents the results of the regression analysis predicting self-disclosure. The intercept indicates that a female college freshman with all continuous variables at their means discloses at a level of 2.629 on a 1 to 7 scale at the start of the school year. As in the relationship change study, the model shows that

Dependent variable: Self-disclosure		
Explanatory variable	Beta	Std. Err.
(Intercept)	2.629 ***	.002
Male	-.252 ***	.002
Facebook tenure (days)¹	.014 ***	.001
Monthly login days¹	-.028 ***	.001
Student type		
0 - Freshmen		
1 - Sophomores	-.023 ***	.002
Linear time (-1, 0, 1)	.001	.001
Linear time * Student type	-.004 **	.001
Quadratic time (-1, 2, -1)	.019 ***	.001
Quadratic time * Student type	-.005 ***	.001
Number of observations	955,800 (318,600 students)	
R-square	0.03	

1: Standardized and centered.

*: p<0.05, **: p<0.01, ***: p<0.001

Table 11. Random-effects linear regression model at the user level predicting machine-coded self-disclosure.

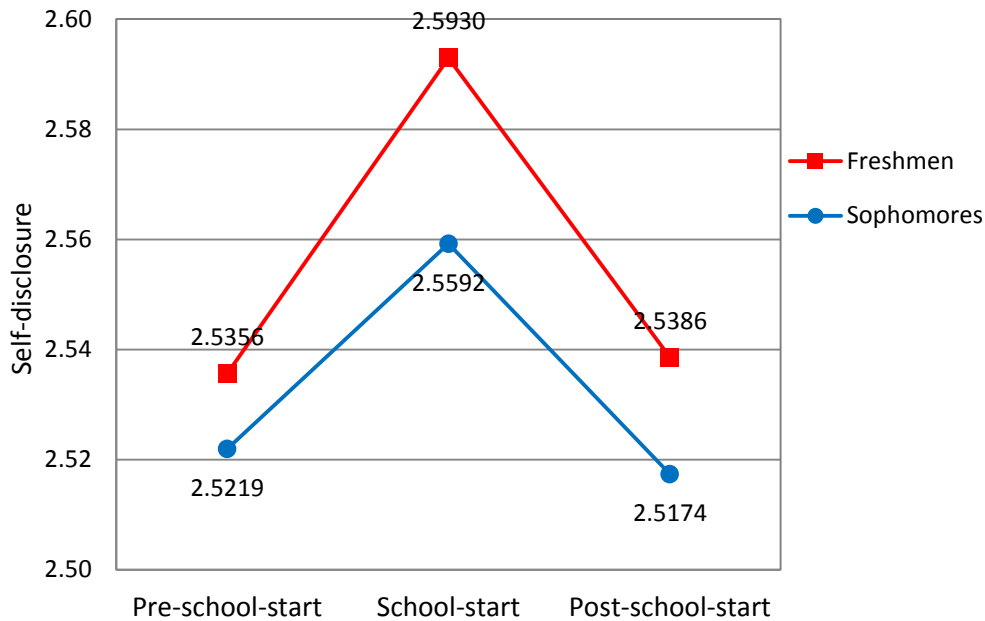


Figure 5. Predictive marginal means of self-disclosure of the two student groups during different time periods.

males revealed significantly less about themselves in their status updates than females, more active Facebook users disclosed less, and students with longer Facebook tenures self-disclosed more. When controlling for demographic information and activity level of these students, the analysis showed that freshmen significantly disclosed more than sophomores. While linear time was not associated with self-disclosure, the significant quadratic trend indicates that students disclosed more during the beginning of the semester compared to before and after the semester started.

Overall, the school-start period was associated with more self-disclosure than either the before- or post-school-start period. The statistically significant interaction of *Quadratic time* with *Student type* shows that the increase in self-disclosure at the start of the school years was stronger for the new college students than for college sophomores. Figure 5 presents these results, showing predictive marginal means of self-disclosure for the two kinds of students before, during and after the school-start period.

4.3.4 Network growth over time

To test the conjecture that new college students' networks grow more when school starts than college sophomores, we used a similar, hierarchical linear regress. The dependent variable is students' Facebook

friend count (logged base 10 after adding 1), measuring it at three time points: school start date, one month after, and two months after. Similar to the previous analysis, this study utilized two time-related predictor variables to test the linear and quadratic effects of time on friend count. The linear contrast tests whether friend count increased linearly with time, whereas the quadratic contrast tests whether the increase immediately after the school start date was larger than expected from the linear growth in friend count. Again, interactions with student type test whether these time trends differ for freshman compared to college sophomores.

Table 12 shows the results. The intercept indicates a typical female college freshman with average logins would have 375 friends ($10^{2.5749}-1$) one month after school started. Male college freshmen had 359 friends ($10^{(2.5749-.0189)}-1$) one month after school started. The significant positive coefficient on *Monthly login days* indicates more active Facebook users had more friends. The effect of *Student type* shows that sophomores had significantly more friends than freshmen.

Dependent variable: Friend count¹			
Explanatory variable	Beta		Std. Err.
(Intercept)	2.5749	***	.00099
Male	-.0189	***	.00105
Monthly login days²	.1299	***	.00061
Student type			
0 - Freshmen			
1 - Sophomores	.0373	***	.00109
Linear time (-1, 0, 1)	.0113	***	.00006
Linear time * Student type	-.0064	***	.00007
Quadratic time (-1, 2, -1)	.0011	***	.00003
Quadratic time * Student type	-.0008	***	.00004
Number of observations	1,944,537 (648,179 students)		
R-square	0.07		

1: Logged (base 10) 2: Standardized and centered.

*: p<0.05, **: p<0.01, ***: p<0.001

Table 12. Random-effects linear regression model at the user level predicting logarithms of friend counts with respect to base 10.

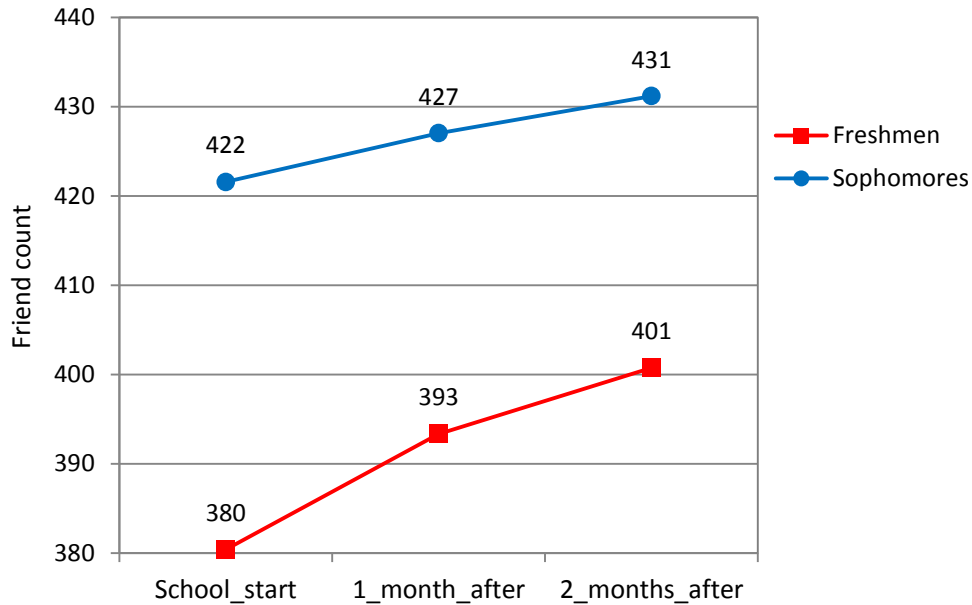


Figure 6. Predictive marginal means of friend counts for the two student types at different time points.

For easy interpretation, Figure 6 shows the predictive marginal means of the friend counts for freshmen and sophomores at the three time points. Along with a main effect of student type, there are effects of the two time-related factors. The positive slope of *Linear time* indicates that students in general had more friends on Facebook over time. The significant positive effect of *Quadratic time* indicates that the increase in friend count between the beginning of the semester and one month after was higher than the increase between one month and two months after the semester started. This plot further demonstrates the interactions between *Student type* and the two time-related variables. The different slopes of the two lines indicate that the change of *Friend count* depends on *Student type*. Although all students acquired more friends over time, freshmen acquired significantly more friends over time than did sophomores (see the *Linear time X Student type* interaction). In addition, freshman had a higher increase in friend count during the first month of the semester (see the *Quadratic time X Student type* interaction).

4.3.5 Self-disclosure and network growth

The previous two analyses showed that freshmen disclosed more and also increased their networks more than sophomores during the beginning of the semester. To test whether self-disclosure predicts changes in friend count, we ran a lagged dependent variable, random-effects, linear regression analysis with friend count as the dependent variable.

In contrast to the previous research that used cross-sectional data to examine the relationship between online self-disclosure and social relationships (e.g., [Barak & Gluck-Ofri, 2007](#); [Park et al., 2011](#); [Special & Li-Barber, 2012](#); [Utz, 2015](#)), our analysis uses longitudinal data to look at the temporal links between online self-disclosure and social relationships. Because the model has a lagged dependent variable (friend count last month) on the right side of the linear equation, it is equivalent to an analysis of change in friend count. In this analysis, friend count at time t is a linear combination of the friend count 30 days prior ($t-30$), the average self-disclosure score of the status updates posted during the intervening month, and the control variables. Although we cannot not make causal claims about the link between self-disclosure and friend count, this analysis fits well with observational studies. By using each student as his or her own control, it eliminates individual difference confounds; by using prior self-disclosure to predict current network size and controlling for prior network size, it controls for reverse causation. The variables in the analysis are:

- Friend count: The dependent variable was the number of Facebook friends a user had at time t , where t is the date one month or two months after school started. Every user had two data points for the two possible values of t .
- Self-disclosure: The independent variable was the average machine-coded self-disclosure score of the status updates posted by a user in the month prior to measuring friend count, between time ($t-30$) and ($t-1$).
- Time: Since there are two data points per user, we introduced a dummy variable to distinguish them: 0 when t is one month after school started, and 1 when t is two months after school started.
- Lagged friend count: The number of friends a user had at time ($t-30$). The correlation between *Friend count* and *Lagged friend count* before and after log-transformed is 0.9955 and 0.9933, respectively.

The analysis also included gender and monthly login days as controls. All continuous variables were standardized. Table 13 shows the descriptive statistics of the continuous variables in this analysis before log-transformation and standardization.

Model 1 in Table 14 is a baseline with *Lagged friend count* (friend count one month ago) and the other control variables. Model 2 adds *Self-disclosure*, *Student type*, and their interaction. The high correlation between *Friend count* and *Lagged friend count* ($r > 0.99$) means that friend count does not change much monthly, so the baseline model explains most of the variance due to the inclusion of *Lagged friend count*. The likelihood ratio test comparing Models 1 and 2 shows that adding *Self-disclosure*, *Student type*, and their interaction significantly improves model fit ($p < 0.0001$). Since the effects of the control variables on

the outcome variable are consistent in the two models, the following discussion focuses on Model 2, which contains both the control and independent variables.

Variable	Mean	Median	S.D.	Min	Max
Monthly login days	26.82	28	4.19	0	28
Lagged friend count	665.40	476	668.23	0	5206
Friend count	676.26	483	680.97	0	5196
Self-disclosure	2.53	2.38	.73	1	7

The mean, median and standard deviation of monthly login days, lagged friend count and friend count were multiplied by a random number close to 1 to protect Facebook's proprietary information but still provide readers a sense of the magnitude of each variable.

Table 13. Descriptive statistics for the numerical variables in the analysis of the relationship between self-disclosure and network growth.

Dependent variable: Friend count ¹	Model 1: Baseline		Model 2: Self-disclosure		
Explanatory variable	Beta	Std. Err.	Beta	Std. Err.	
(Intercept)	2.6683	***	.0001	2.6712	*** .0001
Male	.0017	***	.0001	.0015	*** .0001
Monthly login days³	-.0057	***	.0001	-.0056	*** .0001
Time	-.0033	***	.0001	-.0033	*** .0001
Friend count last month²	.3950	***	.0001	.3951	*** .0001
Self-disclosure				-.0006	*** .0001
Student type					
0 - Freshmen					
1 - Sophomores				-.0044	*** .0001
Self-disclosure * Student type				.0001	.0001
Log likelihood	1187482		1188231		
Number of observations	694,462 (347,231 students)				

1: Logged (base 10) 2: Logged (base 10), standardized, centered 3: standardized and centered.

