Three Essays on Social Media: The Effect of Motivation, Participation, and Sentiment on Performance

Mohammad Salehan

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2015

APPROVED:

Dan Kim, Committee Chair
Mary Jones, Committee Member and Chair of the Department of Information Technology and Decision Sciences
Audhesh Paswan, Committee Member
Victor Prybutok, Committee Member
O. Finley Graves, Dean of the College of Business
Costas Tsatsoulis, Dean of the Toulouse Graduate School
In recent years, social media has experienced tremendous growth in the number of users. Facebook alone has more than 1.3 billion active users and Twitter has attracted over 600 million active users. Social media has significantly changed the way humans communicate. Many people use social media to keep in touch with family and friends and receive up-to-date information about what happens around the world. Politicians are using social media to support their campaigns. Use of social media is not restricted to individuals and politicians. Businesses are now using social media to promote their products and services. Many companies maintain Facebook and Twitter accounts to keep in touch with their customers. Consumers also use social media to receive information about products/services. Online product reviews are now an important source of information for consumers.

This dissertation aims to address one fundamental research question: how do individual differences among users lead to different levels of performance on social media? More specifically, this dissertation investigates the motivations of use and the predictors of performance in the context of social media. We utilize sentiment mining to predict performance in different types of social media including information diffusion in Twitter and helpfulness and readership of online consumer reviews. The results show how different motivations lead to different levels of participation in social media and level of participation consequently influences performance. We also find that sentiment of the messages posted on social media significantly influence their performance.
Copyright 2015

by

Mohammad Salehan
ACKNOWLEDGMENTS

I would never have been able to finish my dissertation without the guidance of my committee members, help from friends, and support from my family. I would like to express my deepest gratitude to my advisor, Dr. Dan Kim, for his excellent guidance, caring, patience, and providing me with an excellent atmosphere for doing research. I would also like to thank Dr. Mary Jones, Dr. Audhesh Paswan, and Dr. Victor Prybutok for guiding my research for the past two years. Finally, I would also like to thank my parents. They were always supporting me and encouraging me with their best wishes.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>PROLOGUE</td>
<td>1</td>
</tr>
<tr>
<td>ESSAY I: USING SOCIAL NETWORKING SERVICES AT WORK AND AT HOME:</td>
<td>5</td>
</tr>
<tr>
<td>MOTIVATION, PARTICIPATION, AND PERFORMANCE</td>
<td></td>
</tr>
<tr>
<td>1.1. Introduction</td>
<td>5</td>
</tr>
<tr>
<td>1.2. Literature Review and Theoretical Background</td>
<td>8</td>
</tr>
<tr>
<td>1.2.1. Online Social Networking Services</td>
<td>8</td>
</tr>
<tr>
<td>1.2.2. Motivation-Participation-Performance Framework</td>
<td>10</td>
</tr>
<tr>
<td>1.2.3. Motivation</td>
<td>11</td>
</tr>
<tr>
<td>1.2.4. Participation</td>
<td>14</td>
</tr>
<tr>
<td>1.2.5. Performance</td>
<td>15</td>
</tr>
<tr>
<td>1.3. Research Model and Hypotheses</td>
<td>16</td>
</tr>
<tr>
<td>1.3.1. Motivations behind Participation</td>
<td>18</td>
</tr>
<tr>
<td>1.3.2. The Effect of Participation on Performance</td>
<td>23</td>
</tr>
<tr>
<td>1.4. Research Methodology and Data Collection</td>
<td>25</td>
</tr>
<tr>
<td>1.4.1. Measures</td>
<td>25</td>
</tr>
</tbody>
</table>
ESSAY III: THE EFFECT OF EMOTIONAL AROUSAL ON INFORMATION DIFFUSION IN SOCIAL MEDIA
LIST OF TABLES

Table 1-Descriptive Statistics and Reliability Coefficients (Essay I) ........................................... 29
Table 2-Discriminant Factor Analysis (Essay I) ........................................................................... 30
Table 3-Hypothesis Testing Results (Essay I) .............................................................................. 31
Table 4-Sobel Test Results (Essay I) ........................................................................................... 33
Table 5-Descriptive Statistics (Essay II) ....................................................................................... 56
Table 6-Correlation Matrix of Independent Variables (Essay II) ................................................. 58
Table 7-Hypothesis Testing Results (Essay II) ............................................................................. 59
Table 8-Group Analysis (Essay II) ............................................................................................... 60
Table 9-Group Analysis of Helpfulness based on Total Number of Votes (Essay II) ................. 60
Table 10-Descriptive Statistics (Essay III) ................................................................................... 78
Table 11-Correlations (Essay III) ................................................................................................. 81
Table 12-Hypothesis Testing Results (Essay III) ......................................................................... 82
LIST OF FIGURES

Figure 1-Research Model (Essay I) .............................................................................................. 17
Figure 2-PLS Analysis Results (Essay I)...................................................................................... 31
Figure 3-Post-hoc Analysis Results (Essay I)............................................................................... 32
Figure 4-Research Model (Essay II) ............................................................................................. 48
Figure 5-System Design of Sentiment Extraction and Mining Process (Essay II) ....................... 54
Figure 6-Distribution of The Measures (Essay II) ........................................................................ 56
Figure 7-Research Model Analysis Results (Essay II) .................................................................. 59
Figure 8-Research Model (Essay III) ............................................................................................ 75
Figure 9-System Design of Sentiment Extraction (Essay III) ........................................................ 80
In recent years, social media has experienced tremendous growth in the number of users. Facebook alone has more than 1.3 billion active users (StatisticsBrain.com, 2014a) and Twitter has attracted more than 600 million active users (StatisticsBrain.com, 2014b). Social media has significantly changed the way humans communicate. Many people use social media to keep in touch with family and friends and receive up-to-date information about what happens around the world. Politicians are using social media to support their campaigns (Naughton, 2012). Use of social media is not restricted to individuals and politicians. Businesses are now using social media to promote their products and services. Many companies maintain Facebook and Twitter accounts to keep in touch with their customers. Customers also use social media to receive information about products/services. Online product reviews are now an important source of information for consumers (Nielsen, 2012).

Previous studies have looked at both antecedents and consequences of social media use. In terms of antecedents, previous studies look at different factors that motivate people to engage in social media. Reputation (Wasko & Faraj, 2005b), expected relationships (Hsu & Lin, 2008), enjoyment in helping others, organizational rewards (H.-F. Lin, 2007), emotional influence (Tsai & Bagozzi, 2014b), social ties (Chai, Das, & Rao, 2011), social identity (C. M. Cheung & Lee, 2010; Tsai & Pai, 2014a), and loneliness (Xu, Ryan, Prybutok, & Wen, 2012) are among the motivations for social media use. In terms of consequences, research looks at how social media use influences individuals and organizations. People and organizations use social media to improve their performance. The significant impact of social media on the success of individuals, businesses, and politicians has created a climate of competition on social media. Those who perform better on social media will gain more benefits. For example, use of social media may
lead to increased firm innovation capability (H.-F. Lin, 2007; Z. Wang & Wang, 2012), improved operational and financial performance (Z. Wang & Wang, 2012), and enhanced knowledge utilization (C.-J. Chen & Hung, 2010). Consequently, investigation of the factors influencing performance on social media is an important research avenue.

Previous research shows three broad categories of predictors of performance on social media. The first category contains structural factors which are related to the structure of the network such as structural capital (Borgatti & Foster, 2003). The second category contains the factors related to the content of the network. A prominent example of content is resource access which explains how different nodes of a network access and benefits from resources available in the network (Borgatti & Foster, 2003). While the first two categories of SNS performance are about the network itself, the third category, individual differences, explains how individual characteristics of each node influence their performance (Kalish & Robins, 2006; Obstfeld, 2005; Sasovova, Mehra, Borgatti, & Schippers, 2010). This dissertation focuses on individual differences that determine performance on social media and aims to address one fundamental research question:

RQ: How do individual differences among users lead to different levels of performance on social media?

This dissertation investigates the motivations of use and the predictors of performance in the context of social media. More specifically, we utilize sentiment mining to predict performance on different types of social media. Sentiment mining refers to the use of natural language processing and computational linguistics to find and extract subjective information from text data. Sentiment mining is usually performed using automated tools that provides benefits such as effective information retrieval, automated cyber risk management, and increased
business profits (Bai, 2011). It also facilitates processing of large amounts of data. Social media is an important source of big data which makes it quite suitable for text mining purposes (Russell, 2013).

**Three Essays in This Dissertation**

Essay I is motivated by lack of an integrated view of how motivations, participation, and performance are interrelated in the context of social media. We use the framework suggested by Kim et al. (2007) to identify three motivations for social media use: hedonic, utilitarian, and social. There are two types of social motivation: vertical and horizontal. We use motivation and performance theory (Campbell & Pritchard, 1976) to investigate the relationship between motivation, participation, and performance. Essay I provides a holistic view of how motivations lead to participation and how participation leads to performance of individuals in the context of social media.

Essay II is motivated by the issue of online consumer reviews. The goal of the second essay is to analyze the predictors of the two performance measures using a sentiment mining approach. We identify two important performance measures in this area: readership and helpfulness. This essay focuses on investigating the effect of emotional valence (i.e., positive and negative emotions) on the performance of online reviews.

Essay III uses sentiment mining to study the performance of information diffusion in social media. This essay focuses on arousal dimension of sentiment (i.e., calm or excited). Previous research suggests that high-arousal emotions significantly influence the decision making process of individuals and provokes action. Low-arousal emotions, in contrast, are shown to have negative impact on the quality of decision making (Kaufman, 1999). Hence, we suggest that messages containing high arousal emotions are more likely to be retweeted.
Moreover, people process high-arousal negative words faster than neutral words while they process low-arousal negative words slower than neutral words (Hofmann, Kuchinke, Tamm, Võ, & Jacobs, 2009; Kuchinke, 2007; Kuchinke, Võ, Hofmann, & Jacobs, 2007). Therefore, we propose that high-arousal negative sentiment improves information diffusion in social media. We also suggest that low-arousal negative sentiment has a negative effect on the diffusion of the message.
ESSAY I: USING SOCIAL NETWORKING SERVICES AT WORK AND AT HOME: MOTIVATION, PARTICIPATION, AND PERFORMANCE

1.1. Introduction

Advancement of information and communication technologies including the emergence of social media has changed our daily life and social activities (Bulmer & DiMauro, 2009; Ellison, Lampe, & Steinfield, 2009; K.-Y. Lin & Lu, 2011). In early days, the Internet was mainly used for searching for information and knowledge. The information flowed mostly in a one-way direction; Information was generated by a limited group of people and most users were consumers rather than producer of information. After the advent of Web 2.0, information flow became more of a two-way one (Kim, et al., 2009). Today, practically any online user can produce content and share it others on the Internet. Weblogs, for instance, have become a place for easy dissemination of personal opinions and ideas. Using online social networking services (SNS), online users can freely communicate and share information with many other users who are scattered all over the world (Chai & Kim, 2011; Gunawardena et al., 2009; Kho, 2007).

SNS are platforms to form and manage personal connections and create a foundation for human relationships (Gunawardena et al., 2009; Valenzuela, Park, & Kee, 2009a). Although social dynamics of community life stimulate the motivations for participation in a social setting, there are significant differences between online communities and communities in real life in terms of factors that predict participation (Parameswaran & Whinston, 2007). SNS have continuously evolved as a space to produce, share, and consume information or content through participation (Bulmer & DiMauro, 2009). People use SNS not only to maintain their relationships with friends and family or to meet new people, but also to participate in groups in order to communicate with people with whom they share similar interests (Ellison et al., 2009;
Gunawardena et al., 2009; Kwon & Wen, 2010). Most SNS users have immersion tendencies where they check various content several times a day, add multimedia content, and modify the look and feel of their profiles (Chai & Kim, 2011). While SNS are still in the early stages of growth, they are influencing power and public culture (Boyd & Ellison, 2008; Ross et al., 2009). Hence, study of the motivations for participation is an important part of SNS research.

Meanwhile, online social networking has become increasingly pervasive in the workplace as a platform to directly produce and distribute content (Wilson, 2009). Many organizations have expressed concerns regarding the negative impact of SNS on employee productivity, and some have blocked their employees’ access to online social media during working hours (Weisul, 2011). However, recent research suggests that SNS can benefit organizations (C.-J. Chen & Hung, 2010; H.-F. Lin, 2007; Z. Wang & Wang, 2012). Those who use SNS tend to prefer to know more about their colleagues’ lives. This kind of information is used to facilitate social interactions that both directly and indirectly support job-related tasks (Ellison et al., 2009). Consequently, an investigation of the effect of SNS use on the performance in terms of both personal and job performance is an important part of the SNS research. In summary, there are two important research questions in the context of SNS: (1) what motivates users to participate in SNS, and (2) how does users’ participation in SNS influence their personal life and job performance?

Several previous studies have tried to address the above research question. However, large gaps remain in the literature. Although previous research identifies different motivations of why SNS users use (Tong, Wang, Tan, & Teo, 2013; Xu et al., 2012), several underlining motivations to participate in SNS activities (for what ultimate ends) remain unexplored. Thus, a profound view of different motivations for SNS use will provide a better understanding of why
people use SNS. Moreover, participation in SNS has been mostly viewed in the form of sharing or knowledge contribution while other types of SNS participation, such as collaboration, remain unexplored (Chai et al., 2011; H. H. Chang & Chuang, 2011; H.-F. Lin, 2007; Tsai & Bagozzi, 2014a; Wasko & Faraj, 2005b). This is particularly important because many users are now collaborating on SNS to attain specific goals. In terms of performance of SNS use, previous research has looked at how SNS influences the individuals and organizations. However, while the effect of SNS use on individual performance has been investigated by previous research, more elaboration is required to explore the effect of SNS use on different dimensions of individual performance (Chiu, Hsu, & Wang, 2006; Phang, Zhang, & Sutanto, 2013). Finally, while the relationship between motivations, participation, and performance has received attention from previous research, previous studies rather look at either how motivations lead to participation or how participation leads to performance (see Appendix A). Thus, a holistic view of why, how, and for what purpose users adopt SNS is not found in the literature. This is particularly important because it will provide a clear picture of how different motivations of users for SNS participation eventually influence their lives.

Based on motivation and performance theory (Campbell & Pritchard, 1976), Roberts et al. (2006) propose a theoretical framework describing the relationships among motivations, participation, and performance in the context of open source software development. They suggest that the mechanics of engagement in open source software development is formed by three main components: motivation, participation, and performance. Motivation refers to the inspiration for joining the network. Participation states the type and level of contribution of an individual in the open source software development network. Performance refers to the final outcome of the process for participations. Because open source software development can be
seen as a collaboration process among the members of a social network (Madey, Freeh, & Tynan, 2005), we apply the theoretical lens provided by Roberts et al. (2006) to investigate our research questions.

Using the theoretical model developed by Roberts et al. (2006), this study has three primary objectives: (1) to propose a research model explaining the relationship between motivation, participation, and performance in the SNS context at individual level of analysis, (2) to empirically test the proposed research model using data collected from SNS users, and (3) to enhance scholarship by providing academicians and practitioners with insight concerning the antecedents and consequences of using SNS on both personal life and the workplace. The results provide managers with greater insights into how SNS can benefit their organization.

The remaining parts of this paper are structured as follows. In the next section, we review previous research dealing with online social networking services and a motivation-participation-performance framework as theoretical background of the study. Then, we propose a research model and hypotheses. Next, we describe the research methodology that includes the measurement instrument development, data collection, and hypothesis testing procedures. After discussing the results of data analysis, finally, we discuss the findings and implications of the study.

1.2. Literature Review and Theoretical Background

1.2.1. Online Social Networking Services

Since the first online social networking service, SixDegrees.com which was launched in 1997, a large number of SNS have emerged (Y. Kim, Sohn, & Choi, 2011; Tokunaga, 2011). Online social networking service is an online service accessible via any Internet-enabled device which facilitates computer-mediated interaction among people for a certain purpose such as
collaboration and sharing interests, shared activities, or real-like connections (Kizza, 2013). While there is no clear classification of different types of SNS, White (2014) classifies seven major categories of SNS: general purpose Facebook-style social networking services for connecting to friends and family members (e.g., Facebook, Cyworld, and MySpace), short message service-based social networks (e.g., Twitter) for message sharing, professional social network services (e.g., LinkedIn and Konnects) for career-related goals, multimedia sharing services (e.g., YouTube, Instagram, and Snapchat) for sharing videos and pictures, informational social networks (e.g., the Student Room) providing opportunities for people to find answers to their questions, educational social networks (e.g., Academic.edu, GoodReads, Shelfari, and LibraryThing) for collaborating with others on academic projects, and hobby and special interest social networks (e.g., ActionProfiles and FanIQ) helping people find other people with similar interests.

Although each type of SNS is customized for few unique features, most SNS support general information sharing and networking functions such as making connections and uploading rich content including photographs, videos, and other digital content for both personal and business purposes (Boyd & Ellison, 2008). Facebook, as the most popular general purpose SNS, has a unique position. It is not only used for connecting to social connections, but also has other applications. Facebook can be used for multimedia sharing, career related tasks (through applications such as Glassdoor), informational and educational goals (through groups), and hobbies (through pages and groups). As a result, Facebook provides a comprehensive view of various applications of SNS.
1.2.2. Motivation-Participation-Performance Framework

Roberts et al. (2006) propose a theoretical model (i.e., motivation-participation-performance framework) to study the behavior of contributors in the open source software development context. They were initially interested in understanding how different motivations of contributors are related because motivations are an important component of individual’s behavior and performance. They were also interested to find out how motivations influence contribution and in turn how contribution influences performance. To attain these goals, they extended the motivation and performance theory proposed by Campbell and Pritchard (1976) to develop their theoretical framework. According to the theory, different people have different motivations which produce different behaviors. Different behaviors consequently lead to different levels of performance. In short, motivations shape behavior and behavior affects performance.

Wasko et al. (2004) also confirm this view and suggest that different motivations shape the participation of people in social networks. The study identifies two broad categories of intrinsic and extrinsic motivations related to open source software development participation. Intrinsic motivations are those that satisfy basic human needs for competence, control, and autonomy. In contrast, extrinsic motivations are inspired by the environment. The study also identifies the main form of behavior in the context of open source software development as participation in projects. Finally, performance is measured as the rankings assigned to each developer by the open source software community.

This framework is appropriate for the objectives of this study for two reasons. First, open source software development networks and SNS share common attributes that make them highly similar. Participation in both is entirely voluntarily and in the form of intellectual contribution.
Furthermore, interaction with others is a vital component of contribution. In fact, open source software development can be seen as a social network of collaboration among software developers (Madey et al., 2005). Second, the framework can be used to answer the questions of “why”, “how”, and “for what purpose” people participate in SNS. Motivations drive behavior and explain why individuals conduct behavior. In the case of this study, motivations explain why people use SNS. Behavior is the external appearance of motivation and describes how people pursue their motivations. In the context of SNS, participations can be observed in different forms such as sharing and collaboration and specifies what users can do to achieve the desired outcomes. Performance is the ultimate outcome of the behavior and illustrates for what purpose people act in a special manner. In the SNS context, performance is what users want to achieve by their SNS participation in both their daily life and their career.

1.2.3. Motivation

Motivations are general characteristics that elicit, control, and sustain actions taken to fulfill a need or want (Bolar, 2009). The Uses and gratifications (U&G) paradigm (Katz, 1959) explains motivations for individual adoption of media. This theoretical paradigm has been successfully used to identify user motivation for adoption of information and communication technologies (U. M. Dholakia, Bagozzi, & Pearo, 2004; Kuehn, 1994; Rafaeli, 1984; Stafford, Stafford, & Schkade, 2004). According to U&G, users are goal-oriented and select specific media based on their needs. Previous research shows three broad categories of motivations for adoption of Internet-related technologies: utilitarian, hedonic, and social (Zhou, Jin, Vogel, Fang, & Chen, 2011).

