

EXPLORING BEHAVIORAL PATTERNS  
OF INFORMATION CONSUMPTION ON SOCIAL MEDIA:  
A CASE STUDY OF PBS NEWSHOUR ON FACEBOOK

By

Yan Song

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Information and Media—Doctor of Philosophy

2016

## **ABSTRACT**

### **EXPLORING BEHAVIORAL PATTERNS OF INFORMATION CONSUMPTION ON SOCIAL MEDIA: A CASE STUDY OF PBS NEWSHOUR ON FACEBOOK**

By

Yan Song

In August of 2013, Facebook introduced new algorithms to promote high quality news. Since then, many newsrooms have been enjoying substantially increased traffic brought by Facebook to their own websites. 30% of US adults now get news from Facebook according to a 2014 survey by Pew Research Center. Facebook remarked that this strategy is based on mutual benefits between itself and news publishers. In economists' words, Facebook, as well as other social media platforms including Twitter and LinkedIn, have recognized their roles as multi-sided platforms (MSPs) and sought ways to enhance such roles. It is not the first time for them to try to play an intermediary role as they, especially Facebook, have been constantly coordinating between users, game and app developers, content providers and advertisers.

For newsrooms, however, this ecosystem is fairly new. Watching search engines and social media emerging as new multi-sided platforms, news publishers are forced to learn how to adapt their journalistic practices and business models for yet another new media format, reminiscent of the past when newspapers learned to differentiate themselves from radio and television, and filmmakers learned to have their works visually restructured to display better on smaller and narrower television screens.

The purpose of this dissertation is to better understand how online audiences respond to and utilize for their own ends news publishers' content placed with social media and how publishers

might adjust their online strategies in response. It aspires to develop a realistic and integrated framework for investigating the mechanisms of content consumption and diffusion on social media by drawing on theories from psychology, communication and economics. While testing applications of these social science theories, I employ newly developed statistical tools to take advantage of the thorough documentation of online activities by social media services to address puzzles newsrooms confront in the everyday practice of online journalism. These include the motivations behind liking, commenting and sharing behaviors; how different news topics, message length, sentiment (positive or negative), and reading ease influence these behaviors; how news publishers can accurately assess the performance of news stories given that users' reactions are heavily shaped by the nature of the content; and, finally, the strategies newsrooms might use to gain more attention from social media users and grow their audiences and revenues by attracting more fans and followers in the new media ecosystem.

## **ACKNOWLEDGEMENTS**

First of all, I want to thank my adviser, Steve Wildman, who has been offering his scholarship, wisdom, encouragement, support and care throughout the past six years. Without him, I would never have enjoyed my doctoral program as much or made as many achievements. At the same time, I want to thank Susan, Steve's wife and best friend, for her sweet words and wonderful cooking to nourish both my mind and my body.

Also, I would like to give special thanks to my dissertation committee. I owe a debt of gratitude to Bob LaRose, and Steve Lacy. Coming from different backgrounds, they have added more insights into this study. Ann Kronrod has also shared with me her expertise in marketing and further enriched this study.

Further, I must express my gratitude to the following people, whose kindness has made this study possible. Dan Sinker and Erika Owens supported me throughout the Knight-Mozilla OpenNews Fellowship and respected my pursuit of researching online readership. The other OpenNews fellows inspired and taught me a lot, especially Brian Abelson and Stijn Debrouwere, who share a similar interest in media metrics. During this fellowship, Joel Abrams at the Boston Globe worked with me closely and offered me the data that I needed to build the first model, which was how this study was initiated. After reviewing my preliminary findings, Joshua Benton decided to feature them on at the website of the Nieman Journalism Lab under the headline "Sharing Fast and Slow", which helped me attract a lot of attention from the community. Among them was Travis Daub at PBS NewsHour, who recognized the value of my research and invited

me to examine the data from his newsroom. The collaboration with PBS NewsHour widened my research from a regional to a nationwide audience.

To the open source community and hacks/hackers in Boston, Berlin, Buenos Aires and other places around the world, I have been benefiting tremendously from their work. I am contributing as much as I can as well.

Finally, I dedicate this dissertation to my parents. They spent their adolescent years in the Cultural Revolution and had their educations truncated and had few opportunities in their lives. Through my experience of the world and travels, they can see the potential they both had and what they may have experienced were they given the chance. They appreciate freedom and have granted me much of it, which is rarely seen in Chinese parents. Although they are unable to read my dissertation, I know they will be proud to hold it in their hands. 爸爸妈妈，我爱你们！