*: p<0.05, **: p<0.01, ***: p<0.001

Table 14. Distributed lag linear regression models at the user level with random-effects predicting changes in friend count. Model 1 contains controls, and Model 2 adds self-disclosure, student type, and their interaction.

The intercept (2.6712) in Model 2 of Table 14 indicates that a female college freshman who logged in to Facebook the average amount, self-disclosed the average amount during the first month of school (2.48 on the 7-point scale) and had an average logged number of friends on the school starting date (451 friends = $10^{2.6555}-1$) would have 468 friends ($10^{2.6712}-1$) friends 30 days after school started. Women who had a standard deviation more friends in the previous month (654) had more friends in the current month (696). This is expected, since friend count this month highly correlates with friend count last month ($r>0.99$).

After controlling for this lagged dependent variable, all the other betas indicate changes in total friend count month-to-month that are associated with other predictors. First, considering the controls variables, male students had a larger increase in friend count during the beginning of school ($\beta=0.0015$, $p<0.001$). The negative coefficient for *Monthly login days* suggests that the friend count of highly active Facebook users increased more slowly than for average users ($\beta=-.0056$, $p<0.001$).

Next, we examine the independent variables. Students who self-disclosed a standard deviation more than average ended up making about one fewer new friend than those who disclosed at the average level ($\beta=-.0006$, $p<0.001$; 467.38 versus 468.03). This result is not consistent with the assumption that self-disclosure would be associated with an increase in friend count. Sophomores ($\beta=-.0044$, $p<0.001$) made fewer new friends than freshmen.

The interaction between *Self-disclosure* and *Student type* in Model 2 tests whether the association of self-disclosure with the change in friend count differs for freshman compared to sophomores. The association of self-disclosure and friend count did not differ between college freshman and sophomores ($\beta=0.0001$, $p=0.282$).

4.4 Discussion

This research examined the longitudinal patterns of broadcast self-disclosure, the relationship between events and self-disclosure, and possible consequences of self-disclosure on the size of one's social network in social networking sites. We measured the degree of self-disclosure in status updates by applying the machine learning model developed in Chapter 2. In Study 1, we showed that the changes in self-disclosure occur in opposite directions for users who started a relationship versus users who ended a relationship. While the self-disclosure level increases gradually during the beginning of a new relationship, it decreases dramatically during the break-up period. In Study 2, we showed that students, especially freshmen, talked more about themselves during the beginning of school.

Both the relationship and college student samples tested the hypothesis that important life changes would lead to an increase of self-disclosure. While the analysis results of the events of starting a relationship and entering college events demonstrated this, the breakup analysis showed the reverse. Among all motivations for self-disclosure, these results support the contention that self-presentation/social approval is the default motive. Reduced self-disclosure for the breakup sample suggests that self-presentation/social approval supersedes other motives because of the negative interpretation of events. A breakup is not simply a negative emotional experience—it implicates that others may view you as too socially inept to hold a relationship together. That is, a breakup damages one’s reputation and makes one look bad to others. Results from the breakup analysis also suggest that the self-disclosure motivation for distress relief was not the important primary motivating factor; otherwise, we would have expected more self-disclosure with negative than positive events.

The second part of Study 2 demonstrates that freshmen not only made the most new friends but also had the highest increase of self-disclosure among the three types of students during the beginning of school. However, it also shows a higher level of self-disclosure is associated with a lower increase in friend count. This finding contradicts our intuition that self-disclosure increases friend count. There are several possible explanations for this result. First, although prior literature indicates that self-disclosure increases tie strength and fosters friendship, it is based on evaluation between dyads ([Cozby, 1972](#); [Collins & Miller, 1994](#)). However, when Facebook users self-disclose in status updates on Facebook, we do not know exactly who sees the updates and thus who are affected. In fact, according to a Facebook news feed sorting algorithm ([Eulenstein & Scissors, 2015](#)), users’ updates are more likely to be seen by strong ties than weak ties, so users who self-disclose in updates may continue to increase their tie strength with strong ties rather than become Facebook friends with acquaintances and thus expanding their social network. The second explanation is regarding the appropriate level of self-disclosure. While self-disclosure is good for friendship, we know little about whether the effect is linear (i.e., the more the better) or is curvilinear, which suggests that there is an optimal level of self-disclosure for friendship (too much and too little self-disclosure are both bad for friendship). Third, this study cannot completely determine causality between self-disclosure and friend count since it is not a controlled experiment. The result can be explained by other pathways or even interpreted in the reverse direction. For example, freshmen usually meet many new people when going to college, and are likely to “friend” those people on Facebook, so the total number of Facebook friends increases. After having more weak ties in their friend list, they may self-disclose less, given that literature suggests individuals disclose more to the ones they are close to ([Collins & Miller, 1994](#)).

This research contributes to the literature on online self-disclosure and its effect on relationship formation. This is one of the few studies that examine longitudinal patterns of self-disclosure language in large samples of SNS posts ([see also Bak et al., 2014](#)). We explored the conditions under which individuals disclose more and events types associated with a higher degree of disclosure. While prior research is inconclusive and contradictory, we demonstrate that both positive daily and major life events trigger self-disclosure while negative events are linked to lower self-disclosure in the broadcasting style of online communication (i.e., Facebook status updates). Moreover, to the best of our knowledge, this is the first work studying the relationship between one-to-many self-disclosure and changes in online social network size. Although the result disconfirmed our assumption, we provided possible explanations.

Practically, the findings can be used to help SNSs designers provide better online environments and improve the quality of personalization. Since self-disclosure is beneficial for friendship maintenance, and our results suggest that users like to talk about themselves when experiencing positive life events, these sites can remind users about these events (e.g., a wedding anniversary), which may encourage users to share what happen during these events and thus increase the bond between users and their friends. In addition, the negative connection between broadcast self-disclosure and increase in network size suggests that SNSs should prompt users to share to a smaller group of friends when the content they reveal might be too much for new friends.

4.4.1 Limitations and Future Directions

A major limitation of this work is that we do not know what people talked about in their self-disclosures, especially whether they discussed the events that prompted them to self-disclose. For example, did people entering a relationship people discuss their partners? Did freshmen discuss the start of school? This requires more detailed coding or analysis of the texts of the status updates.

A second limitation is that while we are speculating that differences in self-disclosure are driven by different motives, we do not have direct evidence. One future direction would be to match the degree of self-disclosure with self-reported measures of motivations.

The findings in this work are based on two special samples of SNS users, students and people who self-reported the change of their relationship status in their profiles. Presumably there are many other people who also changed relationship status but did not update that change on their profiles. In order to ensure these findings generalize, subsequent research should inspect these results on a more heterogeneous sample or other kinds of events, such as change in job status.

This work is limited in self-disclosure broadcast through status updates. SNSs offer participants several types of communication channels which can be distinguished based on how directed and public the interaction is ([Burke et al., 2010](#); [Bazarova et al., 2012](#)). Whereas directedness measures whether the target of the communication is an identified friend, publicness measures the possibility that an individual's message might not have a well-defined audience and could be seen by those who are not the intended audience. Private messages, wall posts, and status updates on Facebook are good examples of the combination of directedness and publicness. Private messages are sent through a private channel and have a particular receiver. Status updates are published to a poster's entire social network and not targeted at any specific person. Wall posts are directed but can also be seen by any friend who are in the sender's or the recipient's network. The degree of directedness and publicness characterizes the distinctions of participation structure in online communication, which refers to the features of interaction and audience ([Herring, 2007](#)). Participation structure can affect language usage during social interactions ([Herring, 2007](#)). However, our current work only considers status updates. One interesting future direction is to study how self-disclosure differs in these channels and the impacts of self-disclosing in these different channels on social relationships.

Finally, as mentioned above, the studies are observational, hence the findings are correlational. Although the results are based on analyses of longitudinal data and we reduced the error inherent in the single wave design by constructing a distributed lag model to examine a three-wave data, we still cannot determine causality.

Chapter

5 Causal Relationships between Context Collapse and Self-Disclosure

5.1 Introduction

Today, the use of online social networking sites (SNSs) has become a majority social practice. For example, in June 2012, Twitter, a popular micro blogging and social networking site, had 500 million accounts worldwide ([SemioCast, 2012](#)); in June 2015, Facebook, the world's biggest social networking website, reached 1.49 billion monthly active users worldwide ([Facebook, 2015](#)). People communicate on SNSs to maintain friendships, form new social connections, seek support and entertain themselves. They can construct and manage social identities through these virtual places by editing their profiles and posting messages for others to see. SNSs not only provide a new platform for social interaction, but also a novel arena for *self-presentation*, a process through which people try to control the images others form about them ([Goffman, 1959](#)). According to [Goffman \(1959\)](#), individuals alter their self-presentation for different audiences. For example, people typically talk and behave differently at work than at home. However, SNSs also raise new questions for self-presentation, since users share information with people from many parts of their lives at once, a phenomenon known as *context collapse* ([boyd, 2008](#); [Marwick & boyd, 2010](#)). In that way, it is difficult to maintain various self-presentations in SNSs, where both colleagues and family members are present. In this study, we aim to answer the following research question:

Research Question: How does context collapse affect self-presentation on SNS?

We focus on how users present themselves through self-disclosures in status updates on Facebook. To understand the causal relationship between context collapse and self-disclosure, we used existing product tests in which the existence of multiple audiences was made more salient for users in the test group and examined how those changes affected self-disclosure. Product tests use A/B testing methodologies to compare two variations of a Facebook webpage design and examine performance effectiveness in terms of a specific outcome. The results show that emphasizing the existence of audiences did not change the

degree of sharing in Facebook status updates. Before discussing the details of the two studies, we first review relevant literature to derive possible assumptions.

5.1.1 Context collapse, self-presentation, and self-disclosure

In the *Presentation of Self in Everyday Life*, Goffman states, “When an individual appears in the presence of others, there will usually be some reason for him to mobilize his activity so that it will convey an impression to others which it is in his interests to convey” ([Goffman, 1959](#)). The process through which people try to control the images others form about them is known as *self-presentation*, or sometimes termed *impression management* ([Goffman, 1959](#); [Schlenker, 1980](#)). In Goffman’s dramaturgical metaphor, self-presenters alter their performances for different audiences in order to deliver disparate roles / impressions to those audiences. Self-presenters control which audiences see a particular performance. This concept is called *boundaries*, which segregate different audiences.

In SNSs, users’ networks are composed of different social clusters ([Parks, 2010](#)). Since social media technologies allow users to share personal information with multiple audiences at the same time, boundaries between audiences become less clear and more permeable. For instance, status updates on Facebook by default can be seen by all of an individual’s friends, regardless of friendship type or the intended audience. This characteristic of SNSs actually creates a new problem for online self-presentation: an individual’s multiple audiences collapse into one single context ([boyd, 2008](#); [Marwick & boyd, 2010](#)). Due to this *context collapse*, it is more difficult for users to manage separate impressions to different audiences in SNSs than in offline settings ([Marwick & boyd, 2010](#)). Researchers have shown that context collapse can potentially cause problems. For example, [Binder et al. \(2009\)](#) found that context collapse increases social tension within in online social networks because people must meet the expectations and interests of their friends, but friends may have different content preferences.

Here we are interested in whether and how context collapse influences users’ self-presentation on SNSs. In particular, we consider self-presentation made via self-disclosure, since self-disclosure is known to be essential for satisfaction, relationship maintenance and formation, intimacy, and thus accumulation of social capital ([Altman & Taylor, 1973](#); [Collins & Miller, 1994](#); [Oswald et al., 2004](#); [Park et al., 2011](#); [Special & Li-Barber, 2012](#)). Furthermore, we focus on the most common type of communication channels on SNSs, *broadcast communication* (e.g., Facebook status updates and Twitter tweets), since balancing self-presentation is especially challenging in this type of channel.

Numerous studies have examined self-presentation made through disclosure in SNSs (e.g., [Krasnova et al., 2010](#); [Nosko et al., 2010](#); [Park et al., 2011](#); [Stutzman et al., 2011](#); [Special & Li-Barber, 2012](#)). [Krasnova et al. \(2010\)](#) and [Stutzman et al. \(2011\)](#) found that users with more privacy concerns disclosed

less in their Facebook profiles. However, self-disclosure on Facebook is positively associated with intimacy among Facebook friends ([Park et al., 2011](#)) and positively correlates to user satisfaction ([Special & Li-Barber, 2012](#)). Nevertheless, few studies consider the influence of context collapse on self-disclosure in SNSs (Chapter 3; [Marwick & boyd, 2010](#); [Vitak, 2012](#)) and the results reported in these studies are contradictory. Specifically, [Vitak \(2012\)](#) analyzed the survey data collected from U.S. graduate students who were Facebook users and found that audience size and diversity positively correlate with the amount of self-disclosure. In contrast, in Chapter 3 we analyzed de-identified, aggregate network information and status updates and found that network size negatively correlates with self-disclosure, whereas average tie strength and network density positively correlate. Even if we disregard the contradiction, these findings are based on cross-sectional studies of the relationship between audience network structure and self-disclosure. As such, we can only make correlational claims, not causal ones. To resolve the inconsistent findings and overcome the limitation of cross-sectional studies, this study makes observations based on existing product tests, which make participants' potential audience more or less salient by reminding them their social network size or privacy settings in their status update composers. We then examined whether and how these reminders affect users' disclosure.