SNS users’ motivation can also be broadly classified as utilitarian motivation and hedonic motivation (Gu, Fan, Suh, & Lee, 2010; Hyllegard, Ogle, Yan, & Reitz, 2011). *Utilitarian*
motivation deals with the use of SNS for goal oriented, mission critical, rational, and decision effective user’s tasks (Gu et al., 2010; Hyllegard et al., 2011; K.-Y. Lin & Lu, 2011) while hedonic motivation refers to usage behaviors in search for happiness, fantasy, awakening, sensuality, and enjoyment. The benefit of hedonic motivation is experiential and emotional. The reason that hedonic SNS users like to use SNS is because they enjoy experiential and emotional pleasure derived from use (Gu et al., 2010; K.-Y. Lin & Lu, 2011; van der Heijden, 2004). Recent SNS tend to have both a utilitarian and hedonic nature, and users use them for the both purposes (Gu et al., 2010; Hyllegard et al., 2011). For example, teenage users prefer to use it mainly for social entertainment, while employees in the workplace spend much time on SNS for job-related purposes (Gu et al., 2010; Hyllegard et al., 2011; K.-Y. Lin & Lu, 2011).

In addition to hedonic and utilitarian motivation, social motivation also plays an important role in adoption and usage of online social media. Social motivation is derived by social benefits resulted from establishing and maintaining social interaction with others (U. M. Dholakia et al., 2004). Social motivation is especially essential in SNS (J. H. Kim, Kim, & Nam, 2010a). SNS is expanding the meeting place of human networks formed through face to face meetings (Bulmer & DiMauro, 2009). People are becoming accustomed to forming and maintaining close and appropriate relationships with various individuals in cyberspace (Kho, 2007). Previous research has looked at different types of social motivation for SNS adoption including seeking friends (Y. P. Chang & Zhu, 2011; Y. Kim et al., 2011), meeting new people (Y. P. Chang & Zhu, 2011), getting social support (Y. Kim et al., 2011), presenting personal identity (Hyllegard et al., 2011), developing reputation (Wasko & Faraj, 2005a), maintaining relationships (Hsu & Lin, 2008), enjoyment in helping others (H.-F. Lin, 2007), social influence (Tsai & Bagozzi, 2014a), social ties (Chai et al., 2011; Chai & Kim, 2011), sense of belonging
(Chai & Kim, 2011), social identity (C. M. Cheung & Lee, 2010; Kwon & Wen, 2010; Tsai & Pai, 2014b), social presence (Xu et al., 2012), telepresence (Kwon & Wen, 2010), building social capital (Ross et al., 2009; Valenzuela et al., 2009a) and others. Although research shows common motivations for SNS use across cultures, these motivations may have different weights in different cultures (D. J. Kim, 2008; Y. Kim et al., 2011).

Considering the above, one may conclude that broad categories of social motivation to use SNS are to maintain existing intimate social relationships such as family and friends, and to make new and professional social relationships for networking and business (Valenzuela et al., 2009a). Thus, social motivations can be either vertical or horizontal. Vertical social motivation refers to forward and backward linkages with close connections such as family and close friends. Recent studies suggest that vertical social motivation is the main objective of general SNS use (Boyd & Ellison, 2008; Stefanone, Lackaff, & Rosen, 2011) and is associated with bonding social capital (Gittell & Vidal, 1998). Horizontal social motivation refers to acquiring and/or maintaining new relationships with people who have similar interests and objectives (Boyd & Ellison, 2008; Byrd & Jasny, 2010). Because there are no face-to-face communications in the SNS context, SNS-based horizontal relationships tend to exclude prejudice (J. H. Kim et al., 2010a). In addition, SNS entails less social restrictions and sanctions; it enables new relationships to be built and allows people to communicate with each other with less formality and constraint (Boyd & Ellison, 2008; J. H. Kim et al., 2010a). SNS also supports horizontal social relationships by allowing users to browse through people and find someone with the intention of a later offline relationship, or to attend an event organized online (Tokunaga, 2011). Horizontal social motivation is associated with bridging social capital (Gittell & Vidal, 1998).
1.2.4. Participation

SNS are open online platforms used to share thoughts, experiences, and viewpoints (Chai & Kim, 2011; Gunawardena et al., 2009). SNS are also expanding the interactive collaboration of human networks that were formed through face to face meetings online (Bulmer & DiMauro, 2009). SNS users' participation consists of two major activities - i.e., sharing and collaboration (Chai et al., 2011; C.-J. Chen & Hung, 2010; Y. Chen & Xie, 2008; Chiu et al., 2006; Cross, Borgatti, & Parker, 2002; H. Kim, Suh, & Lee, 2013; H.-F. Lin, 2007; Ransbotham & Kane, 2011; Stieglitz & Dang-Xuan, 2013; Z. Wang & Wang, 2012; Wasko et al., 2004).

SNS facilitates distribution of content and enables people to share their ideas, knowledge, and digital content such as photos, music, and video with other SNS users (Boh & Wong, 2013; Boyd & Ellison, 2008; DiMicco et al., 2008; Gunawardena et al., 2009). Sharing intends to make information available to others (Olorunniwo & Li, 2010). Although privacy concerns significantly determine level of information sharing on SNS (Salehan, Mousavizadeh Kashipaz, & Xu, 2013), many people and organizations maintain public profiles through which they broadcast information to the general audience. Depending on the privacy settings of the user, it may or may not be possible for other users to comment on the shared information. Hence, the main goal of sharing is broadcasting of information to the general public or a specific group of people.

A second type of SNS participation is collaboration. Collaboration is defined as disposition to work together in order to achieve goals (Seonghee & Boryung, 2008). Collaboration facilitates the continuous and collective effort to achieve goals, aims, and objectives among different participants by accelerating learning and building consensus. SNS can be used for collaboration between different people in different areas such as scientific.
research among scientists (Barabási et al., 2002; Newman, 2001), coordination (Cross et al., 2002), health care (Eysenbach, 2008), and connecting organization to their customers (Gupta, Armstrong, & Clayton, 2011). Collaboration has two main characteristics: 1) is goal oriented and 2) requires group effort.

Sharing and collaboration are two separate dimensions of SNS participation. Sharing and collaboration have differences in terms of goals and audience. The goal of sharing is usually awareness and it is usually a one-way communication of information to the general audience. Many people maintain public SNS profiles through which they broadcast information to the public. Collaboration is, however, reciprocal, more purposeful, and targeted at specific people. People use SNS to collaborate with others in order to achieve different goals. Many people join different SNS interest groups to exchange ideas, to get answer to their questions, to play online games with each other (also called social gaming) (Balint, Posea, Dimitriu, & Iosup, 2011; Putzke, Fischbach, Schoder, & Gloor, 2010), to generate content (Ransbotham & Kane, 2011), and even to arrange parties or political campaigns.

1.2.5. Performance

SNS influences performance at both personal and business/organizational levels. At the personal level, SNS can increase performance of individuals in their personal daily life (personal performance). However, some studies report the negative effect of SNS use on the personal lives of people (Salehan & Negahban, 2013; Zhou, Fang, Vogel, Jin, & Zhang, 2012). On the other hand, SNS has a wide magnitude of beneficial impacts such as increased social relationship, enhanced social image, social trust, personal life satisfaction and enjoyment, and civic and political participations (R. Chen & Sharma, 2013; Valenzuela et al., 2009a).
At the business/organizational level, there are multiple ways that SNS can improve job performance. SNS has become a major communications platform for businesses to establish electronic connections with customers and many companies have shown interest in adopting SNS (Vannoy & Palvia, 2010). The primary objective for enterprises to adopt SNS is the improvement of communication and collaboration among individuals and businesses through different means such as content contribution (Gunawardena et al., 2009; A. Suh, Shin, Ahuja, & Kim, 2011; Tang, Gu, & Whinston, 2012). As SNS enable businesses to be both agile and proactive in responding to customer needs and market opportunities, they have also influenced business outcomes by enhancing the likelihood of high levels of performance (Thomas, Whitman, & Viswesvaran, 2010).

As SNS become an essential component of businesses, managers are also seeking evidence that supports the positive effect of SNS on job performance and business outcomes (C. Yang, Hsu, & Tan, 2010). Employees participate in SNS in order to achieve or enhance personal and job performance, to support their products, to build and enhance their brands, and to generate new products ideas (Crawford, 2011; Lamont, 2011). Moreover, people with more SNS experience tend to have higher satisfaction and gain better outcomes (Hyllegard et al., 2011; J. H. Kim et al., 2010a).

1.3. Research Model and Hypotheses

In this study, drawing upon the motivation-participation-performance framework suggested by Roberts et al. (2006) as our overarching theory, we propose a research model to investigate why, how, and for what purpose people use SNS.

Because we identify two dimensions of SNS participation (i.e., sharing and coloration) and two levels of performance (i.e., individual and business/organizational) based on the
theoretical lens of motivation-participation-performance framework, we conceptualize them as a second-order hierarchical structure. The use of hierarchical construct allows for more theoretical parsimony and reduces model complexity; it also allows matching the level of abstraction for predictor and criterion variables (Edwards 2001). Figure 1 depicts the research model. The model suggests that motivations influence SNS participation and in turn participation leads to performance.

![Figure 1-Research Model (Essay I)](image)

Based on uses and gratifications paradigm proposed by Zhou et al. (2011), we identify three motivations for SNS use: utilitarian, hedonic, and social. Further, social motivation is divided into two different sub-motivations: vertical and horizontal. We also identify two primary ways that SNS users participate in SNS (i.e., sharing and collaboration). Although sharing and collaboration are distinct constructs, they are interdependent. In order words, collaborations may happen through sharing information with others while sharing may also lead to collaboration with others though feedback and exchange of ideas. Hence, one can say SNS participation is reflected in sharing and collaboration. As a result, in this study, we model participation as a
second order reflective construct with two sub-constructs: sharing and collaboration (Freeze & Raschke, 2007).

In addition, we conceptualize SNS users’ individual performance of SNS use as two sub-elements: job performance at work and performance in personal activities. Although these two sub-elements are related to each other, they do not necessarily co-vary. In many cases, choosing between a reflective and a formative indicator may not be an easy task because the directionality of the relationship is not always straightforward. When indicators could be viewed as causing rather than being caused by the latent variable measured by the indicators, the indicators are operationalized by formative means (MacCallum & Browne, 1993). Because individual performance of SNS can be measured as a combinations of two different types of SNS use (i.e., for personal and/or job purpose), we handle individual performance of SNS use as a second-order formative construct, so the direction of causality is from sub-elements to a higher-construct (i.e., formative) and does not qualify for reflective modeling (Edwards, 2001).

1.3.1. Motivations behind Participation

With the growing popularity of media sharing online, SNS enable people to share various kinds of individual information with a variety of groups (Stefanone et al., 2011). Many people use SNS to connect to their strong ties, people with whom they already have an established offline connection, such as family and friends (Valenzuela et al., 2009a). Sharing photos and status messages helps many individuals to keep close emotional proximity with close friends and family especially when the physical distance between them is considerable (Stefanone et al., 2011).

Sharing is not the only way people use to connect to their strong ties. There are several examples of collaborating with strong ties on SNS. When people join Facebook, their
connections are able to suggest that they add people that they may know. When two users have many mutual friends, such collaboration will be more effective. The collaboration can also be in the form of introducing the strong ties to other people with whom they share common interests. People also use SNS to arrange events. Many people now use Facebook to invite friends and family to their parties. SNS is not only used for inviting people, but also can be used for coordination. Many people also play games on SNS. Some SNS games will provide a bonus to the players if they invite their friends. Even some users start playing a game just to help their close friends to make progress in that game.

The motivation to maintain or strengthen close personal relationships is called vertical social motivation. As a result, we expect vertical social motivation to be an important determinant of SNS participation. Consequently, we propose the following hypothesis:

H1: A user’s vertical social motivation will have a positive impact on his/her SNS participation.

SNS have been recognized as an important media affecting information sharing activities such as making profiles, commenting, messaging, blogging, file sharing, and instant messaging among organizational members (Chai & Kim, 2011). SNS support the horizontal social linkages and encourage information sharing among people from different social environments based on shared interests, political views, or activities (Boyd & Ellison, 2008). For instance, people use SNS to share knowledge with their colleagues (Ardichvili, Page, & Wentling, 2003). It is very likely that a user may find a good article shared by one of his/her connections on Facebook. The act of sharing may lead to receiving feedback from peers, trigger a conversation between people, and lead to increased social interaction. Thus, people may share information on SNS in order to strengthen their horizontal social ties. Even politicians use SNS to share the latest information
with people. Many politicians have Twitter and Facebook accounts through which they communicate with people. Therefore, the degree of horizontal interaction would influence the sharing behavior of SNS users (Chai & Kim, 2011).

In addition to sharing, collaboration with horizontal ties is also an important motivation behind SNS participation. Using SNS, more than a billion individuals around the world have created a new world of collaboration (C. M. K. Cheung, Chiu, & Lee, 2011). SNS users generate content within SNS which can be utilized as a collaborative resource for information and knowledge (C. M. K. Cheung et al., 2011). People use SNS to collaborate with whom they share common interests with. Collaboration enables horizontal ties to encourage the continuous discussion and modification of goals and objectives (Byrd & Jasny, 2010). LinkedIn creates a platform of collaboration with loose connections. People use LinkedIn to receive updated information regarding the career of the people they know. People also use LinkedIn to be introduced to other people through shared connections. If you want to connect to someone on LinkedIn with whom you have a shared connection, you can ask your shared connection to introduce you to the other person. This kind of collaboration can benefit all the parties engaged. Two people get to know each other and create a new tie. The shared connection may also benefit through the social interaction with both parties engaged. This kind of collaboration helps people expand their social network and improve their social interactions. LinkedIn also provides interest groups through which people can interact and collaborate on mutual topics of interest. As a result, the horizontal social motivation is an important predictor of SNS participation. The above arguments lead to the following hypothesis:

H2: A user’s horizontal social motivation will have a positive impact on his/her SNS participation.
Hedonic motivation tends to encompass the values derived from the SNS user’s experiential benefits and pleasure-related elements, which sets it apart from productivity-related utilitarian motivation (Gu et al., 2010). Hedonic motivation seeks self-fulfilling value and enjoyment in information systems (van der Heijden, 2004). Sharing on SNS can be quite fun. Sharing photos, links, and video on SNS is usually accompanied by feedback from other users in the form of liking and commenting. In times even the feedback takes the form of a small conversation about relevant or irrelevant topics. Moreover, sharing is a means for self-presentation and impression management (Boyd & Ellison, 2008). Sharing selfies, a type of self-portrait photograph, on SNS is now an important way for people to manage the impression of others toward themselves. The most retweeted message over Twitter was a picture of a group of actors and actresses taken in the 2014 Academy Awards shared on Twitter (Fallon, 2014). This indicates the strength of the hedonic aspect of SNS. Even politicians have started taking and sharing selfies during political events (Larson, 2013; Weinthal, 2014). As a result, sharing behavior can very well fulfill the hedonic motivation of users.

Hedonic motivations not only lead to sharing, but also they often lead to collaboration on SNS. Currently, many SNS groups serve to fulfill the common hedonic interests of their users. A very prominent example is the increasingly popular concept of online social gaming (Balint et al., 2011; Putzke et al., 2010). Gamers use SNS to connect and collaborate with others interested in the same game. In this case, collaboration can be in various forms such as playing online games with other members and informational exchange in groups and forums. It is understandable that such collaboration is mostly for enjoyment and people don’t expect any explicit utility from it. The above arguments lead to the following hypothesis:

H3: A user’s hedonic motivation will have a positive impact on his/her SNS participation.
In addition to hedonic motivation, many users use SNS because of their utilitarian motivation. SNS users are information consumers as well as information producers (Hyllegard et al., 2011). For example, many Wikipedia users are both consumer and producer of information. The objective of these users to participate on Wikipedia is to share utilitarian information with others (Ransbotham & Kane, 2011). As a result, utilitarian motivation may facilitate SNS users’ sharing of information, knowledge, documentation, file, and user generated content (Lamont, 2011).

While the desire behind SNS users’ utilitarian motivation is to increase the users’ task performance, it may have a positive impact on forming collaboration on SNS by producing information and responding to other people’s information and knowledge needs (Gregg, 2010; van der Heijden, 2004). For example, many Wikipedia pages are the result of the joint collaboration of many authors. The authors discuss different viewpoints in Wikipedia talk pages and finally reach at a conclusion which is later displayed in the form of a web page (Wikipedia, 2013). Scientific outputs of scientists may be exchanged through SNS for collaboration in order to promote the achievements and to benefit from those of others (De Roure, Goble, & Stevens, 2009). Software developers use SNS to collaborate on open source software development projects (Madey et al., 2005). SNS are also used by health professionals and health consumers to collaborate on health and medical issues (Godes & Mayzlin, 2004; Kordzadeh & Warren, 2012). It appears that SNS users’ utilitarian motivation is a strong predictor of their participation. Keeping this in mind, the following hypothesis is proposed:

**H4**: A user’s utilitarian motivation will have a positive impact on his/her SNS participation.
1.3.2. The Effect of Participation on Performance

One important purpose of using SNS is to actively share information and content with other users who have similar interests (Bulmer & DiMauro, 2009; Kwon & Wen, 2010). SNS can improve the personal performance of users’ personal lives in several ways. SNS help people to maintain the current relationships, make them stronger, and build new ones. People can gain trust and psychological stability by sharing information about their daily lives with like-minded people (Boyd & Ellison, 2008; Du, Xuan, & Wu, 2010). People can also achieve higher levels of personal productivity by seeking and sharing information and knowledge on SNS (Lamont, 2011). Moreover, SNS are good places for finding knowledge experts in related fields. Finally, people can communicate with others about their personal issues and receive solutions or empathy. Many users share status messages on a daily basis about their feelings, experiences, and needs. The feedback from other users may help them achieve a solution for their problems or help them accepts the things as they are.

The effect of SNS sharing on individuals’ lives is not limited to personal performance. SNS can also improve the job performance of individuals. SNS promote the sharing of knowledge by decreasing the amount of effort required to find information and communicate with other people (J. H. Kim et al., 2010a). Continuous information sharing with other SNS users in the work environment improves knowledge accumulation and organizational intelligence (Du et al., 2010). Moreover, SNS knowledge sharing among organizational members increases the quality of intellectual works, which in turn has a positive influence on job performance (Chai & Kim, 2011). Increased productivity can be achieved by sharing information in different forms such as sharing work-relevant photos, videos, and notes; joining a wide variety of groups; applying task-related information; and exchanging electronic documents (Boyd & Ellison, 2008;
J. H. Kim et al., 2010a). Therefore, information and knowledge sharing can increase individual performance.

SNS collaboration may also improve the performance of individuals. An important goal of SNS is collaboration based on relationships (K.-Y. Lin & Lu, 2011). Collaboration between SNS users reduces psychological distance, which in turn can impact the user’s outcome and improve it (Brown, Broderick, & Lee, 2007). Such collaboration may occur inside the organization with employees, outside of it with customers and partners (Lamont, 2011), or with friends and family. SNS users’ collaboration through SNS might impact their personal performance by effectively finding out the latest information through mutual interactions and improving social human relationships (Bulmer & DiMauro, 2009; K.-Y. Lin & Lu, 2011; Ransbotham & Kane, 2011). For example, people who discuss social issues on SNS will have a better understanding of the issues by listening to different viewpoints on the same topic. Such understanding may help them in their personal lives and when dealing with different people.

In addition to improved personal performance, SNS collaboration may also lead to improved job performance. SNS provide employees with several types of collaboration via blogging and online work-related interactions (K.-Y. Lin & Lu, 2011). SNS allow different forms of collaboration such as intra-organization (within the organization), inter-organization (with other organizations) (Sanders, 2007), and viral marketing (with customers) (Leskovec, Adamic, & Huberman, 2007). In terms of intra-organization activities, SNS users’ collaboration can have a remarkable influence on job productivity, allowing employees to get answers to work-related questions, as well as to build task-related connections (Bulmer & DiMauro, 2009; Ransbotham & Kane, 2011). From an inter-organizational standpoint, job performance is improved by SNS collaboration both directly through enhanced relationships with other
organizations (Gilbert, 2002) and indirectly through improved intra-organizational collaboration (Sanders, 2007). Finally, SNS collaboration can improve relationships with customers through enhanced word of mouth (De Bruyn & Lilien, 2008). Considering the above arguments, the following hypothesis is proposed:

H5: The extent of a user’s participation will have a positive impact on his/her individual performance.