# TABLE OF CONTENTS

LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
CHAPTER 1: INTRODUCTION .....	1
CHAPTER 2: BACKGROUND .....	9
News on Facebook .....	9
Facebook Pages and fans .....	11
Facebook Insights .....	12
Facebook users' behaviors and information flows .....	15
PBS NewsHour .....	20
CHAPTER 3: LITERATURE REVIEW .....	22
The origin of gatekeeping .....	22
Mediated communication through the lens of gatekeeping .....	24
Gatekeepers as decision makers .....	28
Dual process theory .....	30
Two Types of processing vs. two processing Systems .....	32
The mostly seamless collaboration of the two minds .....	33
Dual processing and media use .....	35
Two types of processing and Facebook users' behaviors .....	36
Type 1 processing and cognitive ease .....	39
Text length .....	41
Readability .....	42
Emotion and decision making .....	45
Framing .....	47
Measuring the emotional content of text .....	48
Media avoidance .....	50
Negativity bias .....	50
News preferences .....	52
Reasons for Facebook use .....	54
Need to belong .....	55
Need for self-presentation .....	56
CHAPTER 4: RESEARCH METHODS .....	59
Data collection .....	59
Conceptual model and empirical model .....	59
Independent variables: data collection and variable construction .....	65
Control variables .....	70
Empirical distributions for variables .....	75
Modeling overdispersed count data .....	79

CHAPTER 5: RESULTS AND DISCUSSION.....	87
Type 1 processing and cognitive ease.....	93
Dual process theory tested with Facebook data.....	96
Framing, media avoidance and negativity bias.....	98
News preferences and Facebook use .....	99
Post types and time variables.....	101
Other control variables.....	107
Secondary Gatekeeping on Facebook.....	109
CHAPTER 6: LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH .....	113
REFERENCES .....	116

## LIST OF TABLES

Table 1: “Total visits” and “unique users” of the post in the left panel.....	14
Table 2: Flesch Reading Ease scores (Flesch, 1948).....	43
Table 3: The hypothesized relationships between dependent and independent variables. ....	62
Table 4: List of dependent, independent, and control variables. Reference category lists the default for the categorical variables. ....	62
Table 5: Distribution by Eastern Time hour of the day. ....	75
Table 6: Distribution by day of the week.....	75
Table 7: Distribution by month of the year.....	75
Table 8: Distribution by content category. ....	75
Table 9: Pearson correlation matrix for variables with p-value listed in parentheses. Due to the large sample size ( $N=1,734$ ), the correlations tend to be significant. ....	79
Table 10: Exponentiated coefficients of NB regressions with p-values in parentheses. Each column represents a user behavior (dependent variable) and each row represents an independent variable. Background colors indicate association directions and effect sizes. Red denotes positive correlations and blue negative. For each category, the saturation indicates the effect size (note not the significant level). ....	88
Table 11: coefficients for IVs used to test H1a and H1b.....	93
Table 12: coefficients for IVs used to answer RQ1.....	98
Table 13: coefficients for IVs used to answer RQ2.....	99
Table 14: coefficients for the post type variables. ....	101
Table 15: coefficients for control variables regarding hours.....	102
Table 16: coefficients for control variables regarding day of week. ....	104
Table 17: coefficients for control variables regarding holiday.....	105
Table 18: coefficients for control variables regarding month.....	106
Table 19: coefficients for the remaining control variables. ....	107



## LIST OF FIGURES

Figure 1: News consumption on Facebook is jointly influenced by three phases of gatekeeping: primary gatekeeping, algorithmic gatekeeping and secondary gatekeeping.....	5
Figure 2: An example of PBS NewsHour’s post on Facebook.....	14
Figure 3: The Facebook Page of PBS NewsHour and its fans commenting on one of its posts. .	16
Figure 4: The Timeline of a Facebook user, where a post is shared. When a Facebook user leaves a comment on a Facebook Page, it may not be discovered by his friends because his friends may not follow this Facebook Page. By contrast, a shared post on a user’s own Timeline is visible to all his Facebook friends. ....	16
Figure 5: A screenshot from Facebook that shows a Facebook user’s News feed, a list of syndicated posts from both a user’s friends and the Facebook Pages he follows. When he sees something that he thinks worth sharing, either from a friend or a followed Facebook Page like PBS NewsHour, he shares it on his own Timeline and at the same time this post will be placed on his friends’ News Feeds as well. ....	17
Figure 6: Likes and comments are logged in a sidebar on the far right of the Facebook web interface in a smaller font than News Feed posts. ....	19
Figure 7: A screenshot from Facebook that shows the age-gender distribution of PBS NewsHour’s fans on Facebook. ....	20
Figure 8: The scatterplots for five user behaviors measured as total visits and unique users. The correlation values show they are highly correlated.....	61
Figure 9: A screenshot of Facebook Insights that shows the daily impressions for unique users of PBS NewsHour. ....	67
Figure 10: A screenshot of the political section front page, which was assigned no topic label by PBS editors.....	68
Figure 11: A screenshot of a post on Facebook showing three blocks as message, name and description.....	69
Figure 12: A screenshot of questions raised in a PBS NewsHour Facebook post and fans’ comment responses. ....	72
Figure 13: Percentage distribution of posts across a day (top left), a week (top right), a year (bottom left) and across sections (bottom right) .....	76