There are three possible actions SNS users can take to manage context collapse. These actions are supported by different literatures, which lead to three competing assumptions. First, to understand what strategies SNS participants employ to navigate multiple audiences, ([Marwick & boyd, 2010](#)) interviewed Twitter users and reported that instead of trying to engage different audiences at the same time, Twitter users post a mix of tweets targeted at different audiences to maintain their general popularity. They claimed that although users understand that there is no boundary segregating different audiences on SNSs, they still treat their audiences as if they were bounded. This suggests that users do not care whether a channel is public or private, yet would use it as if it were private. In other words, awareness of context collapse won't change what users are planning to say, that is, context collapse has no effect on self-disclosure.

On the other hand, [Hogan \(2010\)](#) proposed a theory of lowest common denominator culture, which suggests that when people face multiple audiences, they post content that is acceptable for everyone. In this case, context collapse may cause people to self-disclose less, because they would feel uncomfortable sharing intimate information appropriate for family and friends with relative strangers in their networks. That is, they might self-censor and only present information appropriate to the lowest common denominator. Several recent studies present preliminary findings supporting this theory. The interview analysis done by [Brandtzæg et al. \(2010\)](#) shows that people are aware that SNSs are public or semipublic spaces, and thus they self-disclose less due to privacy concerns. Chapter 3 shows that network size is

negatively correlated with self-disclosure, and average tie strength and network density are positively correlated. Hence, the other possibility is that context collapse has negative effects on self-disclosure.

The third possible action that SNS users can take to manage context collapse is that awareness of audience would make users disclose more. While this seems counter-intuitive, [Vitak \(2012\)](#) reported a positive correlation of audience size and diversity with the amount of self-disclosure by analyzing the survey data collected from U.S. graduate students who were Facebook users. Though, she argued that this result might be because users had utilized Facebook's Friend List feature to build boundaries in their network and publish posts to their sub-networks. Since her study is based on a sample of graduate students and limited in its ability to support casual claims, we retest this counter-intuitive finding (i.e., context collapse has positive effects on self-disclosure.) in our studies.

To answer the research question we examined differences in the amount of machine-coded disclosure in the treatment and control groups for people in two product tests designed to make audience more salient. The first experiment presented the users with the number of people who could see their status updates in each privacy category in the status update composer. The second experiment presented the users with a privacy checkup dialogue, which allowed them to review the privacy setting of their status update composer. These experiments proceed from the assumption that users who see the audience counter or privacy checkup reminder would be more aware of the existence of multiple audiences (i.e., context collapse) than those who do not see it. It is important to note that the two experiments were pre-existing experiments designed to improve user experience and usability on Facebook. We analyzed participants' self-disclosure in status updates as a side effect of the usability studies. In other words, the analyses in this chapter were based on existing data from product tests that Facebook was already running. In the next two sections, we first describe the design and results of the audience counter study and then the privacy checkup study.

5.2 Audience Counter Study

5.2.1 Experimental design

In the audience counter study, participants were shown the number of people in several relationship categories, such as friends or close friends, when composing status updates. Specifically, there was a number presented immediately after each option in the "Who should see this?" dropdown menu in the status update composer. The number represents the number of audience members in each relationship category and also the number of people who could see the status update if the option is selected. Figure 7 presents a mockup interface of the audience counter. The goal of this study was to increase audience salience, and thus encourage people to reflect on their privacy controls when sharing information. As a

corollary, we measured whether making audience more salient influenced levels of self-disclosure. The experiment was deployed to 1.5 million English speakers. A control group of 1.5 million users saw the standard composer privacy selector without the audience counter. The experiment was launched for about 8 months spanning across 2013 and 2014. All status updates posted by the treatment and control groups were de-identified and analyzed autonomously. The rest of the analyses were based on a comparison of collected status updates between the control and experimental groups. These analyses were conducted on Facebook's company servers and fell within the site's Terms of Service.

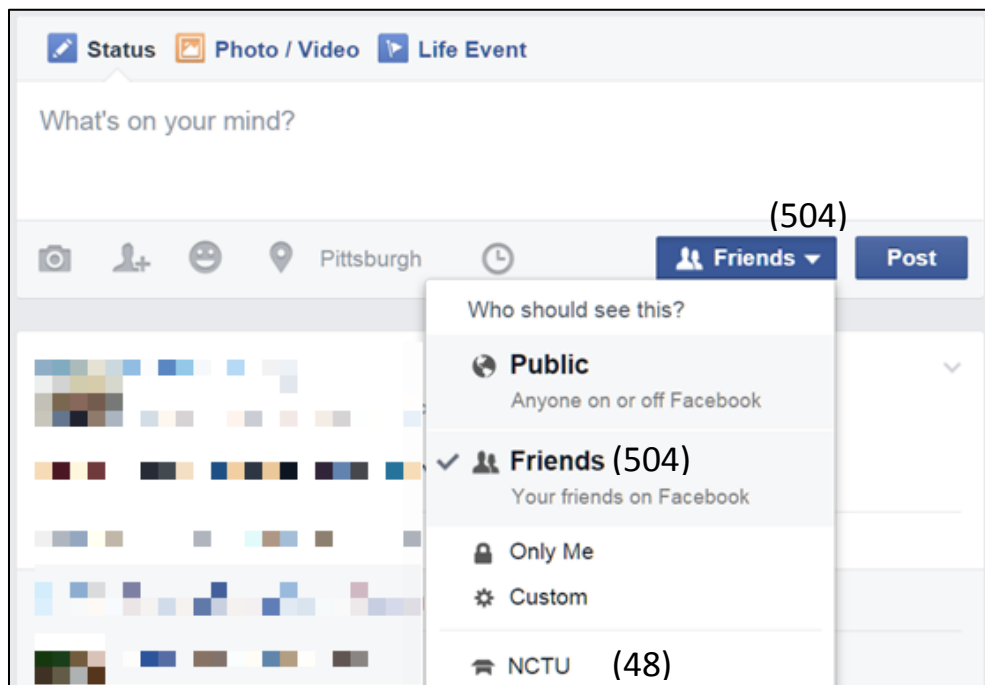


Figure 7. Mock-up interface of the audience counter study.

5.2.2 Analysis and results

To test whether there was a statistical difference in self-disclosure between users in the control group and experimental group, we conducted a linear regression analysis, where the dependent variable was the average self-disclosure score per user, and the independent variable was a binary one indicating in which group a user belonged to (0 for the control group; 1 for the experimental group). The analysis was based on users exposed to the status update composer between 25 December 2013 and 22 February 2014. To ensure that we could gauge self-disclosure level, we restricted the analysis to users who had at least

posted one status update between 11 January 2014 and 22 March 2014. In total, 1,909,808 users posted at least once during that period; 953,648 users were in the experimental group (49.93%). We applied the automatic self-disclosure model introduced in Chapter 2 to measure the disclosure level of each of the 30,733,677 updates written by these users, and computed a self-disclosure score for each user by averaging the machine-coded self-disclosure values of their status updates. All data was de-identified and analyzed in aggregate on Facebook's servers; no text was viewed by researchers.

Because the users were randomly assigned to one of the two conditions, they should have similar characteristics except for the conditions to which they were assigned. Thus, it was not necessary to have control variables in the regression model. However, we still collected their demographic information as control variables, including gender, age, number of days since the users confirmed their Facebook account, number of days they logged into Facebook in the previous four weeks, and number of friends and followers. Except for gender, which was a binary variable (0 for female; 1 for male), the control variables were continuous and standardized with a mean of zero and standard deviation of one. Friend count and follower count were log-transformed before standardization, because their distributions had long tails. The descriptive statistics of these variables before standardization are presented in Table 15.

Table 16 shows the regression result. The intercept of 2.32 is the self-disclosure level of a woman with all continuous variables at their means. Betas are the effect on self-disclosure from a binary variable having a value of 1, or a one standard deviation increase in continuous independent variables. The R-squared value is 0.0278. The value is small because predicting self-disclosure in people's language itself is a difficult task: language use is a subtle characteristic, and many things in a person's life can influence language presentation, such as family, culture, and education level. The result indicates that users who were older or had been a confirmed Facebook member longer significantly disclosed more about themselves. On the other hand, users who were male, more active, or had more friends or followers self-disclosed less. We did not find any significant effect of the audience counter; there is no difference in self-disclosure levels between users who had the counter and users who did not. To examine whether the effect of the counter depends on friend count or follower count, we also considered the interaction between *Has counter* and *Friend count* or *Follower count*, and again found no effect. The findings support the assumption that context collapse has no effect on self-disclosure.

Variable	Mean	Median	Std. Dev.	Min	Max
Age	32.38	28	15.37	13	113
Friend count	435.33	301	498.76	0	5087
Follower count	11.26	0	226.74	0	105530
Self-disclosure	2.23	2.06	0.73	1	7

Table 15. Descriptive statistics for the variables in the regression analysis of the audience counter study.

Dependent variable: Self-disclosure		
Explanatory variable	Beta	Std. Err.
(Intercept)	2.3202 ***	.0009
Male	-.1820 ***	.0011
Age¹	.0627 ***	.0005
Days since registered¹	.0289 ***	.0006
Number of logins¹	-.0430 ***	.0006
Friend count²	-.0030 ***	.0008
Follower count²	-.0042 ***	.0008
Has counter	-.0005	.0010
Has counter X Friend count	.0012	.0011
Has counter X Follower count	-.0007	.0011
Number of observations	1,909,808	
R-square	0.0278	

1: Standardized and centered. 2: Logged (base 10), standardized, centered

*: p<0.05, **: p<0.01, ***: p<0.001

Table 16. Linear regression result of the audience counter study.

5.3 Privacy Checkup Study

5.3.1 Experimental design

The privacy checkup study was designed to encourage users to undergo a privacy checkup and become aware of the audiences that could see the information they share on Facebook. Users in the experiment were shown an introduction dialogue of the privacy checkup when logging into Facebook (Figure 8). The dialogue asked them whether they would be interested in checking out their privacy settings. If users accepted, a second dialogue would show up in which they could go through several steps to inspect the privacy settings for their posts, apps, and profile (Figure 9).

To examine the effectiveness and usability of the feature, a control group was used for comparison. Users were randomly assigned to one of two groups. The users in the control group did not receive the popup dialogue when logging into Facebook. This privacy checkup feature was gradually launched to Facebook users worldwide starting from mid-2014 and known to the public as Facebook's privacy dinosaur because of the blue dinosaur cartoons on the interface elements ([Albergotti, 2014](#)).

5.3.2 Analysis and results

Similar to the analysis for the audience counter study, we built a linear regression of the treatment on self-disclosure (i.e., seeing the introduction dialogue of the privacy checkup). The analysis was based on a random sample of English speakers who logged into Facebook between 29 September 2014 and 10 October 2014. During this period, 206,173 users in the experimental condition logged into Facebook and saw the dialogue; a comparable number of users in the control condition logged into Facebook but did not see the dialogue. Status updates were collected when posted in the four weeks following their exposure date (i.e., the first login date during the timeframe listed above). For instance, for users who were exposed to the dialog on October 5, their status updates between October 5 and November 1 were de-identified and analyzed in aggregate. In total, 4,345,701 updates were analyzed. User information and status updates were de-identified and analyzed autonomously. In this model, we also included the same control variables that had been used in the audience counter analysis. Table 17 shows the descriptive statistics of the variables in the regression. The regression result is reported in Table 18 and is very similar to the result of the audience counter study. In summary, users' self-disclosure level in status updates was not affected by seeing the introductory dialogue of the privacy checkup (Figure 8).