1.4. Research Methodology and Data Collection

We used survey methods to empirically test the proposed theoretical relationships.

1.4.1. Measures

Multi-items with a 5-point Likert scale with anchors ranging from strongly disagree to strongly agree were used to measure each construct. To increase validity of the measures, we adopted most measurement items of each construct from previous tested measures in existing literature. The utilitarian motivation measures were derived from Strahilevitz and Myers (1998), Zeithaml (1988), Gu et al. (2010), and van der Heijden (2004). The items used to measure hedonic motivation were adopted from the studies by Agarwal, Sambamurthy, and Stair (2000), Boyd and Ellison (2008), and van der Heijden (2004). The items for measuring social motivations was inspired by the measurement scales of Boyd and Ellison (2008), Valenzuela et al. (2009a), and Stefanone et al. (2011) for vertical social motivation and De Roure et al. (2009), Byrd and Jasny (2010), and Rubin (1981) for horizontal social motivation. Sharing activity of SNS participation was assessed by Lévy (1997), Solomon and Schrum (2007), and Du et al. (2010). The items used to assess collaboration were adopted from Brown et al. (2007), Sanders (2007), and K.-Y. Lin and Lu (2011). Finally, performance is inspired by the studies of Miller

To improve the validity of research instruments, we asked a panel of experts to review the instruments for any issues on the clarity and the validity of measures. The panel members consisted of several professionals including university professors, researchers, managers of business firms, and members of public organizations. Based on the feedback, some measures were dropped and added; wordings of some items were revised to fit into the research context. After necessary changes were made, a pilot-test was conducted to further refine the survey before collecting data for the field test. The questionnaire was distributed to 50 people who used SNS before. The reliability and validity of measures were assessed using the data collected from the pilot-test; we confirmed that there were no validity issues on the measures. The final measurement items are summarized in Appendix B.

1.4.2. Data Collection

For the data collection of the field test, email surveys were sent over three-month period. Email addresses for the email survey were obtained from the users of Cyworld.com, Facebook.com, Twitter.com, and other social networking services. Cyworld.com is a Korean SNS which has features similar to Facebook. A total of 4,100 survey questionnaires were distributed and the respondents were asked to report any social networking services they had used before. A total of 789 questionnaires were collected (response rate 19.2%). However, 78 invalid samples were excluded – i.e., 41 incomplete responses and 37 respondents who reported they did not have any experiences with social networking services.

The sample shows that SNS users are generally in late 20s and 30s who graduated from a university. About 39% of the respondents said that they use Cyworld, followed by 25.7% for
Facebook, and 19.8% for Twitter; 28% of the respondents visit SNS three times a day or more; 27% of the subjects used SNS twice a day and 34.8% used it once a day. Regarding the user trend, 26.3% were early adopters and 73.7% were later adopters; 63% of the subjects have used SNS for longer than two years. The demographic characteristics of sample are summarized in Appendix C.

For the validation of the proposed model, this study used only 353 respondents who reported they had used Facebook-style SNS (i.e., Facebook and Cyworld) before, because using all diversified types of SNS into a single study it is not applicable to generalize the study results. To assess non-response bias of samples, we compared the data from the last three weeks with that of the first four weeks by conducting a t-test using the level of education and age. The results, combined with the representativeness of the sample, identify reasonable evidence that response bias was not an issue.

1.5. Data Analysis and Results

Structural equation modeling (SEM) technique was used to analyze the data for both the measurement model and the structural model. Compared to a conventional regression analysis that ignores the interrelationships between latent constructs measured by multiple measurement items (Bollen, 1986; Chin, 1998b), SEM is a statistical methodology that takes a confirmatory (i.e., hypothesis-testing) approach to the analysis of causal relationships among latent construct (i.e., a structural theory) (Byrne, 2001). There are two families of SEM techniques: covariance-based techniques (e.g., AMOS) and variance-based techniques (e.g., Partial Least Squares). In this study, we use Partial Least Squares using SmartPLS version 2.0.M3 to test the measurement model and the structural model. PLS analysis was chosen over other analytical techniques for two reasons. First, PLS simplifies the modeling of formative and reflective constructs (Chin,
1995, 1998a) and makes handling of second-order constructs easy (Wetzels, Odekerken-Schröder, & Van Oppen, 2009). Second, it simultaneously tests both the measurement model and the structural model (Wixom & Watson, 2001). It also reports composite reliability (CR) and average variance extracted (AVE) for content validity and discriminant validity.

1.5.1. Measurement Model Testing

To assess the psychometric properties of the instrument, it was tested for reliability and validity of measurement model before the structural model testing. In terms of reliability assessment, formative and reflective constructs should be treated differently (Wixom & Watson, 2001). For the reflective constructs, the assessment of the measurement model includes the estimation of internal consistency for reliability, convergent validity, and discriminant validity (Kabir, 2013). Formative measures don’t need the above assessments because they neither correlate with one another nor exhibit internal consistency (Chin, 1998b). Analysis of reflective constructs show that all items load significantly on their corresponding constructs (Gefen & Straub, 2005) and have a high loading on their corresponding constructs which is higher than the cutoff point of 0.4 (Bickart & Schindler, 2001). Appendix D shows the result of the exploratory factor analysis.

Cronbach’s alpha and Fornell’s composite reliability (Fornell & Larcker, 1981) are used to test the internal consistency of the measurement models. The values of the Cronbach reliability coefficients range from 0.816 to 0.931, which are higher than the minimum cutoff score of 0.7 (Nunnally & Bernstein, 1994). Composite reliability should be greater than the benchmark of 0.7 to be considered adequate (Fornell & Larcker, 1981); the lowest composite reliability is 0.816, which is higher than 0.7, indicating adequate internal consistency. All AVEs are higher than 0.5 (Fornell & Larcker, 1981). The pattern of loadings and cross-loadings
supported internal consistency and discriminant validity. Table 1 shows the summarized reliability indices.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>No. of Items</th>
<th>Mean</th>
<th>S.D.</th>
<th>Cronbach Alpha</th>
<th>Composite Reliability</th>
<th>AVE</th>
<th>Concept &amp; Scales adapted from</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical Social Motivation</td>
<td>5</td>
<td>3.56</td>
<td>.806</td>
<td>0.845</td>
<td>0.890</td>
<td>0.619</td>
<td>Boyd and Ellison (2008), Valenzuela, Park, and Kee (2009b), Stefanone et al. (2011)</td>
</tr>
<tr>
<td>Horizontal Social Motivation</td>
<td>4</td>
<td>3.175</td>
<td>.845</td>
<td>0.879</td>
<td>0.917</td>
<td>0.735</td>
<td>De Roure et al. (2009), Byrd and Jasny (2010), Chai and Kim (2011), C. M. K. Cheung et al. (2011)</td>
</tr>
<tr>
<td>Hedonic Motivation</td>
<td>4</td>
<td>3.244</td>
<td>1.063</td>
<td>0.816</td>
<td>0.878</td>
<td>0.644</td>
<td>Gu et al. (2010), van der Heijden (2004)</td>
</tr>
<tr>
<td>Utilitarian Motivation</td>
<td>4</td>
<td>3.334</td>
<td>.912</td>
<td>0.846</td>
<td>0.897</td>
<td>0.685</td>
<td>Gu et al. (2010), van der Heijden (2004)</td>
</tr>
<tr>
<td>Collective Sharing</td>
<td>6</td>
<td>3.419</td>
<td>.804</td>
<td>0.879</td>
<td>0.909</td>
<td>0.624</td>
<td>J. H. Kim, Kim, and Nam (2010b), Du et al. (2010), Chai and Kim (2011)</td>
</tr>
<tr>
<td>Collective Collaboration</td>
<td>6</td>
<td>3.253</td>
<td>.831</td>
<td>0.894</td>
<td>0.919</td>
<td>0.653</td>
<td>Brown et al. (2007), Sanders (2007), K.-Y. Lin and Lu (2011)</td>
</tr>
<tr>
<td>Personal Performance</td>
<td>6</td>
<td>3.400</td>
<td>.852</td>
<td>0.895</td>
<td>0.919</td>
<td>0.655</td>
<td>Viswesvaran et al. (1999), Wright and Cropanzano (2000)</td>
</tr>
<tr>
<td>Work Performance</td>
<td>7</td>
<td>3.196</td>
<td>.933</td>
<td>0.913</td>
<td>0.931</td>
<td>0.659</td>
<td>Viswesvaran et al. (1999), Wright and Cropanzano (2000)</td>
</tr>
</tbody>
</table>

In order to examine discriminant validity of the constructs, we used the procedure proposed by Fornell and Larcker (1981). This procedure recommends comparing the average variance extracted (AVE) to the variance shared between the construct and other constructs. The values of AVE for all constructs were above 0.50 and the root square of AVE was higher than the correlation of the corresponding construct with other constructs (Fornell & Larcker, 1981) indicating adequate discriminant validity. Table 2 shows the result of discriminant validity analysis.

29
Because survey methodologies may be subject to common method bias (CMB), we ran a PLS test for CMB using the common factor approach described by Liang et al. (2007). We created a common method construct having all the items associated with it; we then modeled each of the 40 indicators as a single-indicator construct and created paths between them and the common method construct as well as the theoretical constructs. Appendix E summarizes the results of the CMB analysis. The results showed that loadings on the theoretical constructs were both high and highly significant. Loadings on the common method construct were low and in most all cases non-significant. This indicates that CMB is not a problem in this research (Liang et al., 2007).

1.5.2. Structural Model Testing and Results

Analysis of the structural model was also done using SmatPLS. Figure 2 summarizes the results of structural model testing. The analysis shows that vertical social motivation positively affects participation ($\beta = 0.24, p < 0.001$), thus supporting H1. Horizontal social motivation significantly affects participation ($\beta = 0.19, p < 0.001$), supporting H2. However, the
hypothesized relationship between hedonic motivation and participation was not significant ($\beta = 0.06$). Thus, H3 is not supported. The analysis also shows that utilitarian motivation is a significant predictor of participation ($\beta = 0.74$, $p < 0.001$), thus supporting H4. The model explains 50.0% of the variation in participation on SNS. The results also show that participation significantly affects individual performance ($\beta = 0.72$, $p < 0.001$), supporting H5. The model explains 52.1% of variance in individual performance. Table 3 summarizes the hypothesis testing results.

![Figure 2-PLS Analysis Results (Essay I)](image)

**Table 3-Hypothesis Testing Results (Essay I)**

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Hypothesized Relationship</th>
<th>Estimates (t-value)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Vertical Social Motivation → Participation</td>
<td>0.24 (5.03)</td>
<td>Supported***</td>
</tr>
<tr>
<td>H2</td>
<td>Horizontal Social Motivation → Participation</td>
<td>0.19 (3.91)</td>
<td>Supported***</td>
</tr>
<tr>
<td>H3</td>
<td>Hedonic Motivation → Participation</td>
<td>0.06 (1.54)</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H4</td>
<td>Utilitarian Motivation → Participation</td>
<td>0.42 (11.21)</td>
<td>Supported***</td>
</tr>
<tr>
<td>H5</td>
<td>Participation → Individual Performance</td>
<td>0.72 (29.25)</td>
<td>Supported***</td>
</tr>
</tbody>
</table>

* Significant at 0.05 level, ** significant at 0.01 level, and significant at *** 0.001 level
1.5.3. Post-hoc Analyses

Because hedonic motivation does not show significant effect on participation, we conducted a post-hoc analysis using a structure model with two separated first-order dimensions of participation (i.e., sharing and collaboration) to further examine if there is any difference at the first-order level. The results confirm that hedonic motivation significantly predicts sharing participation but not collaboration. In contrast, utilitarian motivation is a more important predictor of SNS use than hedonic motivation. Figure 3 shows the results of the post-hoc analysis.

![Figure 3: Post-hoc Analysis Results (Essay I)](image)

In addition, because it is important to check the how much participation mediated the effects from motivations to performance, we also conducted a set of Sobel tests (Sobel, 1982) to examine whether the second-order participation fully mediates the effects of the four motivations (i.e., independent constructs) on the second-order performance (i.e., dependent constructs) of the proposed research model. For the comparison of mediating effects of the first-order and the second-order participation constructs, we also tested the effects of sharing and collaboration as
first-order constructs mediating the relationship between the four motivations and performance. Table 4 summarizes the results of the Sobel Z tests and Appendix F shows detailed results of the Sobel Z tests.

<table>
<thead>
<tr>
<th>Sharing</th>
<th>Collaboration</th>
<th>Participation (second-order)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobel Z</td>
<td>Mediation</td>
<td>Sobel Z</td>
</tr>
<tr>
<td>Vertical</td>
<td>4.44*** Partial</td>
<td>3.85*** Partial</td>
</tr>
<tr>
<td>Horizontal</td>
<td>4.69*** Partial</td>
<td>3.74*** Partial</td>
</tr>
<tr>
<td>Hedonic</td>
<td>3.36*** Full</td>
<td>2.42* Full</td>
</tr>
<tr>
<td>Utilitarian</td>
<td>4.08*** Partial</td>
<td>3.78*** Partial</td>
</tr>
</tbody>
</table>

The results show that sharing and collaboration partially mediate all relationships between the motivations and performance except hedonic motivation which is fully mediated by both sharing and collaboration. Later we tested the mediation effect of the second-order construct (i.e., participation). The result shows that participation, the second-order construct, fully mediates the relationship between motivations and performance. The only exception is utilitarian motivation which is partially mediated by participation. One plausible explanation of this result is that for SNS users with high utilitarian motivation, the level of SNS participation is less important and even low levels of SNS participation may significantly increase their performance. Such people are focused on performance and use their time on SNS wisely rather than wasting it for hedonic purposes. Hence, there is also a direct relationship between motivation and performance. After comparing the mediating effects of the first-order and the second-order participation models, we can conclude that the second-order model better fits to the motivation-performance theory.
1.6. Discussion

There are several findings. The first finding is a possible answer to our first research question which is about how different motivations lead to the participation of social networking services. We have identified four motivations of SNS use: utilitarian, hedonics, vertical social and horizontal social motivations. Most motivations except hedonic significantly explain why SNS users participate in SNS. The finding supports that people use SNS based on different motivations and expectations. Some people use SNS to get in touch with their close friends, relatives and family (i.e., for vertical social motivation). Some others are more interested in connecting their weak ties and increasing the size of their social network (i.e., for horizontal social motivation). Some users may be interested in hedonic aspects of SNS which allows them to enjoy sharing information with their social networking group. For other people, SNS is a tool that helps them increase their performance. It is noteworthy that SNS users possibly have multiple motivations to use SNS at the same time. Some users may use SNS not only because of utilitarian purpose but also social and hedonic purposes.

Although previous research reported that hedonic motivation is an important predictor of SNS use (Zhou et al., 2012), interestingly, the results did not confirm the significant effect of hedonic motivation on the second-order participation. Because it is an interesting but unexpected finding, we further investigated why this happened using a post-hoc analysis. The analysis result partially supported the effect of hedonic motivation on participation: i.e. hedonic motivation strongly leads to sharing but not collaboration. There are two possible interpretations of this unexpected result. Xu et al. (2012) presented that although hedonic motivation is important in the initial stages of SNS adoption, its importance decreases as time passes. Because almost 80% of respondents have used SNS for more than one year, the effect of hedonic motivation is possibly
diminished. Another possible interpretation is that sharing features for hedonic purpose are well equipped; collaboration for hedonic purpose (e.g., collaboration for gaming) is a relatively new in Facebook-style SNS. It seems that many SNS users are not yet familiar with hedonic forms of SNS collaboration; they mainly work together in SNS for utilitarian purposes rather than hedonic.

Another finding is related to the second research question; how users’ participation in SNS affects their personal life and job performance. The results show that when users actively participate in SNS, they perceive higher performance in terms of personal life and work. Participation in SNS can help people improve their personal and work life in several ways such as improving social relationship, receiving up-to-date information, effective completion of personal tasks, and finding experts in areas of interest.

1.6.1. Contributions

This study contributes to both theory and practice. First, this study contributes to the online social media literature by proposing a theoretical model that describes why, how, and for what purpose people use social networking services. Using the theoretical lens provided by Roberts et al. (2006) as an overarching theory in the context of Facebook-style online social networking services, this study provides a more holistic view of how SNS users’ different motivations lead to their participation and how their participation influences their personal performance as well as job performance.

Second, this study identifies different motivations as important predictors of SNS participation, including utilitarian, hedonic, and social motivation. Further, we classify social motivation into two types (i.e., vertical social motivation and horizontal social motivation). Although previous studies (e.g., J. H. Kim et al. (2010a)) show that social motivation is an
important predictor of SNS adoption, it is the first time that social motivations has been categorized into two different types. We believe this is one of the unique contributions of the study.

Third, this study conceptualizes SNS participation as a reflective second-order construct by identifying sharing and collaboration as two major participation methods in Facebook-style SNS. We also introduce another second-order construct, individual performance, which is composed of two sub-dimensions: personal and job performance. While previous studies investigate the effect of SNS use on either personal performance or organizational performance separately, this study proposes a second-order formative construct (i.e., performance) combining personal performance and job performance as two first-order dimensions. We argue that this provides a broader view of SNS performance. Moreover, we further validate the new second-order hierarchical constructs using the empirical data collected from respondents who have used Facebook-style SNS. As Edwards (2001) argued, we confirm the model with hierarchical constructs fits well to the motivation-participation-performance framework in terms of theoretical parsimony and reduced model complexity. The new construct has been validated in this study and can be used by future studies. Finally, this study shows that the ultimate goal of users to participate in SNS is to enhance their personal and job performance through SNS. Our findings suggest that SNS participation significantly influences performance of individuals.

This study also contributes to practice in several ways. First, the results show that people use the Facebook-style SNS not just for one single reason. Because users’ needs and their motivations are diverse, for SNS providers to retain their users, they need to develop various new services beyond the bounds of a simple personal networking platform. Especially, the results show that utilitarian motivation is a more important predictor of SNS use than hedonic
motivation. As people are demanding a more personalized multipurpose SNS, to improve their utilitarian value, SNS providers may offer a SNS platform with a diverse selection of applications and services across multiple devices. To provide a successful SNS platform, it is recommended for SNS providers to build an ecosystem by opening platform for third-party application developers.

In addition, because SNS users share information and collaborate with others to gain benefits not only for personal but also work-related reasons such as improved skills and knowledge, improved business communications, and ultimately improved job performance, organizations can gain benefits (i.e., improved business performance) through sharing and collaboration of their employees on SNS. Thus, we recommend managers of some organizations that have a policy to block SNS access in the workplace to reevaluate their policy. SNS may provide a potential opportunity for organizations to outperform competitors through innovative ways of sharing and collaboration among employees.

1.6.2. Limitations and Future Directions

Although this study provides some interesting results and insights, it is not without limitations. One limitation is that because this study mainly focuses on Facebook-style SNS, the findings are limited to this context. As we discussed, there exist different types of SNS. Thus, one should be careful to generalize the results beyond the scope of the research context. Future research could expand upon the current findings using a different type of SNS. Second, the empirical data collected for this study were from individuals in a single country and therefore it has limited cultural diversity. Future research with a more diverse sample may provide better insights regarding different motivations to use SNS. Furthermore, like other survey-based studies, this study uses self-reported data which is subject to social desirability response bias.
(Arnold & Feldman, 1981). With development of more objective measures of SNS performance, future research should further expand the research boundary of performance measures of SNS. Additionally, while our research shows different motivations for SNS use, it does not show motivations for specific sharing or collaboration actions (e.g., sharing very personal photos, tagging friends’ photos, twitting a specific message, etc.). It would be an interesting future study to investigate this issue by examining which motivations lead to what specific actions.