Figure 14: Distributions of independent variables in histograms. Each observation is a post that was posted by PBS NewsHour on Facebook. ....	77
Figure 15: Distributions of dependent variables (five user behaviors) in histograms. Each observation is a post that was posted by PBS NewsHour on Facebook. The x-axis shows how many users behaviors, e.g., shares, comments, etc., a given post receives. The y-axis shows how many posts receive one, two, or more of each user behavior. The histograms show that the majority of the posts elicited very a few user behaviors and only a few posts elicited a large number of user behaviors. ....	78
Figure 16: Comparison between negative binomial regression and quasi-Poisson regression. The charts on the left show the distributions of the observed and fitted values; those on the right plot the residuals against the fitted values. Among the five behaviors, negative binomial regression explained variation in the current sample better than quasi-Poisson regression, as it has a narrower band of residuals against the fitted values and closely follows the distributions of the observed values except in the cases of “comment” and “like” where QP regression follows the true distribution nearly as well as NB regression. ....	85

## **CHAPTER 1: INTRODUCTION**

In the age of the Internet, a challenge for news producers is “the widening gap between the limitless media and limited attention” (Webster, 2014). In investigating the marketplace for attention, two approaches are adopted. One is on a micro level to study individuals’ reactions to media as stimuli. The other, on a macro level, seeks to explain how people are aggregated or segregated by media and what makes a unit of media content a hit or a bomb (Webster, 2014). The research presented in this dissertation is an attempt to explain on a micro level why social media users react to different content-related stimuli in different ways. The main goal for this study is to develop a better understanding of how users consume and propagate content, especially news, through technology-enabled activities, such as shares, comments, likes, link clicks and negative feedback on Facebook, because these user activities have given Facebook and other social media “a fundamental role in shaping the networked architecture of journalism” (Bastos, 2014).

In their seminal work on media gatekeeping, Shoemaker and Vos pointed out that internet users function as secondary gatekeepers sharing content and providing related feedback to other internet users and stressed that the characteristics of online news messages were underexplored as influences on these behaviors (Shoemaker and Vos, 2009). Donohue, Tichenor and Olien (1972) argued that gatekeeping is not only about information control but is also involved with a wide variety of activities that are aspects of information production and distribution, such as topic selection, transmitting, shaping, repeating and timing. Over the past two decades, this task has been increasingly managed by algorithms designed by various information technology players, such as social media services and search engines (Hamilton, Karahalios, Sandvig &

Eslami, 2014), putting social media services squarely in the gatekeeper role as well. Although content consumers are often aware of editors' gatekeeping, not many of them are aware of algorithmic gatekeeping. For example, 62.5% of Facebook users were not aware that their News Feeds were chosen by filtering algorithms (Eslami et al., 2014).

Today news organizations see two more steps of gatekeeping outside of their control. For them, the ecosystem intermediated by social media services and their users is fairly new. Prior to the advent of the Internet, news publishers ran their own two-sided platforms as they aggregated, and still do, audiences and advertisers using content to bring them together. Watching search engines, social media and other networks emerging as new multi-sided platforms, news publishers are forced to learn how to adapt their journalistic practices and business models to an environment altered by yet another new media format, reminiscent of the past when newspapers learned to differentiate themselves from radio and television, and filmmakers learned to restructure scenes to improve their appearance on TV screens that were smaller and had narrower aspect ratios than cinema screens.

Today, publishers are learning how to exploit social media for two goals: to reach more people and to monetize their audiences through subscriptions and advertising. Regarding the second goal, there is evidence from the music industry that Twitter activities are positively associated with album sales (Joshi, Ma, Rand & Raschid, 2013). Earlier research on Facebook's algorithms has identified a variety of factors as important in determining what items of content are prioritized. For instance, sorting algorithms prioritize content already revealed by earlier users to generate interest, contributing to the rich-gets-richer phenomenon that has been observed on

various websites, such as iTunes and YouTube (Ratkiewicz, Fortunato, Flammini, Menczer, & Vespignani, 2010). In addition, Facebook algorithms also consider friends' sharing behaviors and prioritize their favorites relative to those of other users (Bakshy, Rosenn, Marlow, & Adamic, 2012). However, prior literature has little to say on what types of news are read more and shared more on Facebook and what factors determine the likelihoods that users engage in each of these activities.