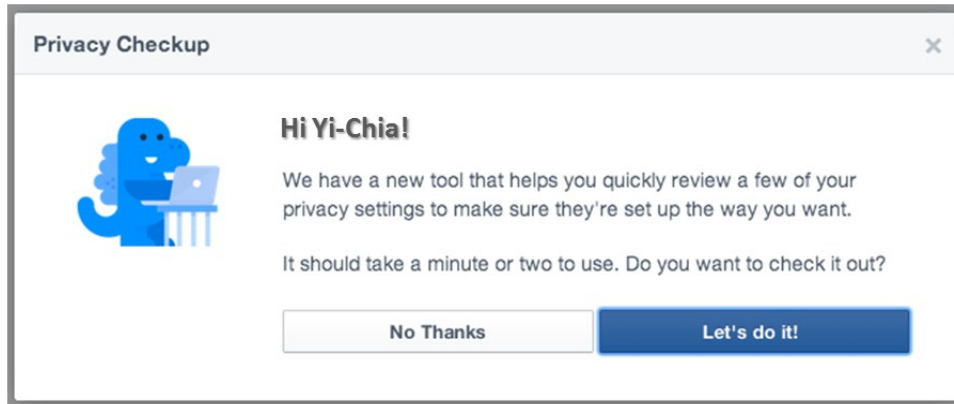


Figure 8. Screenshot of the introduction dialogue of the privacy checkup study.

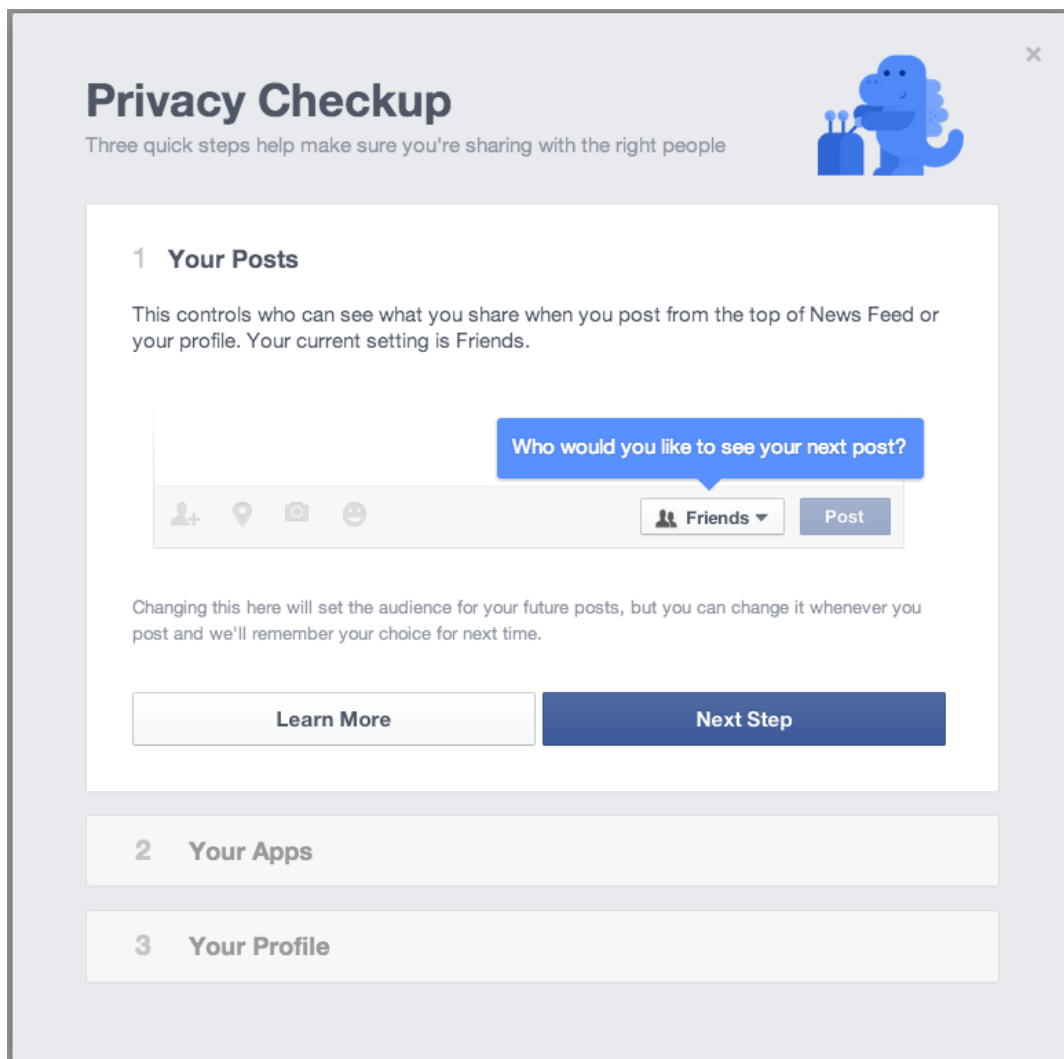


Figure 9. Screenshot of the main dialogue of the privacy checkup study.

Variable	Mean	Median	Std. Dev.	Min	Max
Age	35.65	32	14.24	14	114
Friend count	492.76	329	558.19	0	4,968
Follower count	15.07	0	587.92	0	219913
Self-disclosure	2.50	2.31	.83	1	7

Table 17. Descriptive statistics for the variables in the regression analysis of the privacy checkup study.

Dependent variable: Self-disclosure			
Explanatory variable	Beta		Std. Err.
(Intercept)	2.5952	***	.0020
Male	-.2748	***	.0027
Age¹	.0973	***	.0013
Days since registered¹	-.0080	***	.0013
Number of logins¹	-.0414	***	.0013
Friend count²	-.0048	*	.0019
Follower count²	-.0193	***	.0018
Privacy checkup	-.0002		.0025
Privacy checkup X Friend count	.0025		.0026
Privacy checkup X Follower count	-.0009		.0026
Number of observations			412,346
R-square			0.0436

1: Standardized and centered. 2: Logged (base 10), standardized, centered
 *: p<0.05, **: p<0.01, ***: p<0.001

Table 18. Linear regression result of the privacy checkup study.

5.3.3 Propensity score matching

5.3.3.1 Motivation

One possible explanation for why the introduction dialogue (Figure 8) did not stimulate individuals' sense of audience as we expected, and thus influence disclosure, is that the description does not provide any information about audiences. The figure only states: "We have a new tool that helps you quickly review a few of your privacy settings to make sure they're set up the way you want. It should only take a minute or two to use. Do you want to check it out?". However, it does not contain detailed information about users' privacy settings or audience. Users need to accept the figure before they could walk through their privacy settings in Figure 9, which might be a stronger stimulator of the existence of multiple audiences than Figure 8. Therefore, instead of comparing disclosure of users who did not see Figure 8 with those who did, it might be more appropriate to compare users who saw but rejected Figure 8 with those who accepted it and thus saw Figure 9. However, while the former comparison was conducted based on the data of a true controlled experiment where users were randomly assigned to conditions, we were not able to do the same analysis for the latter since users who saw Figure 8 self-selected into the treatment (i.e., Figure 9). Similar to most events in the world, the treatment was not randomly assigned but was endogenous in the sense that other factors in the system caused it. In our case, some people might be more likely to accept Figure 8. For instance, busy users may just leave the introduction dialogue without reading it. Users who have a longer Facebook membership or trust the service provider more might be more willing to accept the introduction dialogue and continue. Hence, it is important to control confounding factors that affect both the treatment and the outcome, so that they do not bias the estimation of treatment effects.

To address the issue of endogeneity and reduce selection bias, we applied *propensity score matching* (PSM) to our data ([Rosenbaum & Rubin, 1983](#)). PSM is a quasi-experimental method used to make observational studies to approximate randomized ones for causal inference. It matches each treated user with an untreated user by balancing them on baseline confounding variables, and then estimates the average treatment effect on outcomes by comparing the two matched groups. Although PSM can reduce the bias caused by confounders in non-randomized studies, it relies on the researchers to identify a good set of confounding variables, and it is not always possible to control for all variables related to treatment.

In summary, we used propensity score matching methods to examine the effect of accepting the introduction dialogue (i.e., a proxy of audience awareness) on self-disclosure level. Since users' basic demographic network information could be important confounders, PSM was designed to balance treatment and control groups in order to address the problem of endogeneity. Analysis was restricted to

users who saw the introduction dialogue (i.e., users assigned to the experimental group in the original analysis of the privacy checkup study). The data consisted of approximately 200,000 users, a proportion of whom accepted the dialogue. We describe the details of the process and results below.

5.3.3.2 Analysis and results of PSM

Propensity score matching consists of four steps. In the first step, we built a logistic regression to calculate a propensity score (i.e., the probability of accepting the introduction dialogue, Figure 8) for every user in the sample, conditioned on a set of covariates (basic demographic and social network information) that might influence acceptance. In particular, demographic information included gender, age, days since users joined Facebook, and number of days users had logged into Facebook in the past four weeks; social network information included the number of friends and followers, average tie strength between users and their friends, and number of friendship connections among users' friends. The idea behind this was that these factors might correlate with the acceptance of the dialogue as well as changes in self-disclosure. Table 19 presents the logistic regression that estimates the probability of accepting the

Accept introduction dialogue (Figure 8)	Coef.	Std. Err.
(Intercept)	-1.7214 ***	.0077
Male	.0944 ***	.0130
Age¹	.1140 ***	.0062
Days since registered¹	.2965 ***	.0069
Number of logins¹	.0356 ***	.0070
Friend count²	-.1086 ***	.0082
Follower count²	.0842 ***	.0061
Network density¹	-.1091 ***	.0075
Average tie strength¹	.1541 ***	.0068
Number of observations	206116	
Log likelihood	-89527.067	
Pseudo R-square	0.0234	

1: Standardized and centered. 2: Logged (base 10), standardized, centered

*: p<0.05, **: p<0.01, ***: p<0.001

Table 19. Estimate the propensity score (the probability of accepting the introduction dialogue of the privacy checkup) using logistic regression.

introduction dialogue of the privacy checkup (propensity score) using users' demographic and network information as predictors. For example, the result shows that active Facebook users were more likely to accept the dialogue, while users with more friends were not.

In the second step, users who accepted the introduction dialogue (the treatment group) were paired with users who did not (the control group) but had similar propensity scores based on the eight demographic and network predictors. Using propensity scores to match users, we were able to control for multiple factors affecting treatment at the same time with just one variable. We selected k-nearest-neighbor matching as the matching algorithm, one of the most common matching algorithms for PSM. The algorithm was set to find one control neighbor for each treated user without replacement. There were 33,618 pairs of treated and untreated users after matching. Figure 10 shows two density graphs of the propensity score (i.e., the likelihood of accepting the introduction dialogue) for the treated and untreated groups, before and after matching. The overlap of the two distribution lines after matching suggests that the constructed treatment group and control group are balanced on the likelihood of accepting the introduction dialogue, i.e., the distribution of the propensity score is very similar across the two groups.

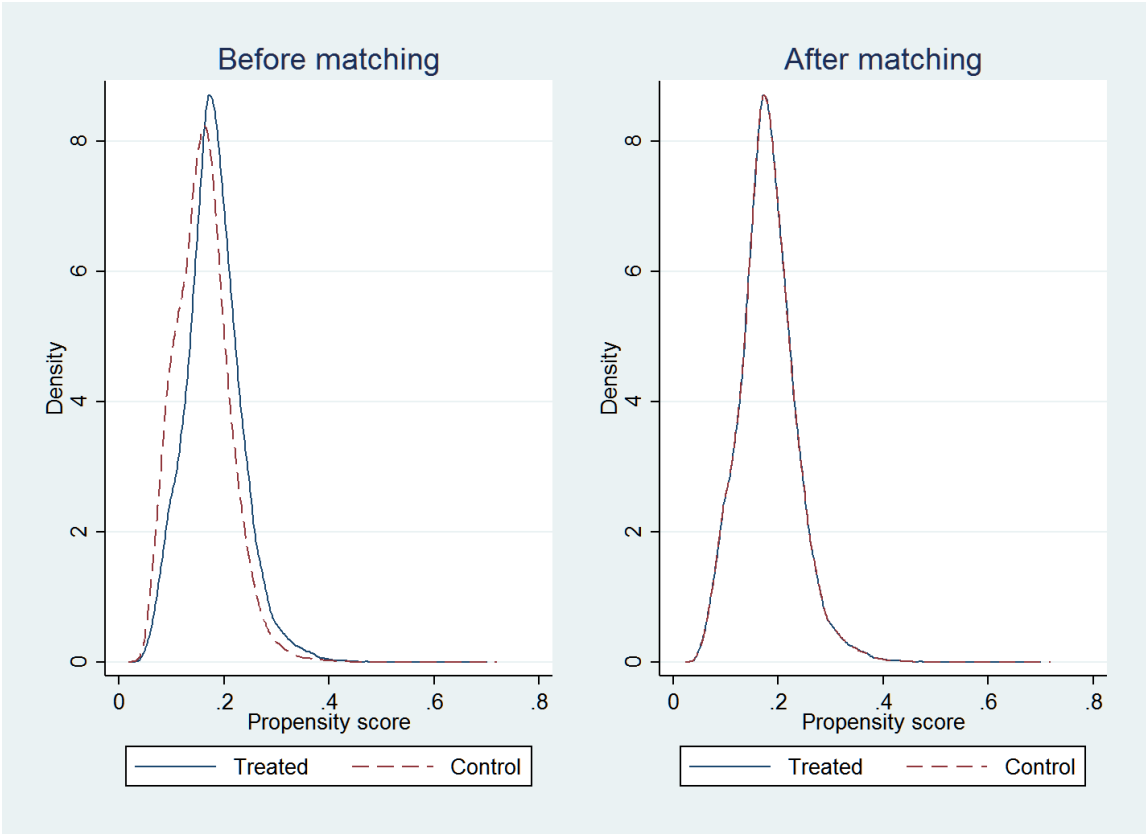


Figure 10. Density graphs of propensity score for the treated group (i.e., accept the intro dialogue) and control group (i.e., not accept the intro dialogue) before matching (left) and after matching (right).