1.7. Conclusions

Despite the above limitations, this paper makes important contributions to both theory and practice. We provide a theory-grounded and holistic view of how motivations lead to participation and how participations lead to performance of SNS users at the individual level of analysis. We identify four different motivations to use SNS and to different types of SNS participation which lead to improved performance of individuals in both their personal life and their work. Thus, this study provides several insights that extend our understanding of social networking services.
ESSAY II: WHAT MAKES A USEFUL ONLINE CONSUMER REVIEW? EMPIRICAL EVIDENCE FROM SENTIMENT MINING OF AMAZON PRODUCT REVIEWS

2.1. Introduction

Businesses are now using social media to promote their products and services. Many companies maintain Facebook and Twitter accounts to keep in touch with their customers. Customers also use social media to receive information about products/services. In many ways the Internet in general and social media in particular, have changed the way customers shop for goods and services. It’s now quite normal for people to enter brick-and-mortar stores, find the product they want, and then order it online. Moreover, online consumer reviews (OCR) have helped customers to learn about the strengths and weaknesses of different products and to find the ones that best suit their needs. Some studies suggest that customers show more interest to user-generated product information on the Internet than vendor information (Bickart & Schindler, 2001). A recent study shows that OCR are the second most trusted source of product information after recommendations from family and friends (Nielsen, 2012). Compared to vendor generated product descriptions, OCR are more user-oriented and describe the product in terms of different usage scenarios and assess it from the user’s perspective (Y. Chen & Xie, 2008). Thus, it has even been suggested that consumers who write OCR serve as “sales assistants” for online retailers (Y. Chen & Xie, 2008).

The process of analyzing OCR can be broken into two steps: the decision to read the review, and the actual processing of the information in the review that deems the decision to use it based on perceived helpfulness of the review (Ahluwalia, 2000). Focusing on the two steps of processing OCR, this study proposes two fundamental research questions:

RQ1: Which factors determine the likelihood of a user paying attention to a review?
RQ2: Which factors determine the perceived helpfulness of a review?

The first research question is about the characteristics of OCR that absorb the attention of online consumers. Many products receive too many reviews that make it difficult for consumers to read all of them. Thus, most consumers have to read them selectively. In summary, the first research question explores the determinants of the readership of OCR. Although reading a review is the first step in determining its helpfulness, previous research overlooks the readership of OCR.

While the first research question remains largely unexplored, the second one has received some attention. Previous research has found strong evidence that OCR influence product sales (Chevalier & Mayzlin, 2006; Dellarocas et al., 2007; Duan, Gu, & Whinston, 2008; Forman, Ghose, & Wiesenfeld, 2008; Ghose & Ipeirotis, 2011; Godes & Mayzlin, 2004; Liu, 2006; Mudambi & Schuff, 2010). However, three important aspects of the second research question require further investigation. First, most studies only use the numerical review ratings (e.g., the number of stars) and the length of the reviews in their empirical analysis, without formally incorporating the information contained in the text of the reviews. Therefore, a deeper analysis of the textual information contained in the OCR can provide greater insight into what constitutes a useful online review (Mudambi & Schuff, 2010).

Second, previous research shows that positive statements are considered to be more helpful by consumers (Schindler & Bickart, 2012). However, research has produced conflicting results regarding the negative comments. For example, Sen and Lerman (2007) found that in the case of utilitarian products, negative comments are more useful than positive ones, whereas Schindler and Bickart (2012) failed to find any significant relationship between negative statements and helpfulness of a review. By investigating the effect of sentiment polarity (i.e.,
negative, positive, or neutral) on helpfulness, this study provides insights into the performance of OCR.

Finally, previous research tends to identify the predictors of the performance of OCR without providing many practical solutions for online vendors. Human subjects are largely employed for categorization of OCR based on the textual information contained in them. Although this method provides a descriptive view of the performance of OCR, it does not facilitate the development of scalable automated systems for classification of OCR.

This study investigates the performance of OCR through analysis of textual information contained in the reviews. Previous research has indicated that computer-mediated communications (CMC) can effectively transfer emotions. For example, the emotions contained in a message transferred through CMC significantly influence how the message is processed and interpreted by the receiver (Berger & Milkman, 2012; Riordan & Kreuz, 2010; Walther & D’Addario, 2001). Sentiment mining can be used for emotional analysis of textual information. Sentiment mining refers to the use of natural language processing and computational linguistics to find and extract subjective information from text data. Sentiment mining is usually done using automated tools that provides benefits such as scalability, effective information retrieval, automated cyber risk management, and increased business profits (Bai, 2011). It also facilitates processing of large amounts of data. Social media is an important source of big data and is quite suitable for text mining purposes (Russell, 2013).

This study contributes to the existing body of knowledge in three unique ways. First, it provides a research model that predicts the performance of OCR in terms of readership and helpfulness of reviews. Second, it uses automated tools to analyze a set of secondary OCR data composed of reviews collected from the Amazon.com website. Finally, it provides insights to
online business managers regarding the design and implementation of scalable automated systems to improve classification and sorting of OCR, which will eventually help them achieve increased sales.

2.2. Theoretical Background

2.2.1. Online Review Performance Measures

Different measures have been used to evaluate the performance of OCR. However, most use helpfulness as the single performance measure of OCR (Aral, 2013; Mudambi & Schuff, 2010; Sen & Lerman, 2007). Helpfulness has also been referred to as the value of the review (Schindler & Bickart, 2012). For studies that use secondary data, helpfulness is measured by dividing the number of people who find a review helpful by the total number of people who voted for that review (Mudambi & Schuff, 2010; Sen & Lerman, 2007).

Purchase intention is another measure of performance. Customer purchase intention is influenced by both quantity and quality of the reviews (Park, Lee, & Han, 2007). Some studies use sales revenue as a performance measure for online reviews. One study used online reviews from the Yahoo Movies website to predict box office revenues and found strong evidence that online reviews influence movie sales (Liu, 2006). Online reviews have also been used to predict the online sales of books (Chevalier & Mayzlin, 2006). Product ratings also indirectly influence sales through sentiment (Hu, Koh, & Reddy, 2013).

2.2.2. Predictors of Online Review Performance

Different measures have been used to predict the performance of OCR. Some studies have focused on the numeric star rating and word count of the reviews to predict their performance. For example, extreme numerical ratings are positively related to sales of books (Chevalier & Mayzlin, 2006). Reviews with extreme numerical ratings are also considered more
helpful (Mudambi & Schuff, 2010). Length of a review may also predict its performance. Length of reviews for a book significantly predict its sales on Amazon.com (Chevalier & Mayzlin, 2006). Length of a review is also positively related to its helpfulness (Mudambi & Schuff, 2010; Schindler & Bickart, 2012). Empirical analysis, however, has failed to find any significant relationship between review length and sales of books on the Barnes & Noble website (Chevalier & Mayzlin, 2006).

The difference between the performance of positive and negative reviews is a controversial research avenue in the context of OCR. Drawing upon negativity bias theory, some studies propose that negative reviews are considered more helpful than positive ones. According to negativity bias theory, people face difficulty in making inferences about the actions of an actor when the actor behaves in an expected fashion. The inference is easier when the actor departs from the norms of behavior (Jones & Davis, 1965; Kelley, 1973). Thus, some researchers have argued that negative comments should be considered more helpful than positive ones because they deviate from the accepted norm of staying positive (Sen & Lerman, 2007). Several studies have tested negativity bias theory in the context of OCR. Their findings, however, are contradictory. While some studies show that consumers find negative reviews more valuable, others find no significant difference between positive and negative reviews. Some others even find that positive reviews are more helpful than negative ones (Sen & Lerman, 2007). However, negative reviews are more useful for utilitarian products (products that focus on task performance) and positive comments are more useful for hedonic products (products that deal with pleasure). Schindler and Bickart (2012) find that the number of positive statements is a significant predictor of the value of a review while the effect of the number of negative statements on value is not significant. A recent study suggests that consumer reviews may be
subject to positive social influence bias (Aral, 2013). Through an experiment, the author manipulated the helpfulness of OCR and found that positive manipulations create a positive social influence that last several months. The negative manipulations, however, are offset by a correction effect that neutralizes the manipulation. The findings suggest that corporations can easily manipulate OCR and the reviews should be considered with some level of skepticism.

Recent studies use text of a review to predict its performance. Schindler and Bickart (2012) look at the wording of online reviews rather than their source. They divide wording factor into two categories: content and style. They define the content as the information the review provides. Style, by contrast, is defined as the choice of words used to convey the information. They find that proportion of product-descriptive statements and proportion of reviewer-descriptive statements are significant predictors of the value of the review. They also find that while use of negative style characteristics decrease the value of the review, use of positive style doesn’t improve its performance. Finally, they didn’t find any significant relationship between negative evaluative statements and value of the review.

Review sentiment is another measure used by previous studies to predict the performance of computer-mediated communications (CMC) in general, and OCR in particular. Previous research has indicated that CMC can effectively transfer emotions. The receiver of a message can detect the sender’s emotions through verbal cues such as emotion words as well as nonverbal cues such as emoticons (Harris & Paradice, 2007). Moreover, the emotions contained in a message transferred through CMC significantly influence how the message is processed and interpreted by the receiver (Riordan & Kreuz, 2010; Walther & D’Addario, 2001). Different methods can be used to extract sentiment out of text. While some studies use human subjects to
extract the sentiment of OCR, others use automated sentiment mining to extract sentiment from the text of reviews (Bai, 2011; Schindler & Bickart, 2012; Sen & Lerman, 2007).

2.2.3. Moderators of Performance

Product type is an important factor in the context of online reviews which has mostly been viewed as the variable moderating the relationship between independent variables and helpfulness. For example, reviews of hedonic products are less likely to be perceived as helpful (Sen & Lerman, 2007). There are also differences between helpfulness of search and experience goods. For search goods, consumers can obtain information about product quality prior to purchase while evaluation of experience goods requires sampling or purchase. While reviews with extremely high or low star rating are considered more helpful for search goods, the effect is different for experience goods. For experience goods, extreme reviews are considered less helpful than the ones with moderate ratings (Mudambi & Schuff, 2010). Product type also moderates the relationship between the number of positive and negative statements with the helpfulness of a review (Sen & Lerman, 2007). The length of a review has greater positive impact on its helpfulness for search goods compared to experience goods (Mudambi & Schuff, 2010).

Customer involvement is another moderating variable in the context of OCR. Low-involvement customers are more influenced by the quantity rather than the quality of reviews, yet high-involvement customers are influenced by the quantity of reviews if the reviews have high quality (Park et al., 2007).

2.2.4. Theories Related to Message Selection

The theory of selective attention posits that people respond to messages selectively because of limited information processing capacity (Treisman, 1969). When people receive
multiple stimuli at once, they need to filter some of the messages because they have limited processing capacity. The same thing happens in the context of online reviews. Many products have large number of reviews which makes people selectively read online reviews. The theory of selective perception helps explain the mechanism underlying selective attention. According to selective perception, people develop belief structures which are a simplified representation of the world and use these structures to filter and interpret information (Walsh, 1988).

Attribution theory explains how people analyze different behaviors. Based on the theory, people identify two types of explanation for different behaviors: actions related to internal (personal) factors and the ones that result from external environment (situational) (Folkes, 1988; Heider, 1958). The theory largely explains the predictors of source credibility and other areas related to consumer perceptions and inference formation (R. R. Dholakia & Sternthal, 1977). It has also been used to explain the perceptions of the consumers regarding the helpfulness of OCR. Readers consider the motivation of the author of a review when deciding about whether to use the information contained in the review. Readers may attribute a review to either external (product-related) or internal (reviewer-related) reasons. If the reader feels that the review is based on external reasons, they are more likely to accept it. If reader believes that the review is based on internal reasons, they are more likely to disregard it (Sen & Lerman, 2007).

2.2.5. Fake Reviews

Because online reviews have the potential to influence customers, they have received significant attention from sellers/providers of goods/services. Some have even tried to influence the perceptions of customers by creating fake reviews (S. Kim, 2013; Streitfeld, 2011; Walker, 2013). One study finds that certain hotels distinguished by location and ownership post more fake positive reviews for themselves and more fake negative reviews for their competitors. The
independent hotels that are owned by single-unit owners are more likely to post fake reviews than branded hotel chains. Moreover, the hotels close to a hotel with a high incentive to post fake reviews have more negative reviews (Mayzlin, Dover, & Chevalier, 2012). Another study reports that 10.3% of book reviews are fake or manipulated (Hu, Bose, Koh, & Liu, 2012).

Previous research has used different methods to identify fake reviews. Some studies compare reviews of products on different websites to investigate fake reviews. One study identifies evidence of fake reviews by comparing the reviews of two travel sites with extensive hotel reviews: Expedia.com and TripAdvisor.com (Mayzlin et al., 2012). While anyone can write a review on TripAdvisor, one needs to have actually booked at least one night for the given hotel on Expedia to be able to post a review for that hotel. Human subjects and automated tools have also been used to identify fake reviews. One study failed to recognize manipulated reviews using sentiment mining but found that consumers are able to recognize them (Hu et al., 2012). Other studies use machine learning methods to identify fake reviews (Jindal & Liu, 2008; Li, Huang, Yang, & Zhu, 2011; Ott, Choi, Cardie, & Hancock, 2011).

2.3. Research Model and Hypothesis

Following the directions provided by Mudambi and Schuff (2010) regarding the need for the analysis of the textual information contained in OCR, we utilize sentiment mining to analyze the text of the reviews and to investigate our two research questions. Each Amazon review has a title and a body text. We extract the sentiment of both the title and the text of reviews and use them to predict performance. We use two measures for OCR performance, readership and helpfulness, each of which corresponds to one of our research questions. Figure 4 shows the proposed research model. In our model, sentiment refers to the total amount of sentiment that exists in a text, both positive and negative. Polarity refers to the direction of the sentiment which
can be positive, negative, or neutral. More information on our measures can be found in the methodology section.

![Figure 4-Research Model (Essay II)](image)

We expect longevity of a review to be positively related to its readership. Older reviews have been on the website for a longer time and hence have higher chance of being viewed and read by consumers. Thus, we can expect older reviews to receive more readership than newer ones and we hypothesize the following:

H1: Longevity of a review has a positive effect on the readership of the review.

According to the theory of selective attention, people respond to messages selectively because individuals possess limited information processing capacity (Treisman, 1969). We believe that OCR are no exception to this theory. Many products/services have large number of reviews and reading all of them is very difficult and time consuming for consumers. Hence, people must pay selective attention toward OCR. We expect people to look for quick signals that enable them to decide about reading a review. Title is a small but important part of a review and
provides users with instant information about the general theme of the review. Thus, we expect the title of a review to be an important predictor of its readership. Moreover, sentiment of a message can be effectively communicated through the text and significantly influences the perceptions of the reader (Harris & Paradice, 2007; Riordan & Kreuz, 2010; Walther & D’Addario, 2001). Finally, affective language in the online environment receives more attention and feedback compared to neutral language (Huffaker, 2010). Hence, we expect the total sentiment contained in the title of a review to be positively related to its readership and we hypothesize that:

H2: The larger the total amount of sentiment (positive or negative) the title of a review exhibits, the more readership it receives.

According to the theory of selective perception, people use their mental structures to filter information. People tend to absorb information that matches their mental model and filter out the information that don’t fit their mental structure (Walsh, 1988). Similarly, we expect people to pay selective attention toward OCR based on their initial attitude about the product. Those with a positive attitude toward the product are expected to look for positive information and those who are more suspicious are expected to look for negative information. Many consumers read OCR to receive reassurance that they have made a good choice (Bailey, 2005; Goldsmith & Horowitz, 2006; Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004). Reassurance will be gained through positive information rather than negative information. Consequently, we expect a larger number of consumers look for positive reviews rather than negative ones. Thus, we expect reviews with positive titles attract more attention and readership. Considering the above arguments, we expect positive sentiment to have a larger effect on the readership of online reviews than negative sentiment. This leads to the following hypothesis:
H3: Title polarity moderates the effect of title sentiment on the readership of the review. The effect will be larger for positive titles than negative and neutral ones.

Although the total amount of sentiment in the title of a review is an important source of information for consumers, length of the title can also be an important factor related to its readership. Short titles are usually not very informative and do not contain much information. A short title usually communicates the general idea of the author about the product such as “I love it” or “not very high quality.” In contrast, longer titles will give the reader more information about the content of the review which subsequently increases the likelihood of it being read by the consumers. For example, a review entitled “the best LED-lit television at a budget price” signals the reader that it contains information about both quality and pricing of the product which can motivate more people to read it. Moreover, length of the title may decrease consumer’s search costs through increased information diagnosticity (Johnson & Payne, 1985). Hence, we expect length of the title of a review to be related to its readership performance and we hypothesize that:

H4: The length of the title of a review has a positive effect on the readership of the review.

The length of a review is an important predictor of its performance (Mudambi & Schuff, 2010; Schindler & Bickart, 2012). Short reviews are more likely to be shallow and lack the comprehensive evaluation of product features. Longer reviews, in contrast, contain more information and are more likely to contain deep analysis of the product, its features, and the context in which it was used. Longer reviews are more likely to receive attention from users. Reading longer reviews may decrease consumer’s search costs through increased information diagnosticity (Johnson & Payne, 1985). Moreover, longer reviews are more likely to be
perceived helpful. An individual’s argument is more persuasive when it provides larger amount of information (Schwenk, 1986). Increased number of reasons for a choice escalates the decision maker’s confidence (Tversky & Kahneman, 1974). Hence, we expect longer reviews to receive more readership as well as more perceptions of helpfulness compared to the shorter ones. As a result, we propose the following hypotheses:

H5: Length of a review has a positive effect on the readership of the review.

H6: Length of a review has a positive effect on the helpfulness of the review.

People tend to find reviews with extreme numerical ratings more helpful (Mudambi & Schuff, 2010). We can expect that reviews with extreme ratings also contain more sentiment because the author is either very satisfied or very unsatisfied. The extreme levels of satisfaction or dissatisfaction are very likely to turn into strong emotions and consequently strong sentiment. The sentiment of a message can be effectively communicated through the text and significantly influences the perceptions of the reader (Harris & Paradice, 2007; Riordan & Kreuz, 2010; Walther & D’Addario, 2001). One can argue that the sentiment contained in the review is the driver of the perceptions regarding its helpfulness rather than just the numerical rating. Different people have different experiences with the same product. Sentiment is the vehicle for people to convey their emotions to others through text. Thus, sentimental reviews are better conveyer of the experience with the product. Consequently, we can expect high-sentiment reviews to be perceived more helpful by the consumers because they are more likely to convey the experience with the product. Thus, we hypothesize that:

H7: The larger the total amount of sentiment (positive or negative) the text of a review exhibits, the more helpful it is perceived to be.
The decision regarding the helpfulness of a review is made after the person reads the review (Ahluwalia, 2000). Reviews with larger number of descriptive statements are considered more helpful (Schindler & Bickart, 2012). One can argue that a good description of an object contains different aspect of it including both positive and negative qualities. This implies that a good review should also contain both positive and negative statements and consequently both positive and negative sentiment. Moreover, the use of positive or negative style characteristics does not increase the value of a review and may even decrease the perceptions regarding its helpfulness (Schindler & Bickart, 2012). In terms of review polarity, this will translate into leaning toward neutral polarity. Neutral polarity does not imply that sentiment is non-existen in the text. It implies that there are balanced levels of positive and negative sentiment. Hence, we expect sentimental reviews with neutral polarity to be perceived more helpful than positive and negative ones. This leads to the following hypothesis:

H8: Review polarity moderates the effect of review sentiment on the helpfulness of the review. The effect will be larger for neutral reviews than for positive and negative ones.