Shoemaker and Vos encouraged future researchers to develop the tools and knowledge base needed to “predict whether and in what form a message passes through a gate” (2009, p. 135). This study responds both to Shoemaker and Vos' observation about the audience taking on an important secondary role in gatekeeping and also their charge to researchers to investigate the impact of message characteristics on the likelihood that a message is passed on by a secondary gatekeeper. To do so, I looked closely at the ways Facebook users respond to PBS NewsHour posts using the dual process theoretical lens (explained below) to examine how news topics, message length, reading ease, and sentiment<sup>1</sup> are associated with the following five ways users may interact with news content: share, comment, like, click and leave negative feedback. This research contributes to a better understanding of the roles users play in larger assemblages of interlinked gatekeepers that determine who is exposed to what units of content.

News consumption on Facebook is jointly influenced by three phases of gatekeeping: primary gatekeeping, application of an algorithmic interface and secondary gatekeeping (see Figure 1). During the primary gatekeeping phase, PBS NewsHour's editors decide what to publish on

---

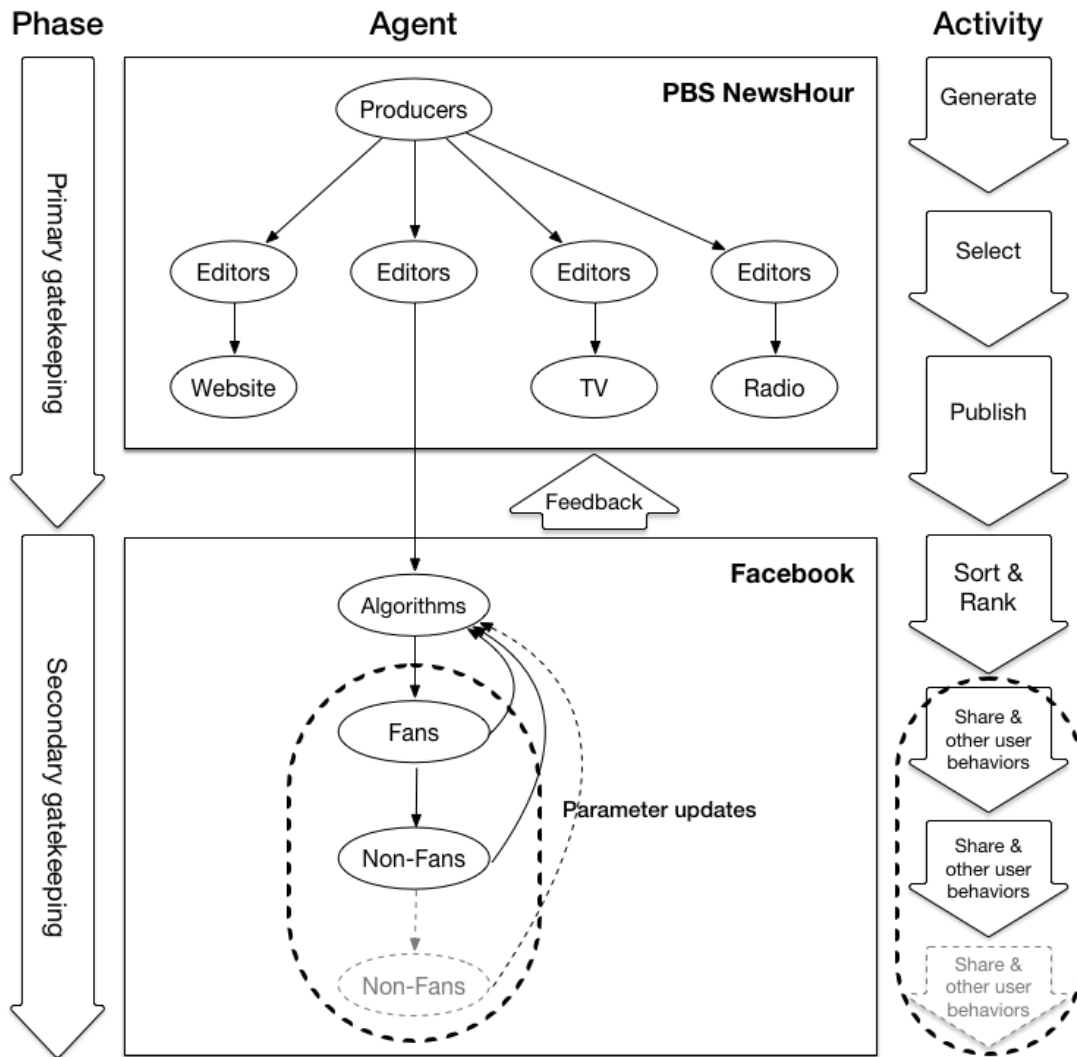
<sup>1</sup> Sentiment is a measurement that quantifies positive and negative expressions of emotions, evaluations, and stances (Wilson, Wiebe & Hoffmann, 2005).

television channels, radio stations, websites and its Facebook Page and other social media accounts. For items posted with Facebook, Facebook's gatekeeping role is performed by an algorithmic interface that automatically sorts NewsHour's posts to determine which might be relevant for each of the news organization's Facebook fans and displays them accordingly. The secondary gatekeeping phase involves both fans and non-fans of a Facebook Page: The fans share the items from a Facebook Page with their friends who may not be fans and may not see these items elsewhere,<sup>2</sup> and some of these friends further share the items with other non-fans. The actions Facebook users take in response to content update the parameters for Facebook's algorithmic selections, as illustrated by the three upward arrows going from fans and non-fans to algorithms in Figure 1.