The third step checks whether the balancing property is satisfied in the matched samples, that is, if the constructed treatment and control groups have similar characteristics in terms of the covariates used for the propensity score calculation. The distributions of individual covariates across the treatment and comparison groups after matching should not differ significantly. One common balance diagnostic is to use t-tests to examine whether the means of these covariates in the treated and untreated groups are different from each other. Good balancing is achieved if the means are not statistically significant after matching. Table 20 shows the results of t-tests for the eight covariates in our propensity score model. We see that the difference in covariate means between the treatment and control groups became insignificant after matching, which suggests good balancing was achieved. Another balance test checks standardized bias. Good balancing is achieved if percent bias is less than 5% after matching. Table 20 indicates that percent bias for the eight covariates were below 5% after matching. The biases were reduced over 90% for six of the eight covariates, which again confirms that the constructed treatment and control groups were well-balanced.

Variable	Sample	Mean		Bias		t-test	
		Treated	Control	%bias	%red bias	t	p> t
Male	Full	.3348	.3372	-0.5		-0.86	0.390
	Matched	.3348	.3316	0.7	-32.6	0.88	0.377
Age	Full	.1499	-.0292	18.0		30.11	0.000
	Matched	.1499	.1520	-0.2	98.8	-0.26	0.795
Days since registered	Full	.2096	-.0406	25.3		42.15	0.000
	Matched	.2096	.2190	-0.9	96.3	-1.26	0.209
Number of logins	Full	.0710	-.0131	8.9		14.13	0.000
	Matched	.0710	.0658	0.5	93.8	0.77	0.440
Friend count	Full	-.0803	.0166	-9.9		-16.31	0.000
	Matched	-.0803	-.0876	0.7	92.5	0.97	0.332
Follower count	Full	.0418	-.0083	5.0		8.40	0.000
	Matched	.0418	.0315	1.0	79.5	1.30	0.194
Network density	Full	-.1207	.0236	-14.9		-24.25	0.000
	Matched	-.1207	-.1143	-0.7	95.6	-0.91	0.365
Average tie strength	Full	.1565	-.0301	18.0		31.58	0.000
	Matched	.1565	.1453	1.1	94.0	1.29	0.197

Table 20. Covariate comparison between users who accepted the intro dialogue (treated) and those who rejected it (control) before and after propensity score matching.

Once propensity scores were computed for users in the data, users were paired based on their propensity scores, and the balancing property was checked and good for all covariates, the last step was to estimate the average effect of treatment (i.e., accepting the introduction dialogue) on self-disclosure in the matched samples. The average difference in self-disclosure between the treated users and matched controls was -0.026 ($p < 0.0001$), indicating that the treatment had a negative impact on self-disclosure. That is, users who accepted the introduction dialogue (Figure 8) and thus continued to the main dialogue of the privacy checkup (Figure 9) did disclose significantly less afterward than users who dismissed the introduction dialogue. This finding suggests that increasing audience salience results in lower self-disclosure.

5.4 Discussion

In recent years, there has been an increase of interest in understanding the effect of context collapse on individuals' self-presentation on social network sites. This study examined how people presented themselves through the disclosures they made in their Facebook status updates, a public online communication channel consisting of multiple audiences. We presented two controlled experiments in which the existence of multiple audiences on SNSs were made salient via adding an audience counter in the status update composer or showing users a privacy checkup tool which allowed them to review the privacy settings of their status update composer. We found that both the audience counter and privacy checkup dialogue had no influence on changing how much users self-disclosed in status updates.

One explanation for this counterintuitive result is that humans adapt to technology. When SNSs were first introduced, they were a new concept. There was no standard method of use. Users simply used whatever features provided by the sites based on their intuitions without worry. As suggested by the qualitative interviews with Twitter users done by [Marwick and boyd \(2010\)](#), although users know that the main communication channel on SNSs is public and there is no boundary separating different audiences, they still use the public channel as if it were private.

Another explanation is that although users were randomly assigned to conditions in the two experiments, they were designed to target a pool of active and long-term Facebook users in the first place. These users might be systematically different from those who do not frequently use or who recently joined SNSs. For instance, active and long-term users have more experience using SNSs and thus have more trust in the service providers. Or, they may be the type of people that are more open to new things. As a result, being aware of context collapse does not change their disclosure.

Furthermore, it is also possible that the treatments in the two experiments were not strong enough to stimulate users' sense of multiple audiences, especially for the privacy checkup study, as explained in the

last section. Hence, we conducted a propensity score matching analysis on a subset of the privacy checkup dataset and demonstrated that users who accepted the privacy checkup significantly self-disclosed less than those who rejected it. Different from the results of the two experiments, this result supports the assumption that audience salience negatively correlates with self-disclosure and also suggests that the treatments in the two experiments might be too subtle for posters to be conscious of their audiences. Nevertheless, even though propensity score matching can ameliorate the endogeneity problem and approximate randomization, this finding is correlational but not causal. There might be some other important confounding factors that we did not account for in the model. Further investigations are needed.

In sum, we believe this work expands our understanding of the link between context collapse and self-disclosure. The results of the two experiments and the propensity score matching analysis together indicate that the issue of context collapse may not be as serious as we thought. We found it either had no effect or a tiny negative effect (-0.026 with respect to a standard deviation of 0.73) on self-disclosure. This finding is stronger than those reported in the prior work, because it was drawn from two very large controlled experiments involving more than 2 million people. Since we suspect that the relationship between audience salience and self-disclosure might not be linear, future experiments could vary the degree of audience salience and examine self-disclosure as the outcome.

Chapter

6 Conclusion and Future Directions

When people interact with others, they disclose information about themselves. Through the disclosing process, they know each other better and develop a closer friendship. Recently, social networking sites (SNSs) have begun to offer people a platform for social interactions and self-disclosure. However, self-disclosure in SNSs may be different from self-disclosure in offline settings, since SNSs allow users to broadcast to a large social network, rather than disclose individually. Many researchers have questioned the influence of SNSs on broadcast self-disclosure and its accompanying social consequences. This dissertation presents a detailed answer that not only extends the existing literature in both social sciences and linguistics but also provides valuable insight into social networking site design. By building an automated model of self-disclosure to analyze millions of de-identified SNS posts and integrating the data with demographic and social network information, this work examines factors that might influence people's broadcast self-disclosure at three levels, including their personal characteristics, the structure of their online social networks, and events happening in their lives. It also studies the relationship between broadcast self-disclosure and social network expansion. In this chapter, we first summarize the findings across the chapters and then discuss limitations of this research as well as suggest next steps for future work.

6.1 Summary of Findings and Contributions

Model of self-disclosure: The machine learning model of self-disclosure developed in Chapter 2 is an important linguistic and methodological advance in self-disclosure work. It shows that post length, emotional valence, social distance, social normativity, and topics are essential language constituents of self-disclosure. One surprising finding is that the social normativity feature is a positive predictor of self-disclosure rather than a negative one as we expected, which indicates that there is a positive norm of self-disclosure on Facebook. Compared to existing machine learning models, which are limited in terms of interpretability, generalizability or accuracy, this model is easy to understand and domain-independent, and performs moderately well. It is also low in computational complexity, so it can be applied quickly and at scale in real time environments.

In addition to the substantive results derived using the model, this research demonstrated the value of automated coding of SNS posts. Most studies of communication in SNSs are based on hand coding relatively small samples of conversations. Using the machine learning techniques in Chapter 2, we were able to largely replicate findings based on human-coded and self-report data and also to discover relationships that underpowered human-coding studies did not find (Chapter 3). In addition, Chapters 4-5 used automated coding to conduct analyses that would be infeasible with human coding; namely, the studies examined how life events and audience salience correlate with self-disclosure and social network growth.

Replication of the previous findings of poster characteristics: This work examines the relationship between individual differences among posters and broadcasting self-disclosure. Using the machine learning model of self-disclosure, it again confirms the patterns found in previous research on self-disclosure but places them in the context of social networking sites and at a large scale. This work shows that women disclose more than men, and individuals who have a stronger desire for impression management disclose less. Successful replication of the results also provides evidence that validates our machine learning model.

Self-disclosure patterns around social network changing events: This work adds to our understanding of the influence of major life events on self-disclosure and the motivations for self-disclosure in SNSs. Since self-disclosure is known to correlate with social network maintenance, we focused on two kinds of events that could affect individuals' social networks, namely, change of romantic relationship status and start of school. This work demonstrates that users disclose more when starting a new relationship and disclose less when breaking up with someone. College students, especially freshmen, self-disclose more during the beginning of school.

These results make two theoretical contributions to self-disclosure research. First, they are consistent with the hypothesis that social approval is the default motive for broadcast self-disclosure and outweighs other motives. While seeking social approval motivates people who start a new relationship or enter college to self-disclose more, it makes people who recently ended a relationship to disclose less because of the negative interpretation of a breakup. Second, the results suggest that positive events (start of a new relationship and school) correlate with more broadcast self-disclosure, whereas negative events (breakups) correlate to less disclosure. However, this claim is based on the analyses of two special events and samples, so further investigation is necessary before it can be fully confirmed.

Ambiguous effect of context collapse on self-disclosure: Context collapse refers to the collapse of an individual's multiple audiences into a single context. One important contribution of this work is that we examined whether and how context collapse influences broadcast self-disclosure by conducting a correlational study and two controlled experiments. In the correlational study in Chapter 3, the concept of context collapse was operationalized as social network size, density, and average tie strength. More friends implies that more social circles might be involved; on the other hand, higher network density and average tie strength suggests that there are fewer social clusters and more strong ties in a network, respectively. So, we assumed that network size is a positive indicator of context collapse, and network density and average tie strength are negative indicators. We discovered that individuals disclose more when their networks are smaller, denser and of higher average tie strength, which implies that context collapse negatively affects self-disclosure.

However, experiments that increased audience salience tell a different story. The audience counter experiment presented to users the number of potential people who could see a status update in the status update composer, and the privacy checkup experiment showed a dialogue that enabled users to review their privacy settings for the status update composer. Results of the two experiments suggest that these design changes had no effect of self-disclosure, but a follow-up analysis of the privacy checkup dataset using propensity score matching shows that it has a small negative effect on self-disclosure, which suggests that if users actually see the details of their privacy settings, they reduce their self-disclosure.

We could consider the explanations for the contradictory findings between the correlational analyses and the experiments from two sides. Assuming the results of the correlational study and propensity score matching analysis are correct (i.e., context collapse is negatively related to self-disclosure), the most plausible explanation for the no effect in the two experiments is that the experimental manipulations are too subtle to make posters aware of multiple audiences. In contrast, if the results of the two experiments are correct, one possible explanation for the negative relationship between context collapse and self-disclosure found by the correlational study is that the three measurements we used to operationalize context collapse might not be appropriate, or there are some important confounds that were not included in the analysis. For example, although we believe that network size can positively indicate the degree of context collapse because more friends potentially mean more different social clusters in a network, it might not accurately measure social clusters or context collapse as we expected. Instead, it is possible that users with more online friends are also experienced Internet users and know the importance of protecting their personal information, so they disclose less in online environments. However, since all three measurements of context collapse suggest the same direction, the likelihood of this explanation is low.