2.4. Methodology

2.4.1. Measurement

Sentiment mining was done using SentiStrength software (Thelwall, Buckley, Paltoglou, Cai, & Kappas, 2010). The software is free for academic research and has been tested and validated in previous research (Garcia & Schweitzer, 2011; Gruzd, Doiron, & Mai, 2011; Stieglitz & Dang-Xuan, 2013; Thelwall & Buckley, 2013; Thelwall, Buckley, & Paltoglou, 2012; Thelwall et al., 2010). SentiStrength is capable of processing different types of information contained in the text including analysis of emoticons and booster words, correction of spelling due to repeated letters, and use of negative words (e.g., not) to flip emotions.
SentiStrength reports two separate numbers for positive and negative emotions. The positive number ranges from 1 (not positive) to 5 (extremely positive). The negative number ranges from -1 (not negative) to -5 (extremely negative). Because both numbers should be considered when evaluating the sentiment of a statement, we use the approach used by Stieglitz and Dang-Xuan (2013) to combine the two numbers. We calculate the sentiment polarity of each statement by creating the following measure, which determines the direction of the sentiment as well as its strength:

\[
\text{polarity} = \text{positive sentiment} + \text{negative sentiment},
\]

Because positive sentiments range from 1 to 5 and negative range from -1 to -5, polarity will have a range of -4 to 4. The other approach to combine the positive and negative numbers is to calculate the total amount of sentiment in a statement regardless of the polarity of positive and negative statements. To attain this, the absolute value of positive and negative sentiments should be added up using the following formula:

\[
\text{sentiment} = (\text{positive sentiment} - \text{negative sentiment}) - 2
\]

Positive sentiment ranges from 1 to 5 and negative sentiment ranges from -1 to -5. Thus, total sentiment has a range of 2 to 10. Therefore we subtracted 2 from (positive – negative) to normalize the range from [2, 10] to [0, 8].

Longevity was measured by counting the number of the days since the review was created. Length of title and review were measured by counting the number of words in the text. Helpfulness of a review was measured as the ratio of “helpful votes” to “total votes” cast by readers in each review (Forman et al., 2008; Ghose & Ipeirotis, 2011; Mudambi & Schuff, 2010). Because we cannot directly measure the readership of OCR, we use the total number of votes of
a review as a measure for its readership. Because people usually vote for a review after they read it, total votes can be a good estimate of readership.

2.4.2. Data Collection

A group of 35000 online reviews of 20 different products were collected from Amazon.com website using crawler software developed by the author. We selected products that had at least 100 reviews. The examined product types were mobile phones, TVs, laptops, tablets, and TV mounts. We eliminated the reviews that had less than 4 votes to ensure that there is a minimum number of votes accumulated for the review (Ghose & Ipeirotis, 2011). The final sample consisted of 2616 reviews. Figure 5 shows the system design of sentiment extraction and mining process.

![Figure 5-System Design of Sentiment Extraction and Mining Process (Essay II)](image-url)
2.4.3. Data Analysis and Results

We first checked the descriptive statistics of our data to determine the proper data analysis approach. Table 5 shows the descriptive statistics of our sample. Variables in our model including review length and total votes represent nonnegative and integer data and their standard deviation is larger than their mean. Hence, the analysis needs to be adjusted for overdispersion using log-transformation (Cameron & Trivedi, 2013). We also checked the distribution of the data. Review polarity, review sentiment, and title polarity follow a normal distribution. However, title sentiment follows a negative binomial distribution with r=1 (Hilbe, 2011). Negative binomial distribution measures the number of failures before a certain number of successes (i.e., r) are achieved. Figure 6 shows the distribution of our different measures.

In terms of considering the fake reviews, we controlled for the effect using Amazon verified purchases. The comments from the authors who have already bought the product from Amazon website are marked as verified purchase. Because those have actually purchased the product, their reviews have a low risk of being fake due to the financial costs (Mayzlin et al., 2012). We controlled for the effect of fake reviews by including VERIFIED_PURCHASE variable in our model which is a binary variable containing 1 if the review is verified purchase and otherwise 0. To analyze hypotheses 1 through 5, we tested the following regression model:

\[
\log (Total \ Votes) \% = \beta_0 + \beta_1 Title \ Sentiment + \beta_2 TITLE\_POSITIVE + \\
\beta_3 Title \ Length + \beta_4 Title \ Sentiment \times TITLE\_POSITIVE + \\
\beta_5 \log(Review \ Length) + \beta_6 \log(longevity) + \beta_7 VEIRIFIED\_PURCHASE
\]

Equation 1 - Prediction of Readership
Table 5-Descriptive Statistics (Essay II)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>Median</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>1-5</td>
<td>2</td>
<td>2.71 (1.70)</td>
</tr>
<tr>
<td>Longevity</td>
<td>15-2196</td>
<td>336</td>
<td>433 (362)</td>
</tr>
<tr>
<td>Total Votes</td>
<td>4-912</td>
<td>7</td>
<td>22.18 (66.89)</td>
</tr>
<tr>
<td>Helpful Votes</td>
<td>0-873</td>
<td>4</td>
<td>16 (60)</td>
</tr>
<tr>
<td>Title Sentiment</td>
<td>0-8</td>
<td>1</td>
<td>1.02 (1.02)</td>
</tr>
<tr>
<td>Title Polarity</td>
<td>-4-4</td>
<td>0</td>
<td>0.18 (1.26)</td>
</tr>
<tr>
<td>Review Sentiment</td>
<td>0-8</td>
<td>3</td>
<td>2.78 (1.48)</td>
</tr>
<tr>
<td>Review Polarity</td>
<td>-4-4</td>
<td>0</td>
<td>0.19 (1.32)</td>
</tr>
<tr>
<td>Title Length</td>
<td>1-24</td>
<td>4</td>
<td>5.06 (3.42)</td>
</tr>
<tr>
<td>Review length</td>
<td>16-2369</td>
<td>84</td>
<td>149.9 (202.8)</td>
</tr>
</tbody>
</table>

Figure 6-Distribution of The Measures (Essay II)
We used negative binomial regression to test the first model in order to control for overdispersion assuming that the data follows a negative binomial distribution (Hilbe, 2011). Title sentiment refers to the total sentiment (positive and negative) available in the title of the review. We created one dummy variable, Title_POSITIVE, to test the moderation effect of polarity on the effect of title sentiment on readership of reviews. The variable has a value of 1 for titles with positive polarity and a value of zero for the others.

Helpfulness is measured as the proportion of helpful votes out of total votes. We used binomial regression with logit transformation to examine Hypotheses 5 through 7 (Baum, 2008):

\[
Helpfulness = \frac{Votes\ Helpful}{Votes\ Total} \% = \\
\quad \beta_0 + \beta_1 Review\ Sentiment + \beta_2 Review\_NEUTRAL \\
\quad + \beta_3 \log (Review\ Length) \\
\quad + \beta_4 Review\ Sentiment \times Review\_NEUTRAL + \beta_5 \text{VERIFIED\_PURCHASE} \\
\quad + \beta_6 \log(\text{Longevity})
\]

Equation 2- Prediction of Helpfulness

We created a dummy variable, Review\_NEUTRAL, to test the moderation effect of polarity on the relationship between review sentiment and helpfulness. The variable has a value of 1 for reviews with neutral polarity (i.e., polarity=0) and a value of zero for the others. We also controlled for the effect of longevity of the review on its helpfulness by including the log transformation of longevity in the model.

We first checked the correlation matrix of independent variables to test for multicollinearity. Table 6 shows the correlation matrix for each regression equation. Because we observed high correlations among some variables, we check the VIF of independent variables. The result of the analysis showed that multicollinearity is not an issue in this study.
Table 6-Correlation Matrix of Independent Variables (Essay II)

<table>
<thead>
<tr>
<th></th>
<th>Equation 1</th>
<th>Equation 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Title Length</td>
<td>Review Longevity</td>
</tr>
<tr>
<td>Title Sentiment</td>
<td>0.09**</td>
<td>1</td>
</tr>
<tr>
<td>Review Length</td>
<td>.15**</td>
<td>.09**</td>
</tr>
<tr>
<td>Longevity</td>
<td>-0.05**</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Review Sentiment</th>
<th>Longevity</th>
<th>Review Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Sentiment</td>
<td>1</td>
<td>-0.01</td>
<td>.12**</td>
</tr>
<tr>
<td>Longevity</td>
<td>-0.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Review Length</td>
<td>-0.01</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

* significant at the 0.05 level, ** significant at the 0.01 level, and *** significant at the 0.001 level

Then we proceeded with model analysis. Both models were significant at the 0.000 level. The first model showed a goodness of fit of 1.160 indicating that negative binomial regression was a good choice for analysis of the proposed model. The proposed relationship between longevity and total votes was significant (b = 0.393, p < 0.001), thus supporting H1. The relationship between title sentiment and total votes was significant but the coefficient was negative (b = -0.087, p < 0.001). Therefore we did not find support for H2. However, the coefficient for the interaction term Title Sentiment × TITLE_POSITIVE was positive and significant (b = 0.439, p < 0.001). Thus we found support for H3. The relationship between title length and total votes was significant but negative (b = -0.017, p < 0.01). Thus, H4 was not supported. Review length was found to be a significant predictor of total votes (b = 0.355, p < 0.001), providing support for H5.

In the second model, review length was a significant predictor of helpfulness (b = 0.407, p < 0.001) providing support for H6. The proposed relationship between review sentiment and helpfulness is significant but negative (b = -0.068, p > 0.001). Thus, H7 is not supported. However, the coefficient for the interaction term Review Sentiment × REVIEW_NEUTRAL is significant and positive (b = 0.162, p < 0.001). Thus, we find support for H8. Surprisingly, both control variables, longevity (b = 0.416, p < 0.001) and VERIFIED_PURCHASE (b = 0.285, p < 0.001),
0.001), had a significant relationship with helpfulness. Figure 7 shows the results of the research model analysis. Table 7 summarizes the hypothesis testing.

![Figure 7-Research Model Analysis Results (Essay II)](image)

**Table 7-Hypothesis Testing Results (Essay II)**

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Hypothesized Relationship</th>
<th>Estimates (Wald Chi-square)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Longevity → Readership</td>
<td>0.393 (227.71) ***</td>
<td>Supported</td>
</tr>
<tr>
<td>H2</td>
<td>Title sentiment → Readership</td>
<td>-0.087 (11.67) ***</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H3</td>
<td>Title sentiment X Title positive → Readership</td>
<td>0.439 (53.27) ***</td>
<td>Supported</td>
</tr>
<tr>
<td>H4</td>
<td>Title length → Readership</td>
<td>-0.017 (7.05) **</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H5</td>
<td>Review length → Readership</td>
<td>0.355 (311.53) ***</td>
<td>Supported</td>
</tr>
<tr>
<td>H6</td>
<td>Review length → Helpfulness</td>
<td>0.407 (1031.03) ***</td>
<td>Supported</td>
</tr>
<tr>
<td>H7</td>
<td>Review sentiment → Helpfulness</td>
<td>-0.068 (37.21) **</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H8</td>
<td>Review sentiment X Review neutral → Helpfulness</td>
<td>0.162 (161.38) ***</td>
<td>Supported</td>
</tr>
</tbody>
</table>

* Significant at 0.05, ** Significant at 0.01, and *** Significant at 0.001

Because the effect of sentiment on performance did not completely match our expectations, we ran post-hoc analyses to explore the difference between positive, negative, and neutral sentiment. We tested the first model for reviews with positive, neutral, and negative title polarity separately. Then we ran the second model for reviews with positive, neutral, and negative sentiment. The results are summarized in Table 8. Positive sentiment is an important
predictor of readership but not helpfulness. Negative sentiment, in contrast, has a non-significant effect on both measures of performance. The results regarding neutral sentiment are mixed. Neutral sentiment in the text of the review improves perceptions of helpfulness. However, reviews with neutral titles attract fewer readers.

Table 8-Group Analysis (Essay II)

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Readership</th>
<th>Helpfulness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Neutral</td>
</tr>
<tr>
<td>Title sentiment</td>
<td>0.302***</td>
<td>-0.209***</td>
</tr>
<tr>
<td>Title length</td>
<td>-0.048***</td>
<td>-0.003</td>
</tr>
<tr>
<td>Review length</td>
<td>0.424***</td>
<td>0.379***</td>
</tr>
<tr>
<td>Review sentiment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longevity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at 0.05, ** Significant at 0.01,  and *** Significant at 0.001

We also tested the second model for different categories of total votes to control for the effect of readership on helpfulness. We categorized the reviews in three categories based on total number of votes: 1-9 votes, 9-99 votes, and 100 votes or higher. We ran the model for the three categories separately and did not find significant differences between them. The only exception is that sentiment does not have a significant effect on reviews with a medium number of votes (10-99). Table 9 shows the results of the analysis.

Table 9-Group Analysis of Helpfulness based on Total Number of Votes (Essay II)

<table>
<thead>
<tr>
<th>Total Votes</th>
<th>1-9</th>
<th>10-99</th>
<th>&gt;99</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Wald Chi-Square</td>
<td>Sig</td>
</tr>
<tr>
<td>Review Length</td>
<td>0.30</td>
<td>95.36</td>
<td>0.00</td>
</tr>
<tr>
<td>Review Sentiment</td>
<td>-0.07</td>
<td>10.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Longevity</td>
<td>0.29</td>
<td>115.08</td>
<td>0.00</td>
</tr>
<tr>
<td>ReviewSentiment * REVIEW_NEUTRAL</td>
<td>0.16</td>
<td>26.54</td>
<td>0.00</td>
</tr>
</tbody>
</table>
2.5. Discussion

This study investigates the effect of review sentiment on readership and helpfulness of online reviews. Despite the proposed positive relationship between sentiment and performance of OCR, we find that sentiment negatively influences both readership and helpfulness of OCR. Emotions have been considered to subvert the rational processes (X.-T. Wang, 2006). In this study we found evidence that consumers may perceive emotional content to be less rational and therefore less useful. The effect, however, is not consistent across various types of emotions. We find a significant positive relationship between the sentiment of title and readership of OCR. This observation can be justified using negativity bias. It is likely that people attribute negative sentiment to the author but relate positive sentiment to the product itself. Because negative emotions are attributed to the author, people are less likely to read reviews with negative titles. However, if the author expresses positive emotions in the title of the review, the likelihood of the review getting attention from consumers increases. This observation also confirms that many consumers look for reviews that reassure them about the choice they have already made (Bailey, 2005; Goldsmith & Horowitz, 2006; Hennig-Thurau et al., 2004).

The effect of sentiment on helpfulness was significant for neutral reviews. It can be argued that a helpful review is a neutral one which communicates both positive and negative aspects of the product. Reviews that lean toward either positive or negative sentiment may be perceived by consumers as biased and consequently less helpful. Although the use of sentiment may help the authors convey their experience with the product to the audience, emotions are effective only where they meet the expectations of the consumers.

Despite the significant effect of title sentiment on the readership of online reviews, length of the title was negatively related to readership. Apparently people use review titles as a quick
source of information about the general theme of the review. Reading and processing longer titles take more time and demotivates people to read them. Unlike title length, review length had a positive effect on both the readership and helpfulness of reviews. Review length can be seen as a signal about the amount of information contained in the review. Longer reviews are perceived to contain more information and thus attract more readerships. Moreover, longer reviews are more likely to analyze different aspects of the product which leads to increased perceptions regarding their helpfulness.

Surprisingly, longevity has a positive effect on the helpfulness of online reviews. In other words, older reviews are perceived to be more helpful. One reason may be the way Amazon sorts the reviews. By default, users view the reviews sorted by most helpful first although Amazon allows them to view the most recent ones first. Because older reviews start receiving votes early, they have a better chance of appearing on the first page which will lead to more votes and perhaps better perceptions regarding their helpfulness. Newer reviews, however, stay at the end of the list and are less likely to receive attention from consumers. Aral (2013) has a similar observation. He observed that positive manipulation of helpfulness of online reviews creates a positive social influence. As a result, longevity not only affects readership, but also influences perception of helpfulness of OCR.

2.5.1. Implications

This study has implications for both theory and practice. From a theoretical perspective, this study introduces a new method of analyzing consumer behavior. Although previous research finds that online reviews significantly influence consumer behavior, few studies have explored the issue using the textual information contained in OCR. Sentiment mining can be utilized to analyze large amount of data that is produced on the Internet every day. Sentiment mining
facilitates processing of big data and is highly replicable. This study also contributes to the body of knowledge by studying both readership and helpfulness of OCR. Furthermore, although previous research has examined the text contained in online reviews, to best of our knowledge this study is the first to also include the title of online reviews in the analysis. We find that the title of reviews contain valuable information and should be considered in the future studies. Finally, this study can be seen as an initial step toward action research in the area of OCR. The findings of this study and future similar studies can be integrated into a single framework to provide a solution for classification and sorting of OCR.

This study also has implications for practice. Most studies use numerical rating and length of the review to investigate performance. Few studies have tried to analyze the textual information contained in online reviews using human subjects. Although the use of human subjects is useful in explaining how people use online reviews, it provides limited usefulness as a solution for online vendors because the process is neither scalable nor replicable through automated systems. To best of our knowledge, this study is the first to provide online vendors with a scalable and replicable solution for analyzing OCR.

The findings of this study can be used by online vendors to develop systems that analyze and classify online reviews based on their potential performance. Our study finds two weaknesses of the current method used for sorting OCR. First, we find that older reviews attract more readers. Many products have thousands of reviews which receive little attention from consumers because they stay at the end of a long list of reviews. Out of the 35000 reviews we collected for this study, 92.5% had less than five votes and 69.7% had no votes. Second, we also observe that older reviews are more likely to be perceived helpful than the recent ones. In other words, perceptions of helpfulness are biased toward older reviews. Although new reviews may
contain up-to-date information about the product and the recent changes made by the manufacturer, such as the software updates, they are likely to be overlooked because of the current sorting methods. These two observations indicate that many potentially useful reviews receive little or no attention from consumers because they are barely seen by consumers. Hence, improving the sorting algorithms of online reviews seems to be necessary. An intelligent system capable of detecting the potentially useful reviews can help both vendors and consumers. It helps consumers by providing them with more valuable information and by saving their time and energy. It also helps vendors to satisfy their customers’ need for information which will allow customers to decide more quickly and may eventually lead to increased sales for the vendors.

2.5.2. Limitations and Future Research

Like other studies, this study has limitations. Although automated sentiment mining is quite useful in analyzing textual information, it still has limitations. The software we used in this study can process different types of textual information. Yet, it lacks processing capability for alternate styles of writing such as sarcasm. Although there are many areas for improvement in the field of natural language processing, future research may provide better insights regarding the information contained in online reviews by using more advanced technology.

The sample used in this study was mainly limited to reviews of electronic products. Previous research shows that product type has an important moderation effect on the performance of online reviews, thus our conclusions may be limited to the product types used in this study. Future research could use a sample of different product types and test the moderation effect of product type on the performance of reviews.

Our sample also lacks language and cultural diversity. The reviews were collected from Amazon US website and all were in English. Different cultures express emotions differently. It is
believed that people in individualistic cultures express negative emotions more freely than those living in collectivist cultures (Takahashi, Ohara, Antonucci, & Akiyama, 2002). Language as the vehicle of information may play a significant role in communicating the message to the reader. Future research may utilize reviews written in other languages to include the effect of language and culture on the performance of online reviews. Future research may also look at the differences in expression of emotions between the real life and the virtual space.

2.6. Conclusions

This research provides insights regarding the predictors of performance of online reviews using a sentiment mining approach. We find that sentiment negatively influences the performance of online reviews with two exceptions: positive sentiment in the title and neutral sentiment in the text of the review. We find longer reviews are more likely to attract readership and to be perceived as helpful. This research can be used by online vendors to create scalable automated systems for sorting and classification of OCR.
ESSAY III: THE EFFECT OF EMOTIONAL AROUSAL ON INFORMATION DIFFUSION IN SOCIAL MEDIA

3.1. Introduction

In recent years, social media has experienced tremendous growth in the number of users. Facebook alone has more than 1.3 billion active users (StatisticsBrain.com, 2014a) and Twitter has attracted more than 600 million active users (StatisticsBrain.com, 2014b). Social media has significantly changed the way humans communicate. Many people use social media to keep in touch with family and friends and receive up-to-date information about what happens around the world. Social media has recently been used to support political campaigns of the candidates of US presidential elections (Naughton, 2012). Politicians are now using Twitter as tool for public diplomacy and to release the up-to-date progress of their negotiations (Kabir, 2013).