PBS NewsHour also uses Facebook to learn about their audience members' preferences by reviewing the responses from its fans on Facebook and this process is illustrated as "feedback" from Facebook to PBS NewsHour in Figure 1. The third phase of the online gatekeeping process, enclosed by the dashed circle, is what I focused on in this study. By examining various attributes of NewsHour's Facebook posts and users' responses to the posted content, I investigated whether and how users responded to different types of content with shares, likes, link clicks or negative feedback through the lens of dual process theory. This study aims to relate Facebook users' responses to a variety of content characteristics as factors influencing this gatekeeping process, using the dual process theoretical framework as an analytical lens. The mechanics of this gatekeeping process are described in detail in chapter 2 and relevant theories are reviewed in chapter 3.

---

<sup>2</sup> Non-fans can view Facebook postings by visiting a Facebook Page directly, such as PBS NewsHour's Page on Facebook. However, direct viewing was statistically rare, accounting for less than 1% of the data collected for this study.



**Figure 1: News consumption on Facebook is jointly influenced by three phases of gatekeeping: primary gatekeeping, algorithmic gatekeeping and secondary gatekeeping.**

Marketing scholars have long been studying how audience members respond to different types and features of content from a cognitive perspective (e.g., Berkowitz, Allen, & Beeson, 1996; Allen, 2005). At least until recently, however, journalists have generally ignored or rejected market research on audience members' responses to news items and instead relied heavily on

their own experience for insight into audience members' needs and preferences (Gans, 1979a, 1979b; Jacobs, 1996). For example, editors used observations of their immediate social contacts, such as family and acquaintances, for guidance on how their news products would be understood. They then generated news content for their imagined audiences based on the very small and unrepresentative samples they personally observed (Sumpter, 2000). This practice has changed rapidly in recent years thanks to the high-quality behavioral data newsrooms are able to collect about the segments of their audiences that access their content online and media operators have benefited from their access to these data. Now the *New York Times* data team can look over the reader's shoulder at what he reads in online channels and how he shares it with other people, and the *Times* are hiring high profile scholars to conduct research on user behavior (Evarts, 2014). As a consequence, user behavior data have substantially improved newsrooms' understanding of their audiences compared to unreliable self-reports collected through questionnaires and focus groups.

In the same spirit, this research aims to discover how Facebook users react to media content with different characteristics and to understand how users serve as gatekeepers by sorting and propagating news through technology-enabled activities, such as sharing, commenting, liking, clicking links and leaving negative feedback. This study explores the relationships between certain content characteristics and user behaviors by examining the responses to PBS NewsHour's Facebook postings through the lens of dual process theories. Briefly, dual process theories posit that human responses to environmental stimuli are directed by two types of cognitive processing, often referred to as Type 1 and Type 2 (Evans, 2009a). Type 1 processing is characterized as fast, automatic and intuitive, while Type 2 is better described as slow, effortful and reflective. Shoemaker and Cohen (2006) have observed that news consumption



decisions may be made at an unconscious level, as much of news consumption is automatic and does not involve conscious effort. This dissertation can in part be viewed as a test of this proposition because Type 1 processing might loosely be described in this way. A fundamental hypothesis is that attributes of news items that facilitate Type 1 processing will favor their selection by Facebook users who are NewsHour fans.

This dissertation is organized as follows. Chapter 2 provides background on how news organizations use Facebook to reach and extend the audiences for their content and the nature of the Facebook data used for this study, all of which is necessary background for the chapters that follow. This chapter pays special attention to how news providers utilize Facebook Pages to interact with their audiences, how Facebook reports these interactions back to news providers to help them better cater to their readers and viewers, and how information flows on Facebook through news providers' publications and users' sharing and other behaviors. Chapter 3 reviews relevant prior literature and uses it to develop the theoretical framework for this study and the hypotheses tested and the research questions addressed. Particularly, I examine information flows on Facebook through the theoretical lens of gatekeeping and I investigate certain of Facebook users' behaviors from the perspective of dual process theory. From the literature review, I derive a set of research questions and hypotheses. Chapter 4 sets up the conceptual model that addresses the research questions and hypotheses and provides an overview of the independent, dependent and control variables included in the empirical models employed in Chapter 5, where the choice and design of these models is also discussed. Chapter 5 reports results from the empirical research and discusses to what extent they address the hypotheses and

research questions. Chapter 6 summarizes the findings from the study, discusses the study's limitations and suggests directions for future research.