Negative correlation between self-disclosure and growth in friend count: I also explored the social consequences of self-disclosure. Since broadcast self-disclosure is meant to be seen by many different people in one's social network simultaneously, I focused on the effect that broadcast self-disclosure has on expanding one's entire social network. The results show that a higher level of broadcast self-disclosure correlates with a smaller growth in friend count on SNSs. This finding is inconsistent with prior research that self-disclosure increases tie strength and our hypothesis that it increases friend count.

This result can be interpreted in three ways. One explanation is that the effect of self-disclosure on fostering friendship might not generalize to the number of one's friends. The Facebook news feed sorting algorithm ranks status updates in a way that they are more likely to be seen by strong ties than weak ties, so users who self-disclose in updates may continue to increase their tie strength with strong ties rather than become Facebook friends with acquaintances and thus increase number of friends. Another explanation is that the positive effect of self-disclosure on friendship is curvilinear, but not linear; therefore, too little and too much disclosure are bad for friendship development. The other explanation is that this study is a correlational one, so there are other pathways to interpret the result. For instance, it is possible that users add new friends first and change self-disclosure afterward. Those who add more new friends disclose less in order to protect their personal information.

Design implications for improving user experience: This work provides guidelines for SNS designers and developers. By knowing how SNS users navigate multiple audiences to present themselves, designers of these sites can improve their services by providing better affordances to users. According to the results of the analysis of audience factors, when network size and diversity become large enough that a person might not feel comfortable sharing personal news with friends, the site might nudge that person to share to a smaller group or custom list of friends.

Since self-disclosure is beneficial for friendship maintenance, and our results suggest that users like to talk about themselves when experiencing positive life events, SNSs can remind users of these events (e.g., a wedding anniversary), which may encourage them to share what happened and thus increase the bonding between users and their friends. In addition, the finding of the negative connection between broadcast self-disclosure and the increase of network size suggests that SNSs should prompt users to share to a smaller group of friends when the content they reveal might be too personal.

Although the current research used automated coding of self-disclosure language to better understand self-presentation behaviors in SNSs, the same technique could be used to improve the way these sites function. Using the machine learning model we have described, it would be possible to provide users

feedback about the degree of self-disclosure in a post that a user is about to share. It can also remind the user the constitution of audience if the post contains a high level of disclosure, so she / he has the opportunity to reconsider whether the post is appropriate for the audience. The model could also be used as the basis for building tools that enable users to show posts with different degrees of self-disclosure to friends with different tie strengths or in different social circles.

Generalizability of the results: The results of this work have better generalizability than past research. Natural language processing techniques and machine learning approaches were used to analyze the large archive of Facebook status updates, and the findings are based on a general sample of online populations and large-scale data analyses.

6.2 Discussion of General Limitations and Future Research

Moderate performance of the machine learning model: Although this thesis introduces a useful machine learning model to automatically measure self-disclosure, it only performs with moderate accuracy. While we have validated the model by demonstrating that it can replicate results reported in prior literature (Section 3.3.1), its annotation errors might limit our ability to discover subtle aspects of self-disclosure. For example, the moderate performance might be the reason that we were not able to identify significant results in the two experiments.

To improve the model, future work could merge similar features or untangle their relationships. We noted that some topic features overlapped with some of the text used to measure emotion valence and social distance features, such as family and first-person words. Combining features that represent similar concepts can reduce the number of features and sparsity in feature values, as well as avoid conflicts between features.

Another possible research direction would be to investigate interactions among features. The model was built using a linear kernel, which assumes features are independent. However, this assumption might be incorrect. The effect of some features on self-disclosure might depend on other features. Consider for example the interaction between the subject and type of emotion in a post: self-disclosure in a post talking about President Obama's emotion would be low regardless of types of emotion (e.g., "Obama is fine" and "Obama is angry" are both low in self-disclosure). In contrast, if the subject is "I", "I am angry" apparently has a higher level of disclosure than "I am fine."

Generalization and internationalization of the machine learning model: Another limitation of the model is that it was trained on a sample of 2,000 English language Facebook status updates collected at a

particular time from consenting workers at Amazon Mechanical Turk, so it may not be applicable to updates posted at other times, posts on other SNS platforms (e.g., Twitter), or posts written in other languages. Future work should gather ratings from a more representative sample and test the generalizability of the model on other SNS platforms.

We can consider using semi-supervised approaches to expand the training data and update the model over time. It is always easier to obtain unlabeled data than labeled data. Semi-supervised learning allows us to utilize both labeled and unlabeled data to build a better model ([Zhu, 2008](#)). One common method of semi-supervised learning is called *label propagation*, which assumes that similar data points are likely to have similar labels. Based on this assumption, it propagates a data point's label to its nearby unlabeled data points and thus increases the size training data. Using the 2,000 updates with self-disclosure labels, we can apply label propagation techniques to assign labels to unlabeled updates posted at other times and retrain the model using both the old and newly labeled updates. We expect this method will improve the model generalizability because it can expand the coverage of the training data.

Another way to make the model more generalizable is to update the three major features—topic, social distance, and social normativity—to incorporate the information from new posts. The three features were trained based on a random sample of eight million updates posted in 2014. However, SNS users produce millions of posts every day, so it is very likely that there are new emerging topics that do not belong to any of the 80 topics or celebrities not listed in the current celebrity dictionary. Social norms in SNSs might also change over time. Fortunately, the three features were extracted or trained with unsupervised approaches, so we can easily update them. For instance, we can select another random sample of updates posted in 2015 and combine it with the 2014 sample to build a new topic model. We can then compare the topics in the new topic model with the current 80 topics, and those that are most dissimilar to the 80 topics could be added to the original feature set. We can also use the same method that was used to construct the celebrity dictionary to update the dictionary. The language model used to compute the social normativity feature can also be rebuilt using both the 2014 and 2015 data samples. Once the feature information is updated, the last step would be to retrain the machine learning model using the new feature set.

Application of the results to other SNSs or social societies: One main limitation of this work is that the results may not directly carry over to other SNSs, like Twitter. This is expected because the nature and utility of SNSs are different. Consider, for example, the differences between Facebook and Twitter: the primary type of social connections on Facebook is mutual friendship, while social connections on Twitter are unidirectional; therefore, people often use the two sites for different purposes. They use Facebook to

interact with friends in their real life and Twitter to follow people they usually don't know but have similar interests. In that way, users are likely to share different types of content with different levels of self-disclosure on the two sites. [Choi and Bazarova \(2015\)](#) have identified differences in several aspects of self-disclosure between Facebook and Twitter, but their results are inconclusive. Though, I believe that the principles we found on Facebook would still hold on Twitter (e.g., gender difference in self-disclosure and the different effects of positive and negative life events on self-disclosure), but further investigation is required to validate this claim.

Furthermore, the results were derived from general populations and might not be applicable to more specialized groups, such as public personalities, movie stars, and politicians. Further, public figures may use handlers or agents, so instead of utilizing SNSs to maintain friendships, they may use SNSs as means to release news, control images, and develop fans or supporters. Unlike the general public, public figures may more strongly emphasize positive self-presentation for impression management rather than as a means disclose their personal information to maintain friendships.

Lack of direct evidence for the pathways examined: Although this thesis presents interesting social science results about the conditions eliciting self-disclosure and its downstream consequences, there is a lack of direct evidence about the pathways involved. For example, our study shows that life events are associated with self-disclosure, but we do not know what people talked about in their self-disclosures, especially whether they talked about the events associated with changes in self-disclosure. Did the people entering new relationships talk about their partners? Did freshmen talk about the start of school? Answers to these questions require more detailed coding or analysis. We also do not know how these events changed their cognitive state, which then caused them to disclose differently. Although we claimed that the reduced self-disclosure for the breakup sample supports the idea that social approval outweighs other motives (e.g., distress relief and identity clarification) because a breakup damages one's reputation and makes one look bad, we still need direct evidence to prove it. This pathway can be directly verified by surveying users who have recently suffered some negative experience to assess their motivation for self-disclosure and analyzing the survey data together with their posts on SNSs.

Although we have conducted both correlational and experimental studies to examine the effect of context collapse on self-disclosure, the results are still unclear. One possible way to more directly test the link between heterogeneity of people's network and their self-disclosure is to examine whether and how users disclose themselves differently in different SNS communication channels ([Bazarova & Choi, 2014](#)). For instance, if the theory of the lowest common denominator is correct, users with many strong and weak

ties would disclose minimally in a broadcast channel, but disclose much more when talking with strong ties in a private channel.

Combination of behavioral and social network information with self-reports: The findings reported in this dissertation are based on analyses of behavioral and social network information. As described above, if we can combine behavioral and social network information with self-reported data, we will be able to directly test some pathways, such as how self-disclosure motivations drive self-disclosing behaviors, and how users in the experiments of audience salience perceive the experimental manipulations.

Expansion of the research scope of self-presentation: This study examined self-presentation made through self-disclosure, since self-disclosure is related to relationship maintenance, the main motivation for people to use SNSs ([Pempek et al., 2009](#); [Barker, 2009](#)). However, another important motivation for them is to use SNSs to present a positive image of themselves and make themselves look good to their audiences ([Barash et al., 2010](#)), so future research can investigate SNS users' positive self-presentation and variables associated with it.

Differences between the attempt of posters and the perceptions of audiences: Another potential future work is to study the perceptions of audiences. So far the studies described in my thesis examined how posters present themselves in SNSs and the outcomes associated with the posters. I am also interested in inspecting the communication dynamics from the audience's point of view, such as how audiences perceive posters' messages and whether they really recognize what the posters try to convey in the messages. For example, people often try to impress their friends online, but we don't know how well they achieve this goal or what they discuss to try to make themselves look good. In an exploratory work, I have found that posters and outsiders agreed only modestly about how good a Facebook update made the poster appear, and posters generally thought that their posts make them look better than did the outsider judges ([Wang, Hinsberger, et al., 2016](#)). In the future, scholars can continue the research in this direction.

Design of better online environments for social interactions: In contrast to social interactions in offline settings, online environments offer researchers an environment where they can influence people's social behavior. This allows us to explore the underlying mechanisms of social dynamics in new situational contexts, which was the main focus of my thesis work. A better understanding of these social dynamics then helps us to improve presentation and communication features. One possible future work is to manipulate the specifics of how communication dynamics is supported to elucidate the mechanisms that lead to better social outcomes.

6.3 Closing Remarks

In conclusion, although social networking sites have become popular places for people to share personal information with others, these sites also pose new challenges for self-disclosure, especially one's audience and social network. Therefore, there has been increasing research interest in the language of broadcast self-disclosure as well as its causes and social consequences. This thesis contributes to this body of research by identifying important language signals of self-disclosure, developing an automatic model of self-disclosure, examining how self-disclosure varies with poster characteristics, life events, audience structure, and audience perceptions, and investigating what effects it has on relationships. It gives insights into the underlying mechanisms and motivations of broadcast self-disclosure and provides a clearer picture of self-disclosure on SNS.

References

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221), 509-514.

Alberghetti, R. (2014, May 22). Facebook's Privacy Dinosaur Is Back: New Members' Posts Aren't Automatically 'Public' Anymore. *Wall Street Journal*. Retrieved from <http://blogs.wsj.com/digits/2014/05/22/facebooks-privacy-dinosaur-is-back/>

Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, crowding*. Monterey, CA: Brooks/Cole.

Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*: Holt, Rinehart & Winston.

Archer, R. L. (1980). Self-disclosure *The self in social psychology* (pp. 183-205). Oxford: Oxford University Press.

Bak, J., Kim, S., & Oh, A. (2012). *Self-disclosure and relationship strength in Twitter conversations*. Paper presented at the ACM ACL '12, Jeju Island, Korea.

Bak, J., Lin, C.-Y., & Oh, A. (2014). *Self-disclosure topic model for classifying and analyzing Twitter conversations*. Paper presented at the Proceedings of the Conference on Empirical Methods in Natural Language Processing, Doha, Qatar.

Balani, S., & Choudhury, M. D. (2015). *Detecting and Characterizing Mental Health Related Self-Disclosure in Social Media*. Paper presented at the Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, Korea.

Barak, A., & Gluck-Ofri, O. (2007). Degree and reciprocity of self-disclosure in online forums. *Cyberpsychol Behav*, 10(3), 407-417. doi: 10.1089/cpb.2006.9938

Barash, V., Ducheneaut, N., Isaacs, E., & Bellotti, V. (2010). *Faceplant: Impression (Mis)management in Facebook Status Updates*. Paper presented at the ICWSM.

Barker, V. (2009). Older Adolescents' Motivations for Social Network Site Use: The Influence of Gender, Group Identity, and Collective Self-Esteem. *CyberPsychology & Behavior*, 12(2), 209-213. doi: 10.1089/cpb.2008.0228

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychol Bull*, 91(1), 3.