Social media have facilitated information sharing in social networks. Previous research shows that several factors such as content related factors and user and network characteristics influence information diffusion in social media. Content related factors include topic (Nagarajan, Purohit, & Sheth, 2010), URL, and hashtags (B. Suh, Hong, Pirolli, & Chi, 2010). User and network characteristics such as social capital perception (Recuero, Araujo, & Zago, 2011), popularity, and homophile (Macskassy & Michelson, 2011) also drive information diffusion in social media. Recent research has focused on the effect of emotions on information diffusion in social media. For example, research shows that that computer-mediated communications (CMC) can effectively transfer emotions. Moreover, the emotions contained in a computer-mediated message significantly influence how the message is processed and interpreted by the receiver (Berger & Milkman, 2012; Riordan & Kreuz, 2010; Salehan & Kim, 2014; Walther &
D’Addario, 2001). Similarly, sentiment of a message may influence its diffusion in social media (Stieglitz & Dang-Xuan, 2013).

Each emotion can be located on a three-dimensional space formed by dimensions of valence (positive–negative), arousal (passive/calm–active/excited), and tension (tense–relaxed) (Wundt, 1908). These three dimensions covary with physical states of the body such as physiological arousal. Although previous research has noted the effect of emotions on information dissemination on social media (Stieglitz & Dang-Xuan, 2013), many aspects of the problem remain unexplored. Among the three dimensions of emotions, valence has received the highest level of attention. Most studies that utilize sentiment mining to study social media, solely focus on emotional valence and the total amount of sentiment in the text. For example, higher levels of total sentiment in a tweet (i.e., both positive and negative) are related to its retweet performance (Stieglitz & Dang-Xuan, 2013).

However, the effect of emotional arousal on information diffusion remains largely unexplored. Emotional arousal influences individuals in several ways. Emotional arousal has been shown to increase action-related behaviors such as moving to help others (Gaertner & Dovidio, 1977), to influence the decision making process of individuals (Kaufman, 1999), and to affect lexical decision response times (Hofmann et al., 2009; Kuchinke, 2007). Considering its significant effect on human behavior, this study considers emotional arousal as a predictor of information diffusion in social media.

This study aims to extend our knowledge on information diffusion in social media by analyzing the diffusion performance of messages in Twitter, a powerful tool for information sharing. We use retweet count as the performance measure of information diffusion in Twitter. Using sentiment mining as our data analysis approach, we suggest that emotions significantly
predict information diffusion in social media. We propose a research model that explains the relationship between different types of sentiment and information diffusion. Then, we test the proposed research model using the data collected from Twitter.com website and discuss the findings and implications of our study. The findings of this study will help individuals, news broadcasting agencies, politicians, and mass media campaigns to improve their performance on social media.

3.2. Theoretical Background

3.2.1. Emotions in Computer-Mediated Communications

Previous research has indicated that CMC can effectively transfer emotions. The receiver of a message can detect the sender’s emotions through verbal cues such as emotion words as well as nonverbal cues such as emoticons (Harris & Paradice, 2007). Moreover, the emotions contained in a message transferred through CMC significantly influence how the message is processed and interpreted by the receiver (Riordan & Kreuz, 2010; Walther & D’Addario, 2001). Sentiment mining can be used to extract sentiment from the text and consequently to predict the behavior of the receiver of the message (Bai, 2011; Salehan & Kim, 2014; Schindler & Bickart, 2012; Sen & Lerman, 2007).

3.2.2. Information Diffusion in social media

Several disciplines from physical to social, and computational sciences have studied diffusion of information in social media. In business and marketing research, information diffusion has also been studied under different titles such as electronic word of mouth and viral marketing (e.g., Berger & Milkman, 2012; Chevalier & Mayzlin, 2006). Although different types of social media platforms such as social networking services (e.g., Bakshy, Rosenn, Marlow, & Adamic, 2012), online review websites (e.g., Salehan & Kim, 2014), photo sharing websites
(e.g., Cha, Mislove, & Gummadi, 2009), weblogs (e.g., Gruhl, Guha, Liben-Nowell, & Tomkins, 2004), and online communities (e.g., Garg, Smith, & Telang, 2011) are used by researchers to study information diffusion, Twitter has received significant attention from academicians because of the way it facilitates diffusion of information in the form of retweets (e.g., Bakshy, Hofman, Mason, & Watts, 2011; Lee, Kwak, Park, & Moon, 2010; Lerman & Ghosh, 2010; Romero, Meeder, & Kleinberg, 2011; Stieglitz & Dang-Xuan, 2012; Jiang Yang & Counts, 2010; Jaewon Yang & Leskovec, 2010). Twitter is particularly important in this context because 75% of Twitter users mainly use it to access information (Recuero et al., 2011).

Previous research looks at both quantity and speed of retweeting as performance measures for information diffusion in Twitter (Stieglitz & Dang-Xuan, 2013). By analyzing over a million tweets, Nagarajan et al. (2010) found that users are more likely to retweet than get involved in direct conversation. They found that popular tweets fall in four categories: call for social action, collective group identity-making, crowdsourcing, and information sharing. They also observed that the retweets from the first three categories create a sparse network (i.e., loosely connected) while the last category, information sharing, has a dense retweet network.

By analyzing 10000 tweets, B. Suh et al. (2010) examined how content features (such as URL inclusion, hashtags, and mentions) and contextual features (including the number of followers and followees, the age of the account, the number of favorited tweets, and the number and frequency of tweets influence diffusion of information in Twitter. They found that tweets containing URL and/or hashtag are more likely to be retweeted. Number of followers/followees and age of the account are also positively related to retweet count. The tweets mentioning other user(s) are less likely to be retweeted.
Macskassy and Michelson (2011) used snowball sampling to collect around 11432 Twitter user accounts from which they collected over 353,000 tweets. Thirty-two percent of the messages were retweets. They analyzed and compared four different models to predict retweet behavior including random model (as benchmark), recent communication model (retweet those recently been in contact with), topic model (retweet topic of interest), and homophily (profile) model (retweet those with similar profiles). They found that the homophily model outperformed the other models followed by recent mode, topic model, and random model.

Recuero et al. (2011) study the effect of social capital on retweet behavior. They find that retweeting not only benefits social network as a whole by spreading information, it but also benefits individual users by allowing them to reach those out of their network while their identity is attached to the message. Because information access is an important motivation for retweeting, timing is an important aspect of retweets. They also find that mentioning the original source when retweeting can add credibility to the message. However, they show that messages that contain too many mentions may look old and hence many users cut the number of mentions so that the information looks fresh. They also believe that many users retweet because it is a convenient way of feeding their followers without actually producing any information. They also find that some users retweet certain people in order to demonstrate their social network. Retweet behavior is also a form of agreement with an expressed idea.

3.2.3. Classification of Emotions

Different categorizations of emotions exist in the literature. Wundt (1908) proposed a dimensional approach for classifying emotions. He suggested each emotion can be located on a three-dimensional space formed by dimensions of valence (positive–negative), arousal (calm–excited), and tension (tense–relaxed). He believed that these three dimensions covary with
physical states of the body such as physiological arousal. Examples of emotions with high arousal and high valence include ambitious, adventurous, self-confident, and delighted. These emotions are all examples of positive emotions that are high in arousal. In the opposite corner is the low valence and low arousal section containing bored, sad, depressed, and doubtful as some examples.

Mehrabian and Russell (1974) proposed a three dimensional model of emotions similar to that of Wundt (1908) composed of pleasure, arousal, and dominance. The model is called PAD emotional state model. The pleasure-displeasure scale represents how pleasant or unpleasant one feels about something. This dimension is similar to what Wundt (1908) calls valence. The arousal dimension is the same as what Wundt (1908) suggested. Finally, the dominance-submissiveness scale measures how controlling and dominant versus controlled or submissive one feels.

Plutchik (1980) proposed a wheel of emotions consisting of eight basic emotions and eight advanced emotions where each advanced emotion is composed of two basic ones. The basic emotions include joy, trust, anticipation, fear, surprise, sadness, disgust, and anger. The advanced emotions include optimism, love, submission, awe, disapproval, remorse, contempt, and aggressiveness. As an example, anger is a highly unpleasant, very aroused, and moderately dominant emotion, while boredom is a little unpleasant, fairly unaroused, and mostly non-dominant.

3.2.4. Negativity Bias

Prior research shows that people respond differently to positive and negative stimuli, and negative events tend to provoke stronger and faster emotional, behavioral, and cognitive reactions than neutral or positive events (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001;
Rozin & Royzman, 2001). More specifically, it has been shown that people are subject to a general bias to show greater weight to negative entities such as events, objects, and personal traits (Rozin & Royzman, 2001). This is generally referred to as negativity bias (Baumeister et al., 2001). Recent studies of communication in the social media context such as Facebook also show that negative sentiment postings induce more feedback in terms of comments compared to those with positive sentiment (Stieglitz & Dang-Xuan, 2012).

3.2.5. Emotional Arousal

Different levels of emotional arousal have different impacts on human behavior. Low levels of arousal are characterized by staying calm and relaxed. Higher levels of arousal, in contrast, are characterized by activity (Heilman, 1997). High levels of arousal increase action-related behaviors such as moving to help others (Gaertner & Dovidio, 1977) and lead to faster response to offers in negotiations (Brooks & Schweitzer, 2011). Stories with high levels of emotional arousal trigger greater emotional reaction in individuals and thus enhance their memory for the story (Cahill & McGaugh, 1995).

Emotional arousal also influences lexical decision tasks. Lexical decision task is an experiment that involves measuring how quickly people identify stimuli as words or nonwords. A mixture of words and logatomes or pseudowords (nonsense strings that respect the phonotactic rules of a language, like trud in English) are presented to subjects. The goal is to recognize whether the presented stimulus is a word or not. Response times of subjects are measured as the dependent variable in the experiment.

Hofmann et al. (2009) ran a series of experiments to investigate the effect of emotional arousal on lexical decision response times. They varied emotional valence and emotional arousal
as experimental factors. They presented negative high-arousing and low-arousing words to subjects and measured their response times. They observed that high-arousing negative words are processed faster than both low-arousing negative words and neutral words. In contrast, low-arousal negative words are processed slower than neutral words. They also observed that negative valence does not influence response times when the treatment is controlled for emotional arousal. Finally, they found out that the effect of emotional arousal on response times is only valid for negative words but not for positive ones. Positive words improve the response time irrespective of their level of arousal.

Another study found that the three discrete emotions of happiness, fear, and disgust significantly predict response times of lexical decision tasks. These three emotions describe as much variance in response times as dimensional and categorical models of emotions (Briesemeister, Kuchinke, & Jacobs, 2011).

Several lexicons have been created for measurement of emotional arousal. A lexicon is a word list where each word has been scored for sentiment strength. The Affective Norms for English Words (ANEW), developed and distributed by the Center for Emotion and Attention (CSEA) at the University of Florida, is a lexicon based on PAD emotional state model (Mehrabian & Russell, 1974) which has measured pleasure, arousal, and dominance for a set of English words (Bradley & Lang, 1999). The initial version included 150 words from Mehrabian and Russell (1974) and 450 words from Bellezza, Greenwald, and Banaji (1986) which together with other words the authors added created a list of 1040 words. The authors used undergrad psychology students to rate each word in terms of the three PAD dimensions on a scale from 1 to 9 (Bradley & Lang, 1999). ANEW reports the mean and standard deviation of each type of sentiment for each word. The latest version is more comprehensive and contains sentiment of
The lexicon has been extensively used and validated by previous research (Bradley, Codispoti, Cuthbert, & Lang, 2001; Jegadeesh & Wu, 2013; Naveed, Gottron, Kunegis, & Alhadi, 2011; Sherdell, Waugh, & Gotlib, 2012; Weymar, Bradley, Hamm, & Lang, 2013; Yu & Hatzivassiloglou, 2003).

Whissell’s Dictionary of Affect in Language is another lexicon developed by University of Columbia scholars (Whissell, 2009). The lexicon contains pleasantness (or valence), activation (or arousal), and imagery for 8742 English words. The authors used 200 volunteers, mostly university students, to rate each word on a three-point scale. The scale for pleasantness includes: (1) unpleasant, (2) in between, and (3) pleasant. Similarly, the scale for activation is: (1) passive, (2) in between, and (3) active. Finally, there are similar scale points for imagery including (1) hard to imagine, (2) in between, and (3) easy to imagine (Whissell, 2010).

3.3. Research Model and Hypothesis

Building upon the previous literature on emotions, we propose a research model that describes the effect of sentiment on information diffusion on social media. We suggest that the levels of total sentiment, emotional valence, and emotional arousal expressed in a Twitter message significantly influence its level of diffusion. Figure 8 shows our proposed research model.

Sentiment of a message can be effectively communicated through the text and significantly influences the perceptions of the reader (Harris & Paradice, 2007; Riordan & Kreuz, 2010; Walther & D’Addario, 2001). Moreover, previous research shows that affective language in the online environment receives more attention and feedback compared to neutral language. For example, people who use affective language in discussion forums receive more
feedback than those who do not (Huffaker, 2010). Both positive and negative sentiment increase participation but in different forms. Positive sentiment increases continued participation while negative sentiment triggers hostile interactions (Joyce & Kraut, 2006). Hence, we expect the total amount sentiment (either positive or negative) expressed in a Twitter message to be positively related to the amount of feedback and attention it receives. A popular way of showing attention toward a Twitter message is to retweet it (Recuero et al., 2011). Thus, we expect tweets with higher amount of sentiment to have a larger number of retweets. Consequently, we hypothesize that:

H1: The larger the total amount of sentiment (positive or negative) a Twitter message exhibits, the more often it is retweeted.

According to negativity bias, people react differentially in response to positive and negative stimuli, and negative stimuli tend to provoke stronger and quicker emotional, behavioral, and cognitive responses than neutral or positive events (Baumeister et al., 2001; Rozin & Royzman, 2001). Moreover, previous research shows that negative sentiment has a similar effect on online users’ behavior. Postings containing negative sentiment elicit more feedback from other users compared to those with positive sentiment (Stieglitz & Dang-Xuan, 2012). Negative sentiment of a message is also more likely to diffuse in subsequent comments.
compared to positive sentiment (Stieglitz & Dang-Xuan, 2012). Drawing on these insights, we argue that the tweets containing negative sentiment are more likely to be retweeted. Negative tweets are more likely to provoke action. Retweeting, as a convenient and quick form of showing reaction to a message (Recuero et al., 2011), is more, likely to be triggered when the original messages contains negative sentiment rather than positive or neutral sentiment. This leads to the following hypothesis:

H2: Polarity moderates the effect of sentiment on retweet count of a Twitter message. The effect will be stronger for tweets with negative sentiment than those with positive and neutral sentiment.

Arousal is a state of mobilization. While low arousal is regarded as relaxation, high arousal is regarded as activity (Berger & Milkman, 2012). Higher levels of emotional arousal increase action-related behaviors such as moving to help others (Gaertner & Dovidio, 1977), enhance long term memory of the events (Cahill & McGaugh, 1995), and lead to faster response in negotiations (Brooks & Schweitzer, 2011). Online users are no exception. A study of New York Times articles shows that articles containing high arousal sentiment are more likely to be shared by email than those with low arousal sentiment (Berger & Milkman, 2012). Hence, we argue that the messages containing high levels of emotional arousal are more likely to elicit reaction from other users. Retweeting a message is a form of reaction to its content (Recuero et al., 2011). Thus, we expect high-arousal messages to elicit greater reaction from online users in the form of retweets. Thus, we hypothesize that:

H3: The larger the amount of emotional arousal a Twitter message exhibits, the more often it is retweeted.
Information processing is a prerequisite to information diffusion. People tend to process information before disseminating them. Previous research suggests that lexical decision response times represent ease of processing of the information (Ferraro, Christopherson, & Douglas, 2006). In other words, words with lower response times are processed more easily. Moreover, emotional facilitation theory proposes that words with greater meaning are processed faster than those with less meaning (Balota & Chumbley, 1984). Drawing on these insights, we argue that people are more likely to spread information that is processed more easily and is more meaningful. Hence, we expect the predictors of lexical decision response times to also be important in the context of information diffusion in social media.

Previous research shows that high-arousal negative content is processed faster than neutral content while low-arousal negative content is processed slower than neutral content (Hofmann et al., 2009; Kuchinke, 2007). We expect the content that is processed more quickly to also be diffused more rapidly. Thus, we expect the expression of high-arousal negative emotions in a messages to enhance its diffusion in social media while expression of low-arousal negative emotions are expected to deteriorate the diffusion of a message. Consequently, we propose the following hypothesis:

H4: Emotional arousal moderates the effect of negative sentiment on retweet count of a Twitter message. The effect will be positive for tweets with high-arousal sentiment and negative for tweets with low-arousal-sentiment.

3.4. Methodology

3.4.1. Data Collection

Twitter API version 1.1 was used to collect tweets. Software was developed by the author to connect to Twitter API. We collected 3219 tweets from CNN’s Twitter page. Data was
collected over a four month period from October 2013 to January 2014. Table 10 shows the
descriptive statistics of our sample. Fifteen percent of the tweets had positive polarity, 40% had
negative polarity, and 45% had neutral polarity.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arousal</td>
<td>0</td>
<td>43.49</td>
<td>10.32</td>
<td>7.76</td>
</tr>
<tr>
<td>Negative Sentiment</td>
<td>-5</td>
<td>-1</td>
<td>-1.67</td>
<td>0.90</td>
</tr>
<tr>
<td>Positive Sentiment</td>
<td>1</td>
<td>4</td>
<td>1.28</td>
<td>0.60</td>
</tr>
<tr>
<td>Polarity</td>
<td>3</td>
<td>-4</td>
<td>-0.40</td>
<td>1.12</td>
</tr>
<tr>
<td>Retweet Count</td>
<td>1</td>
<td>147128</td>
<td>293.51</td>
<td>2668.54</td>
</tr>
<tr>
<td>Total Sentiment</td>
<td>0</td>
<td>5</td>
<td>0.95</td>
<td>1.04</td>
</tr>
</tbody>
</table>

3.4.2. Measurement

The Affective Norms for English Words (ANEW) lexicon was used for analysis of
emotional arousal. To calculate emotional arousal of a text using ANEW, we calculated arousal
of each word separately and then added those values up for each tweet.

SentiStrength software was used for measuring total sentiment and emotional valence
(Thelwall et al., 2010). The software is free for academic research and has been tested and
validated by previous studies (Garcia & Schweitzer, 2011; Gruzd et al., 2011; Stieglitz & Dang-
Xuan, 2013; Thelwall & Buckley, 2013; Thelwall et al., 2012; Thelwall et al., 2010). SentiStrength is capable of processing different types of information contained in the text
including correction of spelling due to repeated letters, analysis of emoticons and booster words,
and use of negative words (e.g., not) to flip emotions.

SentiStrength reports two distinct numbers for positive and negative sentiment. The
positive number varies from 1 (not positive) to 5 (extremely positive). The negative number
varies from -1 (not negative) to -5 (extremely negative). Because both numbers indicate the
sentiment of a statement, we use the approach used by Stieglitz and Dang-Xuan (2013) to
combine the two numbers. We compute the sentiment polarity of each statement by the following formula, which determines the direction of the sentiment (i.e., emotional valence) as well as its strength:

$$\text{polarity} = \text{positive sentiment} + \text{negative sentiment},$$

Because positive sentiment varies from 1 to 5 and negative sentiment varies from -1 to -5, polarity has a range of -4 to 4. The other approach to combine the positive and negative numbers is to calculate the total amount of sentiment in a statement regardless of it is positive or negative. To achieve this, the absolute value of positive and negative sentiments should be added up using the following formula:

$$\text{total sentiment} = (\text{positive sentiment} - \text{negative sentiment}) - 2$$

Positive sentiment varies from 1 to 5 and negative sentiment varies from -1 to -5. Thus, total sentiment has a range of 2 to 10. Hence, we subtracted 2 from (positive – negative) to change the range from [2, 10] to [0, 8]. Figure 9 shows system design and sentiment extraction process.

3.4.3. Data Analysis

We first checked the correlation of our items. Arousal, total sentiment, and polarity are significantly correlated. That is possibly because all these variables are a type of sentiment. Therefore, some level of correlation is not unexpected. All other correlations were insignificant. Table 11 shows the calculated correlations.