## **CHAPTER 2: BACKGROUND**

This study examines how Facebook users consume PBS NewsHour's posts and help NewsHour propagate its content on this social media platform. This study is a timely contribution to communication research because a new media ecosystem has been evolving around Facebook and other social media and these platforms are being utilized by traditional news providers to interact with members of their audiences and to grow those audiences as members of social media audiences share content and their thoughts on the content with friends and acquaintances the news providers would not have been able to contact directly.

An empirical investigation of the ways traditional news organizations have utilized Facebook to interact with and expand their audiences requires data on individual news providers' use of Facebook services and Facebook users' responses to the content these organizations post on Facebook. Some of this information can be acquired directly from Facebook, but there is also information specific to each news organization with a presence on Facebook that can only be acquired directly from those news organizations. This study was made possible by access to PBS NewsHour-specific data made available by PBS to the author of this dissertation. To understand, interpret and contextualize the empirical findings reported in this study, is critical that readers start with a certain baseline level of understanding of (1) the ways news organizations utilize Facebook to reach and extend their audiences and (2) the nature of the data employed to conduct this study. This chapter was written to provide that background and context.

### **News on Facebook**

Although they started as vehicles for connecting people, for their users social media have become increasingly important sources for news content, especially Facebook, a platform that

serves over 1.3 billion users (Carr, 2014). According to Pew Research, in 2013, 15% of U.S. adults got most of their news from friends and families over social media, 77% of them followed links to full stories (Pew Research, 2013), and in 2014, 30% of U.S. adults got news from Facebook (Anderson & Caumont, 2014). For major U.S. news websites, 9% of their traffic came from Facebook in 2013 (Sasseen, Olmstead & Mitchell, 2013). This percentage has been soaring since August of 2013 when Facebook introduced new algorithms to promote news publishers (Backstrom, 2013). In early 2014, Facebook contributed 26% of all traffic for online news websites, according Parse.ly, a traffic monitoring service (Wihbey, 2014).

The promotion of news content on Facebook benefits both news publishers and the social media platform. Facebook executives, in an interview with the *New York Times*, stated that “when publishers promote their content on Facebook, [Facebook’s] users have more engaging material to read, and the publishers get increased traffic driven to their sites” (Somaiya, 2014). Or using economic terminology, Facebook has recognized its role as a multi-sided platform (MSP) and sought ways to enhance its performance in that role. It is not the first time for Facebook to try to play an intermediary role as it has been coordinating relationships among users, game and app developers, content providers and advertisers for a number of years. For instance, Zynga started its close relationship with Facebook in 2010, in which Zynga games exclusively used Facebook credits and Facebook helped Zynga grow users (Gannes, 2011). Their partnership ended in 2012 and since then Zynga has been bound by standard Facebook policies with no special treatment (BBC, 2012). While social games are fading in importance, news and other content perceived as key by Facebook are gaining more popularity on Facebook, possibly because Facebook has recognized that the demand for games is not as broad or persistent as the demand for news, hard

or soft. Therefore, it is a natural step for Facebook to identify and promote news on its platform. Besides professionally generated content from news media, Facebook also makes an effort, via Facebook Newswire, to promote popular content produced by users. This strategy is similar to how YouTube surfaces popular content and then further exploits its popularity. Unlike YouTube, which shares ad revenues with amateur and professional content producers (Song & Wildman, 2014), so far Facebook has not started sharing its revenues with any kind of content contributors; but it is speculated that Facebook may experiment with this option with *The New York Times*, BuzzFeed, and *National Geographic* (Somaiya, Isaac & Goel, 2015).

### **Facebook Pages and fans**

For business users, Facebook has created “Facebook Page” for brands and organizations to share stories and connect people. Facebook Pages look like personal pages but are enabled with more features, such as retrieving readership data and selecting Facebook users with specific demographic traits for targeted messages from a brand or organization (Facebook, 2014). Once Facebook users “like” a Page, they have become its “fans” (a Facebook term) and receive its updates on their News Feeds, mixed together with updates from their personal friends. By 2014, there were 40 million active Pages on Facebook and an average user was connected to about 80 of them (Facebook, 2014). Posts from Facebook Pages and updates from users’ friends are mixed together on Facebook’s News Feeds with the sequencing determined by Facebook’s undisclosed algorithms. Facebook claims that the algorithms take into account individual users’ previous interactions with Facebook Pages and friends, which presumably means the more frequently users “like” posts from some Facebook Pages or from friends, the more posts these users will see from the liked “sources” in the future. However, individuals and organizations can pay Facebook for a higher listing (Facebook 2013). The difference is that sponsored content is not marked

explicitly to help end users distinguish it from unpaid content on Facebook, but is only known to Pages administrators.