Bazarova, N. N., & Choi, Y. H. (2014). Self-Disclosure in Social Media: Extending the Functional Approach to Disclosure Motivations and Characteristics on Social Network Sites. *Journal of Communication*, 64(4), 635-657. doi: 10.1111/jcom.12106

Bazarova, N. N., Taft, J. G., Choi, Y. H., & Cosley, D. (2012). Managing Impressions and Relationships on Facebook: Self-Presentational and Relational Concerns Revealed Through the Analysis of Language Style. *Journal of Language and Social Psychology*. doi: 10.1177/0261927x12456384

Binder, J., Howes, A., & Sutcliffe, A. (2009). *The problem of conflicting social spheres: effects of network structure on experienced tension in social network sites*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3, 993-1022.

boyd, d. m. (2008). *Taken Out of Context: American Teen Sociality in Networked Publics*: University of California, Berkeley.

Branckaute, F. (2010). Facebook Statistics: The Numbers Game Continues. Retrieved August 22, 2015, from <http://www.blogherald.com/2010/08/11/facebook-statistics-the-numbers-game-continues/>

Brandtzæg, P. B., Lüders, M., & Skjetne, J. H. (2010). Too Many Facebook “Friends”? Content Sharing and Sociability Versus the Need for Privacy in Social Network Sites. *International Journal of Human-Computer Interaction*, 26(11-12), 1006-1030. doi: 10.1080/10447318.2010.516719

Burke, M., & Kraut, R. E. (2014). *Growing closer on facebook: changes in tie strength through social network site use*. Paper presented at the ACM CHI 2014, Toronto, Ontario, Canada.

Burke, M., Marlow, C., & Lento, T. (2010). *Social network activity and social well-being*. Paper presented at the ACM CHI 2010, Atlanta, Georgia, USA.

CareerBuilder. (2012). Thirty-Seven Percent of Companies Use Social Networks to Research Potential Job Candidates, According to New CareerBuilder Survey. Retrieved August 22, 2015, from <http://www.prnewswire.com/news-releases/thirty-seven-percent-of-companies-use-social-networks-to-research-potential-job-candidates-according-to-new-careerbuilder-survey-147885445.html>

Chen, S. F., & Goodman, J. (1996). *An empirical study of smoothing techniques for language modeling*. Paper presented at the ACM ACL '96, Santa Cruz, California.

Choi, Y. H., & Bazarova, N. N. (2015). Self-Disclosure Characteristics and Motivations in Social Media: Extending the Functional Model to Multiple Social Network Sites. *Human Communication Research*, 41(4), 480-500. doi: 10.1111/hcre.12053

Clarkson, P., & Rosenfeld, R. (1997). *Statistical Language Modeling Using the CMU-Cambridge Toolkit*. Paper presented at the Proceedings ESCA Eurospeech. <http://www.speech.cs.cmu.edu/SLM/toolkit.html>

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: a meta-analytic review. *Psychol Bull*, 116(3), 457-475.

Cozby, P. C. (1972). Self-Disclosure, Reciprocity and Liking. *Sociometry*, 35(1), 151-160. doi: 10.2307/2786555

Cozby, P. C. (1973). Self-disclosure: A literature review. *Psychol Bull*, 79(2), 73-91. doi: 10.1037/h0033950

Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., & Potts, C. (2013). *No country for old members: user lifecycle and linguistic change in online communities*. Paper presented at the WWW '13, Rio de Janeiro, Brazil.

Derlega, V. J., & Berg, J. H. (1987). *Self-Disclosure: Theory, Research and Therapy*: Springer.

Derlega, V. J., & Chaikin, A. L. (1977). Privacy and Self-Disclosure in Social Relationships. *Journal of Social Issues*, 33(3), 102-115. doi: 10.1111/j.1540-4560.1977.tb01885.x

Derlega, V. J., & Grzelak, J. (1979). Appropriateness of self-disclosure. In G. J. Chelune (Ed.), *Origins, patterns and implications of openness in interpersonal relationships* (pp. 151-176). San Francisco: Jossey-Bass.

Dindia, K., & Allen, M. (1992). Sex differences in self-disclosure: a meta-analysis. *Psychol Bull*, 112(1), 106-124.

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, 22(6), 469-493. doi: [http://dx.doi.org/10.1016/0047-2484\(92\)90081-J](http://dx.doi.org/10.1016/0047-2484(92)90081-J)

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. *Journal of Computer-Mediated Communication*, 11(2), 415-441. doi: 10.1111/j.1083-6101.2006.00020.x

Eulenstein, M., & Scissors, L. (2015). News Feed FYI: Balancing Content from Friends and Pages. Retrieved August 7, 2015, from <http://newsroom.fb.com/news/2015/04/news-feed-fyi-balancing-content-from-friends-and-pages/>

Facebook. (2010). What's on your mind? Retrieved December 25, 2015, from www.facebook.com/notes/Facebook-data-team/whats-on-your-mind/477517358858

Facebook. (2012). Prospectus. Retrieved August 22, 2015, from <http://www.sec.gov/Archives/edgar/data/1326801/000119312512240111/d287954d424b4.htm>

Facebook. (2015). Stats. Retrieved August 13, 2015, from <http://newsroom.fb.com/company-info/>

Garrison, A. M., Kahn, J. H., Sauer, E. M., & Florczak, M. A. (2012). Disentangling the effects of depression symptoms and adult attachment on emotional disclosure. *J Couns Psychol*, 59(2), 230-239. doi: 10.1037/a0026132

Gilbert, E., & Karahalios, K. (2009). *Predicting tie strength with social media*. Paper presented at the ACM CHI '2009, Boston, MA, USA.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*: Doubleday.

Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, 40(3-4), 237-264. doi: 10.1093/biomet/40.3-4.237

Greene, K. (2006). *Self-disclosure in personal relationships*: Cambridge University Press.

Grieve, R., Indian, M., Witteveen, K., Tolan, G. A., & Marrington, J. (2013). Face-to-face or Facebook: Can social connectedness be derived online? *Computers in Human Behavior*, 29(3), 604-609. doi: <http://dx.doi.org/10.1016/j.chb.2012.11.017>

Gross, R., & Acquisti, A. (2005). *Information revelation and privacy in online social networks*. Paper presented at the Proceedings of the 2005 ACM workshop on Privacy in the electronic society, Alexandria, VA, USA.

Hancock, J. T., Toma, C., & Ellison, N. (2007). *The truth about lying in online dating profiles*. Paper presented at the ACM CHI 2007, San Jose, California, USA.

Herring, S. C. (2007). A Faceted Classification Scheme for Computer-Mediated Discourse. *Language@Internet*, 4(1).

Hogan, B. (2010). The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, 30(6), 377-386. doi: 10.1177/0270467610385893

Jiang, L. C., Bazarova, N. N., & Hancock, J. T. (2011). The Disclosure–Intimacy Link in Computer-Mediated Communication: An Attributional Extension of the Hyperpersonal Model. *Human Communication Research*, 37(1), 58-77. doi: 10.1111/j.1468-2958.2010.01393.x

Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, 31(2), 177-192. doi: 10.1002/ejsp.36

Jourard, S. M. (1971). *Self Disclosure: An Experimental Analysis of the Transparent Self*. Oxford, England: John Wiley.

Jourard, S. M., & Lasakow, P. (1958). Some factors in self-disclosure. *J Abnorm Psychol*, 56(1), 91-98.

Kennedy, P. (2003). *A Guide to Econometrics* (5th ed.): MIT Press.

- Kivran-Swaine, F., & Naaman, M. (2011). *Network properties and social sharing of emotions in social awareness streams*. Paper presented at the Proceedings of the ACM 2011 conference on Computer supported cooperative work, Hangzhou, China.
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802-5805. doi: 10.1073/pnas.1218772110
- Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online social networks: why we disclose. *J Inf technol*, *25*(2), 109-125.
- Kruger, J., Epley, N., Parker, J., & Ng, Z.-W. (2005). Egocentrism over e-mail: Can we communicate as well as we think? *Journal of Personality and Social Psychology*, *89*(6), 925-936.
- Larzelere, R. E., & Huston, T. L. (1980). The Dyadic Trust Scale: Toward Understanding Interpersonal Trust in Close Relationships. *Journal of Marriage and Family*, *42*(3), 595-604. doi: 10.2307/351903
- Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychol Bull*, *107*(1), 34.
- Leimeister, J. M., Krcmar, H., Köbler, F., Vetter, C., & Riedl, C. (2010). *Social connectedness on Facebook-An explorative study on status message usage*. Paper presented at the 16th Americas Conference on Information Systems, Lima, Peru.
- Lin, H., Tov, W., & Qiu, L. (2014). Emotional disclosure on social networking sites: The role of network structure and psychological needs. *Computers in Human Behavior*, *41*, 342-350. doi: <http://dx.doi.org/10.1016/j.chb.2014.09.045>
- Marwick, A. E., & boyd, d. m. (2010). I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society*. doi: 10.1177/1461444810365313
- McAdams, D. (2011). Narrative Identity. In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of Identity Theory and Research* (pp. 99-115): Springer New York.
- McLean, K. C., Pasupathi, M., & Pals, J. L. (2007). Selves creating stories creating selves: a process model of self-development. *Pers Soc Psychol Rev*, *11*(3), 262-278. doi: 10.1177/1088868307301034
- Mehdizadeh, S. (2010). Self-presentation 2.0: narcissism and self-esteem on Facebook. *Cyberpsychol Behav Soc Netw*, *13*(4), 357-364. doi: 10.1089/cyber.2009.0257
- Meo, P. D., Ferrara, E., Fiumara, G., & Provetti, A. (2014). On Facebook, most ties are weak. *Commun. ACM*, *57*(11), 78-84. doi: 10.1145/2629438
- Miller, L. C., Berg, J. H., & Archer, R. L. (1983). Openers: Individuals who elicit intimate self-disclosure. *Journal of Personality and Social Psychology*, *44*(6), 1234-1244.

- Morton, T. L. (1978). Intimacy and reciprocity of exchange: A comparison of spouses and strangers. *Journal of Personality and Social Psychology*, 36(1), 72-81.
- Nguyen, M., Bin, Y. S., & Campbell, A. (2012). Comparing online and offline self-disclosure: a systematic review. *Cyberpsychol Behav Soc Netw*, 15(2), 103-111. doi: 10.1089/cyber.2011.0277
- Nosko, A., Wood, E., & Molema, S. (2010). All about me: Disclosure in online social networking profiles: The case of FACEBOOK. *Computers in Human Behavior*, 26(3), 406-418. doi: <http://dx.doi.org/10.1016/j.chb.2009.11.012>
- Omarzu, J. (2000). A Disclosure Decision Model: Determining How and When Individuals Will Self-Disclose. *Personality and Social Psychology Review*, 4(2), 174-185. doi: 10.1207/s15327957pspr0402_05
- OpenSource. (2010). The Apache OpenNLP library. from <https://opennlp.apache.org/>
- Oswald, D. L., Clark, E. M., & Kelly, C. M. (2004). Friendship Maintenance: An Analysis of Individual and Dyad Behaviors. *Journal of Social and Clinical Psychology*, 23(3), 413-441.
- Park, N., Jin, B., & Jin, S.-A. A. (2011). Effects of self-disclosure on relational intimacy in Facebook. *Computers in Human Behavior*, 27(5), 1974-1983. doi: <http://dx.doi.org/10.1016/j.chb.2011.05.004>
- Park, N., Lee, S., & Kim, J. H. (2012). Individuals' personal network characteristics and patterns of Facebook use: A social network approach. *Computers in Human Behavior*, 28(5), 1700-1707. doi: <http://dx.doi.org/10.1016/j.chb.2012.04.009>
- Parks, M. (2010). *Who are Facebook friends? Exploring the composition of Facebook friend networks*. Paper presented at the Proceedings of the Annual Meeting of the International Communication Association, Singapore.
- Pasupathi, M. (2001). The social construction of the personal past and its implications for adult development. *Psychol Bull*, 127(5), 651-672.
- Pasupathi, M., McLean, K. C., & Weeks, T. (2009). To tell or not to tell: disclosure and the narrative self. *J Pers*, 77(1), 89-123. doi: 10.1111/j.1467-6494.2008.00539.x
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA, US: Academic Press.
- Pearce, W. B., & Sharp, S. M. (1973). Self-disclosing communication. *Journal of Communication*, 23, 409-425.
- Pempek, T. A., Yermolayeva, Y. A., & Calvert, S. L. (2009). College students' social networking experiences on Facebook. *Journal of Applied Developmental Psychology*, 30(3), 227-238. doi: <http://dx.doi.org/10.1016/j.appdev.2008.12.010>

Pennebaker, J. W. (1997). *Opening Up: The Healing Power of Expressing Emotion*. New York: Guilford Press.