We used regression analysis to test the proposed research model. The dependent variable in our model, retweet count, represents nonnegative and integer data and its standard deviation is larger than its mean. Hence, the analysis needs to be adjusted for overdispersion using log-
transformation (Cameron & Trivedi, 2013). To analyze hypotheses 1 through 3, we use the following regression model:

\[
\log(\text{retweet\_count}) = \\
\beta_0 + \beta_1 \text{Total sentiment} + \beta_2 \text{Arousal} + \beta_3 \text{NEGATIVE\_POLARITY} \\
+ \beta_4 \text{Total sentiment} \times \text{NEGATIVE\_POLARITY}
\]

We used negative binomial regression to analyze the model because the dependent variable is a count measure (Hilbe, 2011). Negative binomial regression uses log-transformation which addresses the overdispersion problem of the dependent variable. In the above equation, \textit{Arousal} refers to the level of arousal expressed in the message. \textit{NEGATIVE\_POLARITY} is a dummy variable containing 1 if the message has negative polarity and 0 otherwise. We used the interaction term \textit{Total sentiment} \times \textit{NEGATIVE\_POLARITY} to test the second hypothesis.
Table 11-Correlations (Essay III)

<table>
<thead>
<tr>
<th></th>
<th>Arousal</th>
<th>Polarity</th>
<th>Retweet</th>
<th>Sentiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arousal</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polarity</td>
<td>-0.085**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retweet</td>
<td>0.022</td>
<td>-0.20</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sentiment</td>
<td>0.268**</td>
<td>-0.347**</td>
<td>0.029</td>
<td>1</td>
</tr>
</tbody>
</table>

** Significant at 0.01

To test the fourth hypothesis, we used the following regression model which we analyzed using negative binomial regression. We used the polar extremes approach to break arousal into three levels (George & Prybutok, 2014). AROUSAL_LEVEL is a dummy variable set to 2 if the value for arousal is in the top one-third, to one if arousal level is in the bottom one-third, and 0 otherwise. AROUSAL_LEVEL is a factor in this model and the effect of its different levels (i.e., zero, one, or two) will be analyzed.

\[ \log(\text{retweet count}) = \]

\[ \beta_0 + \beta_1 \text{Total sentiment} + \beta_2 \text{AROUSAL LEVEL} + \beta_3 \text{NEGATIVE POLARITY} \]

\[ + \beta_4 \text{Total sentiment} \times \text{NEGATIVE POLARITY} \times \text{AROUSAL LEVEL} \]

The overall model was significant at \( p < 0.001 \). The results show that sentiment is significantly related to retweet count (\( b = 0.099, p < 0.001 \)) providing support for H1. The coefficient for total sentiment \( \times \) NEGATIVE POLARITY was significant (\( b = 0.216, p < 0.001 \)) thus we find support for H2. The analysis also shows that emotional arousal significantly predicts retweet behavior (\( b = 0.016, p < 0.001 \)), thus H3 is supported. Analysis of the second model provides partial support for our fourth hypothesis. Descriptive statistics show that 15% of the tweets were categorized as high arousal (> \( \mu + \sigma \)) and 16% were categorized as low arousal (< \( \mu - \sigma \)). While high-arousal-negative content is more likely to be retweeted (\( b = 0.351, p < 0.001 \)), low-arousal-negative content negatively influenced retweet count (\( b = -0.157, p < 0.02 \)). We also checked the effect of different levels of arousal on diffusion of positive tweets. While positive
tweets are less likely to be retweeted compared to negative and neutral ones \((b = -0.302, p < 0.001)\), different levels of arousal does not significantly influence the retweet likelihood of positive tweets. The coefficient was significant and negative for high arousal \((b = -0.334, p < 0.001)\), but not significant for low arousal \((b = -0.133, p > 0.1)\). Table 12 shows a summary of hypothesis testing.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Hypothesized Relationship</th>
<th>Estimates (Wald Chi-square)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Sentiment (\rightarrow) Retweet Count</td>
<td>0.10 (12.75) ***</td>
<td>Supported</td>
</tr>
<tr>
<td>H2</td>
<td>Total sentiment * NEGATIVE_POLARITY (\rightarrow) Retweet Count</td>
<td>0.22 (20.99) ***</td>
<td>Supported</td>
</tr>
<tr>
<td>H3</td>
<td>Arousal (\rightarrow) Retweet Count</td>
<td>0.02 (37.35)***</td>
<td>Supported</td>
</tr>
<tr>
<td>H4</td>
<td>Total sentiment * NEGATIVE_POLARITY * [AROUSAL_LEVEL = 2] (\rightarrow) Retweet Count</td>
<td>0.351 (49.68) ***</td>
<td>Supported</td>
</tr>
<tr>
<td></td>
<td>Total sentiment * NEGATIVE_POLARITY * [AROUSAL_LEVEL = 1] (\rightarrow) Retweet Count</td>
<td>-0.157 (6.18)*</td>
<td></td>
</tr>
</tbody>
</table>

* Significant at 0.05, ** Significant at 0.01, and *** Significant at 0.001

3.5. Discussion

This study shows how the sentiment of a message influences its diffusion in social media. We test two dimensions of sentiment, valence and arousal, and show how they influence diffusion of information in Twitter. We find that total amount sentiment contained in a message significantly predicts its diffusion in social media. Drawing upon negativity bias, and in line with previous research, we show that negative content is more likely to be retweeted than positive content (Stieglitz & Dang-Xuan, 2013). We also show that emotional arousal is a significant predictor of diffusion of Twitter messages, and high-arousal content is more likely to be retweeted than low arousal content. We also find that level of arousal specifically influences diffusion of negative messages, and high-arousal negative sentiment positively influences diffusion of Twitter messages. Low-arousal negative sentiment, however, negatively influences
retweet count of a message. We find no significant evidence that the level of arousal influences the diffusion of positive tweets. It seems that most of the variation in retweet count caused by emotional arousal can be attributed to high-arousal negative sentiment.

This study contributes to both theory and practice in several ways. From a theoretical perspective, we show valence is not the only important dimension of sentiment affecting information diffusion in social media and the other dimensions of sentiment can also significantly predict user behavior. We also show there is a significant effect of emotional arousal on information diffusion in social media. We also show how different dimensions of sentiment may interact to jointly predict user behavior on social media. Finally, we use findings from neuro-science to explain the different effect of emotional arousal on diffusion of positive and negative messages. We argue that information that is processed easily and is more meaningful is more likely to be disseminated in social media. Thus, we show how findings from lexical decision tasks can be used to predict users’ retweet behavior. This approach may be used by future research to further explain user behavior on social media.

From a practical perspective, we show how different forms of sentiment influence user behavior on social media. Many organizations are trying to reach a broader audience on social media and to improve the diffusion of their messages. Many of them post links to their websites on social media and try to get the users visit their website. Many of these organizations post a status message with the posted link that is intended to persuade the users to click on the link or forward it to other users. Our findings suggest that selection of the words is very important for organizations. Those responsible for managing the social media accounts of organizations should carefully select the status message that comes with the links they post on social media.
Emotional messages are more likely to be dispersed in social media. Moreover, a combination of negative and high arousal sentiment may also improve the diffusion of the message.

Like any other studies, this study has limitations. First, the software we used in this study is capable of processing different types of text. However, it lacks the processing capability for alternate styles of writing such as sarcasm. While research in the area of natural language processing continues, future studies may provide better insights into determinants of information diffusion in social media by using advanced text analytics techniques. Our sample also lacks language and cultural diversity. The tweets used in this study were collected from CNN Twitter page and are all in English. Previous research, however, shows that different cultures have differences in terms of how they express emotions. For example, it is known that people in individualistic societies express negative emotions more freely than those living in collectivist ones (Takahashi et al., 2002). On the other hand, language as the medium of communication may play a significant role in how the message is processed by the reader. Although our findings regarding the effect of total sentiment and negative emotions are similar to those of Stieglitz and Dang-Xuan (2013), who examined tweets written in German, the case of emotional arousal may be different. Future research may utilize messages from other cultures, which are written in other languages, to control for the effect of language and culture on the diffusion of information in social media. Future research may also look at the differences in expression of emotions between the real life and the virtual space.

3.6. Conclusions

This study investigates the effect of emotions on information diffusion in social media. We find that the level of sentiment a Twitter message carries significantly influences its retweet performance. We also find that negative messages have higher retweet performance than positive
ones. Finally, we find that high-arousal negative sentiment significantly improves information diffusion in social media while low-arousal negative sentiment is negatively related to the diffusion of tweets. This study contributes to the practice by providing individuals, broadcasting agencies, political campaigns, and non-for-profit organizations with strategies to increase their area of impact on social media.
This dissertation investigates the concept of performance in social media. The Merriam-Webster dictionary defines performance as the ability “to do an action or activity that usually requires training or skill” (Merriam-Webster, 2015a, 2015b). Hence, this dissertation looks at the qualities that determine the likelihood of good performance in social media. There are different types of social media. White (2014) provides 7 major categories of social media which are social connections websites, multimedia sharing websites, professional SNS, informational networks, educational networks, hobbies networks, and academic SNS. In this dissertation, we investigate performance in social connections SNS, such as Twitter and Facebook, and informational networks, such as online consumer review websites.

The first essay investigates how users’ motivations lead to their participation in SNS and how participation leads to improved performance at the individual level of analysis. This study identifies four types of motivations, (i.e., hedonic, utilitarian, horizontal social, and vertical social) and shows how they influence two types of SNS participation, namely sharing and collaboration. Finally, the study shows that SNS participation may improve performance of individuals in their personal lives as well as in their workplace.

The second essay investigates performance of online consumer reviews. The study finds two measures of performance for online reviews: readership and helpfulness. The study shows that older and longer reviews attract more readership and are perceived to be more helpful. It also shows that reviews with lengthy titles attract fewer readers. The study also finds that while use of sentiment generally deteriorate performance of most reviews, use of positive sentiment in the title and neutral sentiment in the body of the review may improve its performance. This study provides several directions for further research in this area.
The third essay focuses on performance of information diffusion in Twitter. Twitter is a popular SNS for receiving information (Recuero et al., 2011) and many of its users are trying to reach a larger audience by making a larger number of people retweet their messages. Retweeting a message benefits the original author in two ways. It helps the author to reach people who are not in his/her social network. It also gives the original author some credibility because his/her name is attached to the retweeted message. Hence, retweets are an important indicator of performance on twitter. This essay shows that tweets containing more sentiment are more likely to be retweeted. It also shows that negative messages and those with higher levels of arousal are more likely to be retweeted. Finally, the study finds that high-arousal negative messages are more likely to be retweeted than positive and low-arousal negative messages.
APPENDIX A

DEMOGRAPHIC ANALYSIS OF ESSAY I
<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td>245</td>
<td>69.4</td>
</tr>
<tr>
<td>Woman</td>
<td>108</td>
<td>30.6</td>
</tr>
<tr>
<td>Total</td>
<td>353</td>
<td>100.0</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20~25</td>
<td>36</td>
<td>10.2</td>
</tr>
<tr>
<td>26~30</td>
<td>87</td>
<td>24.6</td>
</tr>
<tr>
<td>31~35</td>
<td>103</td>
<td>29.2</td>
</tr>
<tr>
<td>Total</td>
<td>353</td>
<td>100.0</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under High School</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>High School Graduates</td>
<td>65</td>
<td>18.4</td>
</tr>
<tr>
<td>College Students</td>
<td>24</td>
<td>6.8</td>
</tr>
<tr>
<td>College Graduates</td>
<td>238</td>
<td>67.4</td>
</tr>
<tr>
<td>Graduate School Students</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>Graduate School Graduates</td>
<td>19</td>
<td>5.4</td>
</tr>
<tr>
<td>Others</td>
<td>3</td>
<td>0.8</td>
</tr>
<tr>
<td>Total</td>
<td>353</td>
<td>100.0</td>
</tr>
<tr>
<td>Annual Income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 1,100</td>
<td>10</td>
<td>2.8</td>
</tr>
<tr>
<td>1,101~3,300</td>
<td>124</td>
<td>35.1</td>
</tr>
<tr>
<td>3,301~6,600</td>
<td>178</td>
<td>50.4</td>
</tr>
<tr>
<td>Over 6,601</td>
<td>41</td>
<td>11.6</td>
</tr>
<tr>
<td>Total</td>
<td>353</td>
<td>100.0</td>
</tr>
<tr>
<td>SNS Usage Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CyWorld</td>
<td>253</td>
<td>38.2</td>
</tr>
<tr>
<td>Facebook</td>
<td>171</td>
<td>25.7</td>
</tr>
<tr>
<td>Twitter</td>
<td>131</td>
<td>19.8</td>
</tr>
<tr>
<td>me2day</td>
<td>28</td>
<td>4.2</td>
</tr>
<tr>
<td>YouTube</td>
<td>61</td>
<td>9.2</td>
</tr>
<tr>
<td>QQSpace</td>
<td>6</td>
<td>0.9</td>
</tr>
<tr>
<td>Other</td>
<td>13</td>
<td>2.0</td>
</tr>
<tr>
<td>Multiple Responses</td>
<td>663 (353)</td>
<td>100.0</td>
</tr>
<tr>
<td>Number of Visits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three times a day</td>
<td>54</td>
<td>15.3</td>
</tr>
<tr>
<td>Over three times a day</td>
<td>46</td>
<td>13.0</td>
</tr>
<tr>
<td>Total</td>
<td>353</td>
<td>100.0</td>
</tr>
<tr>
<td>User Trend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early adopter</td>
<td>93</td>
<td>26.3</td>
</tr>
<tr>
<td>Later adopter</td>
<td>260</td>
<td>73.7</td>
</tr>
<tr>
<td>Total</td>
<td>353</td>
<td>100.0</td>
</tr>
</tbody>
</table>
APPENDIX B

MEASUREMENT ITEMS FOR ESSAY I
<table>
<thead>
<tr>
<th>Construct</th>
<th>Measurement Items</th>
</tr>
</thead>
</table>
| Vertical Social Motivation | The reasons that I use the SNS are -  
1: To interact with people that I know  
2: To maintain a closed friendship that I know  
3: To intimate relationship with people I meet at school or work  
4: To interact with people who have the same thoughts and/or opinions  
5: To reconnect people I had forgotten |
| Horizontal Social Motivation | 1: To form a personal online network  
2: To expand offline human network to online  
3: To strengthen my personal network in general  
4: To form social relationship with general others |
| Hedonic Motivation         | 1: To do something interesting  
2: To escape from everyday boring life  
3: To get excited  
4: To feel enjoyment |
| Utilitarian Motivation     | 1: To obtain useful information  
2: To be helpful in my work  
3: To be useful for my business  
4: To improve my personal performance |
| Collective Sharing         | Using the SNS,  
1: I can share my knowledge with many others  
2: I believe that it would be helpful to everyone by sharing knowledge with others  
3: It is recommended to share information within an organization.  
4: I can actively participate in sharing knowledge and experience with others  
5: I can get different knowledge about the same matters from experienced others  
6: I can participate in a special interest group for sharing opinions |
| Collective Collaboration   | Using the SNS,  
1: I can collaborate with others to create knowledge  
2: I can jointly perform tasks effectively  
3: I can plan group events or activities efficiently  
4: I can easily collaborate with others to finish tasks  
5: It is important to have system and culture to encourage collaboration  
6: I can actively collaborate with others who have different backgrounds and experiences |
| Personal Performance       | Based on my experience with the SNS,  
1: (dropped)  
2: I can improve my social relationships  
3: I can effectively update the required information for my life  
4: I can complete my personal tasks more effectively  
5: I can easily share personal information with others  
6: I can easily find people who have expert knowledge for my personal interests  
7: I can effectively communicate with others about various personal issues |
| Work Performance | Based on my experience with the SNS,  
| 1: I can improve business processes  
| 2: I can improve work performance  
| 3: I can improve sharing job related information  
| 4: I can improve communication in work  
| 5: I can improve my understating about the organizational goals and objectives  
| 6: I can find a better way to do business activities  
| 7: I can easily learn new knowledge and information needed for business |
APPENDIX C

LITERATURE COMPARISON FOR ESSAY I
<table>
<thead>
<tr>
<th>Study</th>
<th>Motivation</th>
<th>Participation</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>K.-Y. Lin and Lu (2011)</td>
<td>Perceived complementarily</td>
<td>Continued intention to use</td>
<td>Usefulness, Enjoyment</td>
</tr>
<tr>
<td>Wasko and Faraj (2005a)</td>
<td>Reputation, Enjoy helping</td>
<td>Knowledge contribution</td>
<td></td>
</tr>
<tr>
<td>Hsu and Lin (2008)</td>
<td>Expected relationships</td>
<td>Intention to use</td>
<td></td>
</tr>
<tr>
<td>H. H. Chang and Chuang (2011)</td>
<td>Social interaction, Identification</td>
<td>Knowledge sharing</td>
<td></td>
</tr>
<tr>
<td>(H.-F. Lin, 2007)</td>
<td>Enjoyment in helping others, Organizational rewards</td>
<td>Knowledge sharing</td>
<td>Firm innovation capability</td>
</tr>
<tr>
<td>Tsai and Bagozzi (2014a)</td>
<td>Social influence, Emotional influence</td>
<td>Contribute to the friendship group</td>
<td></td>
</tr>
<tr>
<td>Chai and Kim (2011)</td>
<td>Social ties, Sense of belonging</td>
<td>Knowledge sharing</td>
<td></td>
</tr>
<tr>
<td>Fang and Neufeld (2009)</td>
<td>Initial motivation</td>
<td>Sustained participation</td>
<td></td>
</tr>
<tr>
<td>C. M. Cheung and Lee (2010)</td>
<td>Social identity</td>
<td>SNS use</td>
<td></td>
</tr>
<tr>
<td>Tsai and Pai (2014b)</td>
<td>Affective social identity, Cognitive social identity</td>
<td>Participation</td>
<td></td>
</tr>
<tr>
<td>C.-J. Chen and Hung (2010)</td>
<td>Perceived relative advantage</td>
<td>Knowledge sharing</td>
<td>Knowledge utilization, Community promotion</td>
</tr>
<tr>
<td>Phang et al. (2013)</td>
<td></td>
<td>Participation</td>
<td>Consumption of niche products</td>
</tr>
<tr>
<td>Xu et al. (2012)</td>
<td>Hedonic motivation, Utilitarian motivation, Loneliness, Social presence</td>
<td>SNS use</td>
<td></td>
</tr>
<tr>
<td>Tong et al. (2013)</td>
<td>Intrinsic motivation, Extrinsic motivation, Internalized intrinsic motivation, Internalized extrinsic</td>
<td>Contribution to online reviews</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Motivation</td>
<td>Use of collaborative online shopping</td>
<td>Pre-adoption intention</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------------</td>
<td>--------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>H. Kim et al. (2013)</td>
<td>Shopping enjoyment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y. P. Chang and Zhu (2011)</td>
<td>Information, Entertainment, Connecting with old friends, Meeting new people, Conformity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valenzuela et al. (2009b)</td>
<td>SNS intensity, Facebook group use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross et al. (2009)</td>
<td>CMC motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiu et al. (2006)</td>
<td>Quantity of knowledge sharing, Knowledge quality, Social Interaction ties</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhou et al. (2012)</td>
<td>Utilitarian value, Hedonic Value, Relational Capital, Personalization, Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chai et al. (2011)</td>
<td>Social ties, Reciprocity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y. Kim et al. (2011)</td>
<td>Seeking friends, Social support, Entertainment, Information, Convenience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wasko et al. (2004)</td>
<td>Individual motivation, Social controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J. H. Kim et al. (2010b)</td>
<td>Social motivation, Non-social motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current study</td>
<td>Social vertical, Social horizontal, Hedonic, Utilitarian</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX D