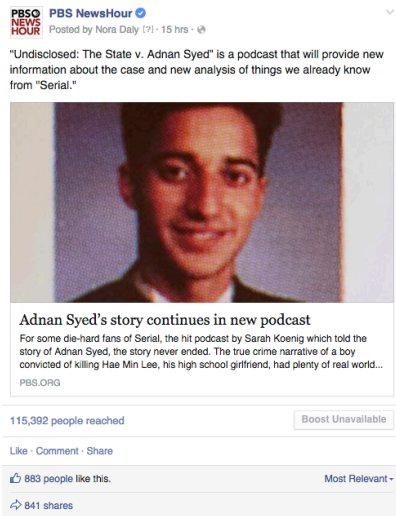
### **Facebook Insights**

Facebook provides a variety of ways for users to interact with its platform and documents and reports measures of the corresponding user interactions with Facebook Insights. These interactions are also called user activities or user behaviors and these terms are often used interchangeably in the academic literature and in industry reports. This dissertation also employs these terms. Two measures are reported for every user behavior: unique users and total visits. A user engaged in the same activity multiple times is counted as one in the count of unique users but more than one in the activity count, i.e., total visits. Both total visits and unique users are used to measure site exposure but the difference between them lies in the units of measure: either it is about individual people or sessions (Bhat, Bevans, and Sengupta, 2002). A session is a connection established between a user and a computer connected to the Internet. It starts when a user lands on a website and expires when a user stays inactive on the website for a certain amount of time, which can be 5, 10, or 30 minutes, depending on a website's configuration. When a person's previous session has expired on a website and that person comes back for another visit, he contributes one more count to the number of "total visits". In this case, he contributes only one count to the number of "unique users", because this metric measures the overall size of an audience regardless of how frequently users come back to a website. These metrics are not extremely accurate because there are caveats in measuring them. For example, when several members in a household share one computer, they may be counted as one "unique user" by a website if they do not use different accounts to log in to this website because in this case they all appear from the same computer and they do not reveal their different identities with

additional information. In the case of Facebook, when a user comes back five times during lunch, commute and work breaks, he generates five visits. At the same time, he contributes only one to the count for unique users Facebook on that day.

Each item published on Facebook is called a *post* and a variety of metrics are developed around posts. During a visit in a session, a user can be exposed to a number of posts, and for each of these posts one “impression” is added to its impression count, “whether the post is clicked or not” (Facebook, 2013). Impressions can be counted either as “total visits” or “unique users”. The metric of total visits focuses on sessions and considers each session as an opportunity to expose people to some content, whereas the metric of unique users focuses on people and considers each user as an opportunity to expose that user to some content. More specifically, total visits, based on sessions, refers to the number of times all the individuals exposed to at least some content are exposed to content during a specified period, while unique users refers to the number of people exposed to content independent of the number of times they were exposed. The total visit metric is the number of exposures generated, while the unique user metric is the number of people exposed. For example, three people visit a Facebook post. One reads it once and never comes back, one reads it twice, and the other reads it three times. In this case, the number of unique users is three, while the number of total visits is  $1+2+3=6$ . It is not hard to see that, for a post, the number of unique users never exceeds the number of total visits. In particular, every activity generates its own visit count, and for a given post, a like and a comment can be counted as separate visits. Table 1 shows that, as of April 8, 2015, a NewsHour post generated 204,696 total visits from 115,392 unique users, which translates into 1.8 visits per user. Similarly, other user behaviors can also be measured in both ways, counted either as unique users or total visits. For

example, one user finds a Facebook post interesting and leaves a comment on it. Later, he sees other people responding to his comment and he leaves three additional comments as his responses to other users. In this case, this user contributes one to the number of unique users and  $1+3$ , or four, to the total visits. The example in Table 1 lists 1,001 total visits and 760 unique users for the number of comments, which translates to 1.32 comments per unique user.



**Table 1: “Total visits” and “unique users” of the post in the left panel.**

	Total Visits	Unique Users
Impressions	204,696	115,392
Link clicks	4,621	4,092
Comments	1,001	760
Shares	851	816

**Figure 2: An example of PBS NewsHour’s post on Facebook.**

Beyond passive exposure measured by impressions, more active user interactions are measured as “consumptions”, a term coined by Facebook, which measures “the number of people who clicked anywhere in your post.” The clicks can reflect a variety of interactions with a post, such as like, comment, share, view a photo, play a video, navigate to an external website like PBS, or leave negative feedback. Further, negative feedback includes “hide a post”, “hide all the posts” and “unlike a Page”. These metrics reflect user behaviors that range from passive consumption to proactive interactions and from favorable to unfavorable. In this study, five user behaviors are investigated—share, comment, like, link click and negative feedback. These terms are fairly self-