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). *The development and psychometric properties of LIWC2007*. LIWC.net, Austin, Texas.

Pennebaker, J. W., Kiecolt-Glaser, J. K., & Glaser, R. (1988). Disclosure of traumas and immune function: health implications for psychotherapy. *J Consult Clin Psychol*, *56*(2), 239-245.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, *54*(1), 547-577.

Porter, M. (2006). Porter stemmer. from <http://tartarus.org/martin/PorterStemmer/>

Rimé, B., Finkenauer, C., Luminet, O., Zech, E., & Philippot, P. (1998). Social Sharing of Emotion: New Evidence and New Questions. *European Review of Social Psychology*, *9*(1), 145-189. doi: 10.1080/14792779843000072

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55. doi: 10.1093/biomet/70.1.41

Rust, J., & Golombok, S. (2009). Psychometric assessment of personality in occupational settings *Modern Psychometric: The Science of Psychological Assessment* (Third ed., pp. 165-182). New York, NY: Routledge.

Ryan, M. B., & Kahn, J. H. (2015). Talking about daily emotional events: Psychological well-being moderates the intensity–disclosure link. *Personality and Individual Differences*, *83*(0), 223-227. doi: <http://dx.doi.org/10.1016/j.paid.2015.04.027>

Schlenker, B. R. (1980). *Impression Management: The Self-concept, Social Identity, and Interpersonal Relations*: Brooks/Cole Publishing Company.

Segal, J. A. (2014). Social Media Use in Hiring: Assessing the Risks. *HR Magazine*, *59*.

Semiocast. (2012). Twitter reaches half a billion accounts more than 140 millions in the U.S. Retrieved May 1, 2015, from http://semiocast.com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US

Shevade, S. K., Keerthi, S. S., Bhattacharyya, C., & Murthy, K. R. K. (2000). Improvements to the SMO algorithm for SVM regression. *Neural Networks, IEEE Transactions on*, *11*(5), 1188-1193. doi: 10.1109/72.870050

Smyth, J. M., Stone, A. A., Hurewitz, A., & Kaell, A. (1999). Effects of writing about stressful experiences on symptom reduction in patients with asthma or rheumatoid arthritis: a randomized trial. *JAMA*, *281*(14), 1304-1309.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, *30*(4), 526-537. doi: 10.1037/h0037039

Special, W. P., & Li-Barber, K. T. (2012). Self-disclosure and student satisfaction with Facebook. *Computers in Human Behavior*, *28*(2), 624-630. doi: <http://dx.doi.org/10.1016/j.chb.2011.11.008>

Spera, S. P., Buhrfeind, E. D., & Pennebaker, J. W. (1994). Expressive Writing and Coping with Job Loss. *Academy of Management Journal*, *37*(3), 722-733. doi: 10.2307/256708

Sprecher, S., Treger, S., & Wondra, J. D. (2013). Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions. *Journal of Social and Personal Relationships*, *30*(4), 497-514.

Stutzman, F., Capra, R., & Thompson, J. (2011). Factors mediating disclosure in social network sites. *Computers in Human Behavior*, *27*(1), 590-598. doi: <http://dx.doi.org/10.1016/j.chb.2010.10.017>

Tamir, D. I., & Mitchell, J. P. (2012). Disclosing information about the self is intrinsically rewarding. *Proceedings of the National Academy of Sciences*, *109*(21), 8038-8043. doi: 10.1073/pnas.1202129109

Thelwall, M. (2008). Social networks, gender, and friending: An analysis of MySpace member profiles. *Journal of the American Society for Information Science and Technology*, *59*(8), 1321-1330. doi: 10.1002/asi.20835

Tidwell, L. C., & Walther, J. B. (2002). Computer-Mediated Communication Effects on Disclosure, Impressions, and Interpersonal Evaluations: Getting to Know One Another a Bit at a Time. *Human Communication Research*, *28*(3), 317-348. doi: 10.1111/j.1468-2958.2002.tb00811.x

Utz, S. (2015). The function of self-disclosure on social network sites: Not only intimate, but also positive and entertaining self-disclosures increase the feeling of connection. *Computers in Human Behavior*, *45*, 1-10. doi: 10.1016/j.chb.2014.11.076

Vitak, J. (2012). The Impact of Context Collapse and Privacy on Social Network Site Disclosures. *Journal of broadcasting & electronic media*, *56*(4), 451-470. doi: 10.1080/08838151.2012.732140

Wang, Y.-C., Burke, M., & Kraut, R. E. (2013). *Gender, topic, and audience response: an analysis of user-generated content on facebook*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

Wang, Y.-C., Burke, M., & Kraut, R. E. (2016). *Modeling Self-Disclosure in Social Networking Sites*. Paper presented at the Proceedings of the ACM 2016 conference on Computer Supported Cooperative Work and Social Computing, San Francisco, CA.

Wang, Y.-C., Hinsberger, H., & Kraut, R. E. (2016). *Does Saying This Make Me Look Good? How Posters and Outsiders Evaluate Facebook Updates*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA.

Wang, Y.-C., Kraut, R. E., & Levine, J. M. (2015). Eliciting and Receiving Online Support: Using Computer-Aided Content Analysis to Examine the Dynamics of Online Social Support. *J Med Internet Res*, 17(4), e99. doi: 10.2196/jmir.3558

Wikipedia. (2015). List of emoticons. Retrieved August 22, 2015, from https://en.wikipedia.org/wiki/List_of_emoticons

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*: Morgan Kaufmann Publishers Inc.

Zhao, C., Hinds, P., & Gao, G. (2012). *How and to whom people share: the role of culture in self-disclosure in online communities*. Paper presented at the Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, Seattle, Washington, USA.

Zhu, X. (2008). Semi-Supervised Learning Literature Survey.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.

Appendix A: Informed Consent Used in the Amazon Mechanical Turk Task

Consent form to provide and rate your recent Facebook text post

This Amazon Mechanical Turk HIT is part of a research study conducted by Robert E Kraut and his doctoral student, Yi-Chia Wang at Carnegie Mellon University.

The purpose of the research is to collect Facebook status updates so that you and other people can classify the amount and type of information revealed in them. The collected data will be used to guide our research on social networking sites.

Procedures

In this HIT you will be asked to copy and paste your most recent Facebook status update into a form and rate them in terms of self-disclosure (i.e., how much personal information they reveal about you) and self-enhancement (i.e., how good they make you appear). You should finish the task in less than three minutes.

In a later stage in this research, we will show a sample of these status updates to research assistants or other observers so that they can also rate the degree of self-disclosure and self-enhancement they contain. The status-updates we show to others will be anonymous, so that no one can link these updates to you or your Facebook account.

Participant Requirements

Participants in this study must be 18 or older. The study is only for people who have a Facebook account and post status updates on Facebook in English.

Risks

The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life or during other online activities. The risk is minimal because the information you share is anonymous.

Benefits

There may be no personal benefit from your participation in the study but the knowledge received may be of value to humanity.

Compensation and Costs

You will be paid \$0.50 to complete the HIT.

There will be no cost to you if you participate in this study.

Confidentiality

The data captured for the research does not include any personally identifiable information about you, except for your Amazon Mechanical Turk worker identification number, which will only be used to pay you. Amazon Mechanical Turk worker identification number will not be stored after you are paid and will not be linked to the status updates you provide.

By participating in this research, you understand and agree that Carnegie Mellon may be required to disclose your consent form and data as required by law, regulation, subpoena or court order.

By participating, you understand and agree that the anonymous data and information gathered during this study may be used by Carnegie Mellon and published and/or disclosed by Carnegie Mellon to others outside of Carnegie Mellon.

Right to Ask Questions and Contact Information

If you have any questions about this study, you should feel free to ask them by contacting the Principal Investigator now at:

Yi-Chia Wang
Ph.D. Student
Language Technologies Institute
School of Computer Science
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
412-268-6591
yichiaw@cs.cmu.edu

Robert E Kraut
Herbert A. Simon Professor of Human-Computer Interaction
Human-Computer Interaction Institute
Tepper School of Business
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
412-268-7694
robert.kraut@cmu.edu

If you have questions later, desire additional information, or wish to withdraw your participation please contact the Principal Investigator by mail, phone or e-mail in accordance with the contact information listed above.

If you have questions pertaining to your rights as a research participant; or to report objections to this study, you should contact the Office of Research integrity and Compliance at Carnegie Mellon University. Email: irb-review@andrew.cmu.edu . Phone: 412-268-1901 or 412-268-5460.

The Carnegie Mellon University Institutional Review Board (IRB) has approved the use of human participants for this study.

Voluntary Participation

Your participation in this research is voluntary. You may discontinue participation at any time during the research activity.

Appendix B: Screenshot of the MTurk Task

Provide and rate your recent Facebook text posts
Requester: Carnegie Mellon University - Social Computing Lab **Reward:** \$0.50 per HIT **HITs Available:** 1 **Duration:** 20 minutes
Qualifications Required: HIT approval rate (%) is not less than 98; Location is US

Instructions

This Amazon Mechanical Turk HIT is part of a research study conducted by Robert E Kraut and Yi-Chia Wang at Carnegie Mellon University. The purpose of the research is to collect a sample of anonymous Facebook status updates and evaluate the type of information revealed in them. You will copy and paste your most recent Facebook status update into a form and rate it in terms of self-disclosure (i.e., how much personal information they reveal about you) and self-enhancement (i.e., how good they make you appear). You should finish the task in less than five minutes.

In a later stage in this research, we will ask research assistants at Carnegie Mellon University to also rate the degree of self-disclosure and self-enhancement the status update contains. The status update will be anonymous, so that no one can link the update to you or your Facebook account.

To find out more information go to the full consent form by clicking [this link](#).

Participants in this study must be 18 or older. The study is only for native English speakers who are active Facebook users and have posted English text content in the last month on Facebook.

We would like to know about the content you share on Facebook.

How many days in the past week did you use Facebook?

How many friends do you have on Facebook? (You can find the number by going to your Facebook timeline. The number is shown next to your profile picture.)

How many photos do you have on Facebook? (You can find the number by going to your Facebook timeline. The number is shown next to your profile picture.)

On what date did you post it? (You must enter a date with the format yyyy-mm-dd., e.g., 2014-01-20):

Please copy and paste your most recent English text post on Facebook here:
(Note: we only accept a pure text post that has more than five words and does not contain any photos, videos, or links)

To what extent does this post involve:

	1 (Not at all)	2	3	4	5	6	7 (Completely)
personal information about yourself or people close to you, such as accomplishments, family, or problems you are having?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
personal thoughts on past events, future plans, appearance, health, wishful ideas, etc.?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
your feelings and emotions , including concerns, frustrations, happiness, sadness, anger, and so on?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
what is important to you in life?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
your close relationships with other people?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

You must ACCEPT the HIT before you can submit the results.

Appendix C: Familiar Nickname Dictionary

Angel	Darlington	Luuuuve	Sugar
Babe	Dear	Luvv	Sugarpie
Babes	Dearest	Ma	Sugarplum
Baby	Dearheart	Mami	Sunshine
Babycheeks	Dimples	Merlin	Sweet
Babylicious	Doll	Mistress	Sweetheart
Bear	Doobie	Muffin	Sweetie
Boo	Doodle	Munchkin	Sweetness
Braveheart	Dove	Pa	Sweets
Bubba	Dreamlover	Papa	Sweetums
Bubbles	Firecracker	Papi	Sweety
Bunny	Firefly	Poohbear	Tadwinks
Buttercup	Fluffy	Pookie	Teddy
Butterfly	Gorgeous	Precious	Tempress
Butterscotch	Gumdrop	Prince	Tiger
Canoodle	Handsome	Princess	Tigress
Captain	Hon	Puddin	Tulip
Chickadee	Honey	Pumpkin	Tweetums
Cookie	Honeybunch	Punkin	Twinkie
Cowboy	Honeypot	Puppy	Waffles
Cowgirl	Hunny	Pussycat	Wiggles
Cuddle	Kitten	Scrumptious	Wifey
Cuddles	Kitty	Sexy	Witchy
Cuddly	Love	Snickerdoodle	Woobie
Cupcake	Loveable	Snookie	Wookie
Cupid	Lovealump	Snookums	Wookums
Cutie	Lovebird	Snowflake	Wuggle
Cutiehead	Lovely	Snuggle	Wuggles
Cutiepie	Lover	Snuggles	
Darlin	Lovie	Starfish	
Darling	Lubs	Starshine	