EXPLORATORY FACTOR ANALYSIS FOR ESSAY I
<table>
<thead>
<tr>
<th></th>
<th>Vertical</th>
<th>Horizontal</th>
<th>Hedonic</th>
<th>Utilitarian</th>
<th>Shaping</th>
<th>Collaboration</th>
<th>Personal</th>
<th>Performance</th>
<th>Work Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR1</td>
<td>0.798</td>
<td>0.500</td>
<td>0.297</td>
<td>0.161</td>
<td>0.353</td>
<td>0.308</td>
<td>0.298</td>
<td>0.323</td>
<td></td>
</tr>
<tr>
<td>FR2</td>
<td>0.844</td>
<td>0.574</td>
<td>0.303</td>
<td>0.305</td>
<td>0.451</td>
<td>0.398</td>
<td>0.416</td>
<td>0.431</td>
<td></td>
</tr>
<tr>
<td>FR3</td>
<td>0.835</td>
<td>0.505</td>
<td>0.343</td>
<td>0.228</td>
<td>0.383</td>
<td>0.375</td>
<td>0.309</td>
<td>0.347</td>
<td></td>
</tr>
<tr>
<td>FR4</td>
<td>0.716</td>
<td>0.424</td>
<td>0.375</td>
<td>0.358</td>
<td>0.441</td>
<td>0.315</td>
<td>0.347</td>
<td>0.325</td>
<td></td>
</tr>
<tr>
<td>FR5</td>
<td>0.732</td>
<td>0.411</td>
<td>0.227</td>
<td>0.231</td>
<td>0.374</td>
<td>0.290</td>
<td>0.372</td>
<td>0.356</td>
<td></td>
</tr>
<tr>
<td>UM1</td>
<td>0.447</td>
<td>0.827</td>
<td>0.360</td>
<td>0.402</td>
<td>0.427</td>
<td>0.379</td>
<td>0.430</td>
<td>0.383</td>
<td></td>
</tr>
<tr>
<td>UM2</td>
<td>0.528</td>
<td>0.882</td>
<td>0.346</td>
<td>0.395</td>
<td>0.457</td>
<td>0.414</td>
<td>0.420</td>
<td>0.393</td>
<td></td>
</tr>
<tr>
<td>UM3</td>
<td>0.561</td>
<td>0.891</td>
<td>0.378</td>
<td>0.384</td>
<td>0.460</td>
<td>0.426</td>
<td>0.418</td>
<td>0.421</td>
<td></td>
</tr>
<tr>
<td>UM4</td>
<td>0.578</td>
<td>0.827</td>
<td>0.340</td>
<td>0.309</td>
<td>0.463</td>
<td>0.374</td>
<td>0.426</td>
<td>0.442</td>
<td></td>
</tr>
<tr>
<td>HM1</td>
<td>0.408</td>
<td>0.345</td>
<td>0.809</td>
<td>0.189</td>
<td>0.342</td>
<td>0.224</td>
<td>0.250</td>
<td>0.218</td>
<td></td>
</tr>
<tr>
<td>HM2</td>
<td>0.170</td>
<td>0.276</td>
<td>0.743</td>
<td>0.293</td>
<td>0.260</td>
<td>0.206</td>
<td>0.233</td>
<td>0.206</td>
<td></td>
</tr>
<tr>
<td>HM3</td>
<td>0.248</td>
<td>0.312</td>
<td>0.780</td>
<td>0.205</td>
<td>0.248</td>
<td>0.207</td>
<td>0.198</td>
<td>0.225</td>
<td></td>
</tr>
<tr>
<td>HM4</td>
<td>0.403</td>
<td>0.389</td>
<td>0.872</td>
<td>0.249</td>
<td>0.356</td>
<td>0.211</td>
<td>0.313</td>
<td>0.262</td>
<td></td>
</tr>
<tr>
<td>PM1</td>
<td>0.289</td>
<td>0.327</td>
<td>0.305</td>
<td>0.777</td>
<td>0.537</td>
<td>0.451</td>
<td>0.530</td>
<td>0.462</td>
<td></td>
</tr>
<tr>
<td>PM2</td>
<td>0.284</td>
<td>0.372</td>
<td>0.190</td>
<td>0.873</td>
<td>0.473</td>
<td>0.454</td>
<td>0.537</td>
<td>0.501</td>
<td></td>
</tr>
<tr>
<td>PM3</td>
<td>0.242</td>
<td>0.346</td>
<td>0.205</td>
<td>0.850</td>
<td>0.414</td>
<td>0.423</td>
<td>0.532</td>
<td>0.525</td>
<td></td>
</tr>
<tr>
<td>PM4</td>
<td>0.276</td>
<td>0.395</td>
<td>0.246</td>
<td>0.807</td>
<td>0.431</td>
<td>0.415</td>
<td>0.509</td>
<td>0.505</td>
<td></td>
</tr>
<tr>
<td>CI1</td>
<td>0.443</td>
<td>0.475</td>
<td>0.354</td>
<td>0.483</td>
<td>0.777</td>
<td>0.448</td>
<td>0.527</td>
<td>0.464</td>
<td></td>
</tr>
<tr>
<td>CI2</td>
<td>0.445</td>
<td>0.464</td>
<td>0.327</td>
<td>0.533</td>
<td>0.838</td>
<td>0.541</td>
<td>0.539</td>
<td>0.479</td>
<td></td>
</tr>
<tr>
<td>CI3</td>
<td>0.357</td>
<td>0.351</td>
<td>0.319</td>
<td>0.399</td>
<td>0.776</td>
<td>0.523</td>
<td>0.446</td>
<td>0.413</td>
<td></td>
</tr>
<tr>
<td>CI4</td>
<td>0.349</td>
<td>0.388</td>
<td>0.321</td>
<td>0.419</td>
<td>0.797</td>
<td>0.503</td>
<td>0.530</td>
<td>0.460</td>
<td></td>
</tr>
<tr>
<td>CI5</td>
<td>0.393</td>
<td>0.419</td>
<td>0.269</td>
<td>0.455</td>
<td>0.809</td>
<td>0.546</td>
<td>0.563</td>
<td>0.486</td>
<td></td>
</tr>
<tr>
<td>CI6</td>
<td>0.435</td>
<td>0.387</td>
<td>0.213</td>
<td>0.377</td>
<td>0.742</td>
<td>0.553</td>
<td>0.485</td>
<td>0.479</td>
<td></td>
</tr>
<tr>
<td>CC1</td>
<td>0.373</td>
<td>0.426</td>
<td>0.277</td>
<td>0.507</td>
<td>0.634</td>
<td>0.797</td>
<td>0.547</td>
<td>0.527</td>
<td></td>
</tr>
<tr>
<td>CC2</td>
<td>0.380</td>
<td>0.382</td>
<td>0.230</td>
<td>0.408</td>
<td>0.494</td>
<td>0.804</td>
<td>0.467</td>
<td>0.475</td>
<td></td>
</tr>
<tr>
<td>CC3</td>
<td>0.361</td>
<td>0.372</td>
<td>0.209</td>
<td>0.441</td>
<td>0.471</td>
<td>0.841</td>
<td>0.465</td>
<td>0.463</td>
<td></td>
</tr>
<tr>
<td>CC4</td>
<td>0.296</td>
<td>0.351</td>
<td>0.218</td>
<td>0.403</td>
<td>0.511</td>
<td>0.810</td>
<td>0.466</td>
<td>0.457</td>
<td></td>
</tr>
<tr>
<td>CC5</td>
<td>0.300</td>
<td>0.341</td>
<td>0.181</td>
<td>0.398</td>
<td>0.539</td>
<td>0.788</td>
<td>0.516</td>
<td>0.492</td>
<td></td>
</tr>
<tr>
<td>CC6</td>
<td>0.380</td>
<td>0.374</td>
<td>0.158</td>
<td>0.396</td>
<td>0.514</td>
<td>0.809</td>
<td>0.490</td>
<td>0.505</td>
<td></td>
</tr>
<tr>
<td>UPE2</td>
<td>0.369</td>
<td>0.372</td>
<td>0.258</td>
<td>0.508</td>
<td>0.533</td>
<td>0.460</td>
<td>0.785</td>
<td>0.578</td>
<td></td>
</tr>
<tr>
<td>UPE3</td>
<td>0.353</td>
<td>0.378</td>
<td>0.255</td>
<td>0.495</td>
<td>0.522</td>
<td>0.486</td>
<td>0.816</td>
<td>0.612</td>
<td></td>
</tr>
<tr>
<td>UPE4</td>
<td>0.361</td>
<td>0.401</td>
<td>0.229</td>
<td>0.535</td>
<td>0.549</td>
<td>0.509</td>
<td>0.839</td>
<td>0.605</td>
<td></td>
</tr>
<tr>
<td>UPE5</td>
<td>0.346</td>
<td>0.425</td>
<td>0.279</td>
<td>0.554</td>
<td>0.481</td>
<td>0.519</td>
<td>0.831</td>
<td>0.614</td>
<td></td>
</tr>
<tr>
<td>UPE6</td>
<td>0.368</td>
<td>0.440</td>
<td>0.235</td>
<td>0.510</td>
<td>0.538</td>
<td>0.512</td>
<td>0.787</td>
<td>0.564</td>
<td></td>
</tr>
<tr>
<td>UPE7</td>
<td>0.368</td>
<td>0.380</td>
<td>0.265</td>
<td>0.498</td>
<td>0.554</td>
<td>0.480</td>
<td>0.796</td>
<td>0.577</td>
<td></td>
</tr>
<tr>
<td>OPE1</td>
<td>0.376</td>
<td>0.431</td>
<td>0.277</td>
<td>0.539</td>
<td>0.524</td>
<td>0.514</td>
<td>0.634</td>
<td>0.840</td>
<td></td>
</tr>
<tr>
<td>OPE2</td>
<td>0.365</td>
<td>0.416</td>
<td>0.275</td>
<td>0.547</td>
<td>0.512</td>
<td>0.489</td>
<td>0.640</td>
<td>0.859</td>
<td></td>
</tr>
<tr>
<td>OPE3</td>
<td>0.370</td>
<td>0.359</td>
<td>0.221</td>
<td>0.491</td>
<td>0.511</td>
<td>0.485</td>
<td>0.619</td>
<td>0.856</td>
<td></td>
</tr>
<tr>
<td>OPE4</td>
<td>0.375</td>
<td>0.360</td>
<td>0.211</td>
<td>0.374</td>
<td>0.422</td>
<td>0.428</td>
<td>0.512</td>
<td>0.751</td>
<td></td>
</tr>
<tr>
<td>OPE5</td>
<td>0.337</td>
<td>0.349</td>
<td>0.190</td>
<td>0.489</td>
<td>0.421</td>
<td>0.452</td>
<td>0.570</td>
<td>0.762</td>
<td></td>
</tr>
<tr>
<td>OPE6</td>
<td>0.359</td>
<td>0.388</td>
<td>0.205</td>
<td>0.469</td>
<td>0.459</td>
<td>0.501</td>
<td>0.586</td>
<td>0.820</td>
<td></td>
</tr>
<tr>
<td>OPE7</td>
<td>0.410</td>
<td>0.410</td>
<td>0.230</td>
<td>0.491</td>
<td>0.478</td>
<td>0.552</td>
<td>0.580</td>
<td>0.786</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX E

COMMON METHOD BIAS TESTING FOR ESSAY I
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Theoretical Construct Loading</th>
<th>T-stat</th>
<th>P-value</th>
<th>Common Method Factor Loading</th>
<th>T-stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR1</td>
<td>0.89</td>
<td>36.32</td>
<td>P &lt; 0.001</td>
<td>-0.11</td>
<td>3.47</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>FR2</td>
<td>0.80</td>
<td>16.56</td>
<td>P &lt; 0.001</td>
<td>0.07</td>
<td>1.38</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>FR3</td>
<td>0.88</td>
<td>30.03</td>
<td>P &lt; 0.001</td>
<td>-0.06</td>
<td>1.64</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>FR4</td>
<td>0.62</td>
<td>12.15</td>
<td>P &lt; 0.001</td>
<td>0.10</td>
<td>1.94</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>FR5</td>
<td>0.73</td>
<td>19.91</td>
<td>P &lt; 0.001</td>
<td>0.00</td>
<td>0.07</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UM1</td>
<td>0.84</td>
<td>30.11</td>
<td>P &lt; 0.001</td>
<td>-0.01</td>
<td>0.22</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UM2</td>
<td>0.90</td>
<td>41.80</td>
<td>P &lt; 0.001</td>
<td>-0.03</td>
<td>1.14</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UM3</td>
<td>0.90</td>
<td>41.19</td>
<td>P &lt; 0.001</td>
<td>-0.01</td>
<td>0.30</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UM4</td>
<td>0.79</td>
<td>23.21</td>
<td>P &lt; 0.001</td>
<td>0.05</td>
<td>1.1</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>HM1</td>
<td>0.77</td>
<td>33.00</td>
<td>P &lt; 0.001</td>
<td>0.03</td>
<td>1.08</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>HM2</td>
<td>0.77</td>
<td>26.12</td>
<td>P &lt; 0.001</td>
<td>-0.03</td>
<td>0.78</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>HM3</td>
<td>0.83</td>
<td>34.96</td>
<td>P &lt; 0.001</td>
<td>-0.05</td>
<td>1.67</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>HM4</td>
<td>0.84</td>
<td>48.08</td>
<td>P &lt; 0.001</td>
<td>0.04</td>
<td>1.57</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>PM1</td>
<td>0.64</td>
<td>15.00</td>
<td>P &lt; 0.001</td>
<td>0.14</td>
<td>3.21</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>PM2</td>
<td>0.91</td>
<td>32.04</td>
<td>P &lt; 0.001</td>
<td>-0.04</td>
<td>1.24</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>PM3</td>
<td>0.93</td>
<td>38.32</td>
<td>P &lt; 0.001</td>
<td>-0.08</td>
<td>2.51</td>
<td>P &lt; 0.01</td>
</tr>
<tr>
<td>PM4</td>
<td>0.82</td>
<td>25.79</td>
<td>P &lt; 0.001</td>
<td>0.00</td>
<td>0.06</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>CI1</td>
<td>0.71</td>
<td>13.38</td>
<td>P &lt; 0.001</td>
<td>0.08</td>
<td>1.38</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>CI2</td>
<td>0.82</td>
<td>19.10</td>
<td>P &lt; 0.001</td>
<td>0.02</td>
<td>0.44</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>CI4</td>
<td>0.84</td>
<td>17.41</td>
<td>P &lt; 0.001</td>
<td>-0.06</td>
<td>1.09</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>CI5</td>
<td>0.87</td>
<td>21.61</td>
<td>P &lt; 0.001</td>
<td>-0.04</td>
<td>0.95</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>CI6</td>
<td>0.76</td>
<td>14.47</td>
<td>P &lt; 0.001</td>
<td>0.00</td>
<td>0.08</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>CC1</td>
<td>0.64</td>
<td>11.86</td>
<td>P &lt; 0.001</td>
<td>0.21</td>
<td>3.94</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>CC2</td>
<td>0.88</td>
<td>26.91</td>
<td>P &lt; 0.001</td>
<td>-0.05</td>
<td>1.36</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>CC3</td>
<td>0.97</td>
<td>37.04</td>
<td>P &lt; 0.001</td>
<td>-0.13</td>
<td>3.73</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>CC4</td>
<td>0.87</td>
<td>23.05</td>
<td>P &lt; 0.001</td>
<td>-0.07</td>
<td>1.71</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>CC6</td>
<td>0.73</td>
<td>14.82</td>
<td>P &lt; 0.001</td>
<td>0.06</td>
<td>1.18</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UPE2</td>
<td>0.77</td>
<td>14.72</td>
<td>P &lt; 0.001</td>
<td>0.02</td>
<td>0.26</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UPE3</td>
<td>0.85</td>
<td>14.82</td>
<td>P &lt; 0.001</td>
<td>-0.04</td>
<td>0.63</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UPE4</td>
<td>0.87</td>
<td>19.44</td>
<td>P &lt; 0.001</td>
<td>-0.04</td>
<td>0.77</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UPE5</td>
<td>0.84</td>
<td>18.42</td>
<td>P &lt; 0.001</td>
<td>-0.01</td>
<td>0.25</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UPE6</td>
<td>0.73</td>
<td>11.79</td>
<td>P &lt; 0.001</td>
<td>0.07</td>
<td>1.16</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>UPE7</td>
<td>0.78</td>
<td>18.13</td>
<td>P &lt; 0.001</td>
<td>0.01</td>
<td>0.29</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>OPE1</td>
<td>0.77</td>
<td>18.34</td>
<td>P &lt; 0.001</td>
<td>0.08</td>
<td>1.84</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>OPE2</td>
<td>0.84</td>
<td>23.30</td>
<td>P &lt; 0.001</td>
<td>0.01</td>
<td>0.37</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>OPE3</td>
<td>0.92</td>
<td>23.89</td>
<td>P &lt; 0.001</td>
<td>-0.07</td>
<td>1.73</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>OPE4</td>
<td>0.83</td>
<td>17.13</td>
<td>P &lt; 0.001</td>
<td>-0.09</td>
<td>1.79</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>OPE5</td>
<td>0.77</td>
<td>14.81</td>
<td>P &lt; 0.001</td>
<td>-0.01</td>
<td>0.05</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>OPE6</td>
<td>0.86</td>
<td>22.72</td>
<td>P &lt; 0.001</td>
<td>-0.05</td>
<td>1.35</td>
<td>P &gt; 0.05</td>
</tr>
<tr>
<td>OPE7</td>
<td>0.68</td>
<td>13.46</td>
<td>P &lt; 0.001</td>
<td>0.12</td>
<td>2.42</td>
<td>P &lt; 0.01</td>
</tr>
<tr>
<td>IV</td>
<td>Vertical</td>
<td>Horizontal</td>
<td>Hedonic</td>
<td>Utilitarian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>------------</td>
<td>---------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mediator</td>
<td>Participation (2nd order)</td>
<td>Sharing</td>
<td>Collaboration</td>
<td>Participation (2nd order)</td>
<td>Sharing</td>
<td>Collaboration</td>
</tr>
<tr>
<td>Direct no mediation</td>
<td>0.487</td>
<td>0.487</td>
<td>0.487</td>
<td>0.522</td>
<td>0.522</td>
<td>0.522</td>
</tr>
<tr>
<td>Direct w/ mediation</td>
<td>0.153</td>
<td>0.190*</td>
<td>0.254**</td>
<td>0.184</td>
<td>0.229**</td>
<td>0.286**</td>
</tr>
<tr>
<td>IV -&gt; Mediator beta</td>
<td>0.517</td>
<td>0.519</td>
<td>0.438</td>
<td>0.543</td>
<td>0.535</td>
<td>0.466</td>
</tr>
<tr>
<td>Mediator -&gt; DV Beta</td>
<td>0.642</td>
<td>0.571</td>
<td>0.526</td>
<td>0.622</td>
<td>0.547</td>
<td>0.505</td>
</tr>
<tr>
<td>IV -&gt; Mediator SE</td>
<td>0.093</td>
<td>0.086</td>
<td>0.089</td>
<td>0.085</td>
<td>0.077</td>
<td>0.090</td>
</tr>
<tr>
<td>Mediator -&gt; DV SE</td>
<td>0.073</td>
<td>0.087</td>
<td>0.085</td>
<td>0.087</td>
<td>0.086</td>
<td>0.093</td>
</tr>
<tr>
<td>Sobel Z test Statistics</td>
<td>4.70</td>
<td>4.44</td>
<td>3.85</td>
<td>4.75</td>
<td>4.69</td>
<td>3.74</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: IV – independent variable, DV – dependent variable, SE – Standard Error, * significant at the 0.05 level, ** significant at the 0.001 level
REFERENCES


Fallon, K. (2014). Ellen's epic Oscar selfie is now the most retweeted tweet ever. [http://www.thedailybeast.com/articles/2014/03/02/ellen-s-epic-oscar-selfie-is-now-the-most-retweeted-tweet-ever.html](http://www.thedailybeast.com/articles/2014/03/02/ellen-s-epic-oscar-selfie-is-now-the-most-retweeted-tweet-ever.html)


Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research (JMR), 18*(1).


110


Russell, M. A. (2013). *Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More* : O'Reilly Media, Inc.


Tokunaga, R. S. (2011). Friend me or you'll strain us: Understanding negative events that occur over social networking sites. *CyberPsychology, Behavior, and Social Networking, 14*(7-8), 425-432.