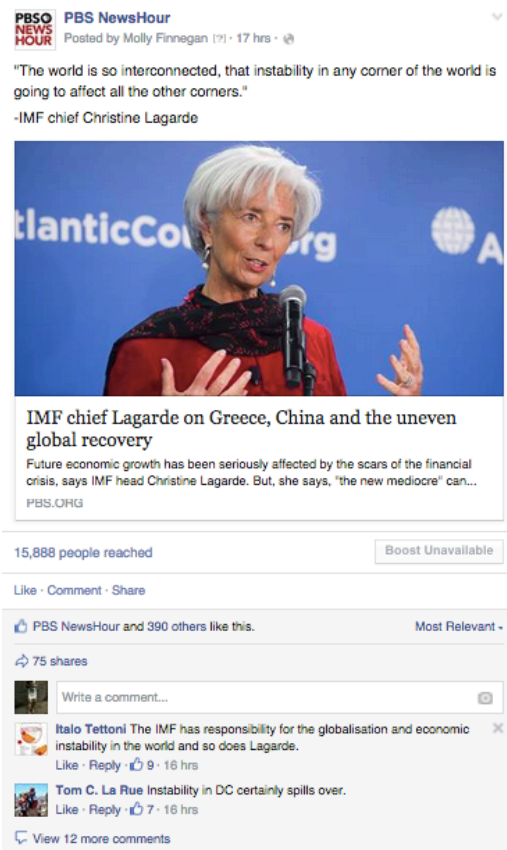
explanatory in terms of the associated activities. While the totals for shares, comments and likes are public to all Facebook users, the records for link clicks and negative feedback are only available to the news organizations themselves through Facebook Insights. Like impressions, more active user behaviors are also counted in both total visits and unique users. For instance, Table 1 shows that 760 users left 1,001 comments to the post, which translates into 1.32 comments per user.

While user behaviors are enabled with limitations determined by technical architecture (Lessig, 1999), they may be responses to a variety of different motivations and goals. For instance, during the 2012 U.S. presidential election, Twitter users worked to propagate information by increasing their retweets and hashtag uses, although this came at the cost of reduced personal interactions (Lin, Keegan, Margolin & Lazer, 2014). Likewise, on Facebook, users also like, share and comment for their own purposes. Also, users can express their approval by clicking the “like” button and disapproval by leaving a negative comment or even hiding a particular post or all the posts from a given publisher on their News Feeds. When people enjoy a post, they pass it on to their friends and families by “sharing”.

### **Facebook users’ behaviors and information flows**

The social functions of Facebook are facilitated by two major components, *Timeline* and *News Feed*. Timeline is a thread of posts a user publishes under his own name, and it is directly accessible to him and to his friends under Facebook’s default privacy settings (see Figure 4). News feed is a list of syndicated posts from both a user’s friends and the Facebook Pages he follows. When he sees something that he thinks worth sharing, either from a friend or a followed Facebook Page like PBS NewsHour, he shares it on his own Timeline and at the same time this

post will be placed on his friends' News Feeds as well. For example, in Figure 5, the first user, Julia, shares a post from HuffPost Parents and the second user Rebecca shares one from Global Voices and both shared posts are shown on their friend's News Feed.



**Figure 3: The Facebook Page of PBS NewsHour and its fans commenting on one of its posts.**



**Figure 4: The Timeline of a Facebook user, where a post is shared. When a Facebook user leaves a comment on a Facebook Page, it may not be discovered by his friends because his friends may not follow this Facebook Page. By contrast, a shared post on a user's own Timeline is visible to all his Facebook friends.**



**Figure 5: A screenshot from Facebook that shows a Facebook user’s News feed, a list of syndicated posts from both a user’s friends and the Facebook Pages he follows. When he sees something that he thinks worth sharing, either from a friend or a followed Facebook Page like PBS NewsHour, he shares it on his own Timeline and at the same time this post will be placed on his friends’ News Feeds as well.**

Through sharing, a post hops from one user's Timeline to another's Timeline on Facebook, and this activity appears on the News Feed where Facebook friends' posts are visible to each other.<sup>3</sup> The user behavior of sharing is an important force driving information flow on Facebook. By contrast, the non-sharing user behaviors (e.g., comment, like, link click, and negative feedback) do not directly propagate a post from one user's Timeline to another's, but, as discussed below, comments and likes do contribute directly to Facebook information flows and all four contribute indirectly because the Facebook sorting algorithm take them into account in determining what information from other sources users see on their News Feeds. Further, notices of "likes" and "comments" (but not the likes and comments themselves) are logged in small print in a sidebar on the far right side of the Facebook web interface (see Figure 6), which may lead some users to explore the associated posts. Because the other two non-sharing behaviors (link clicks and negative feedback) are not logged in a sidebar and are invisible to a Facebook user's friends, they should contribute less to Facebook information flows than likes and comments.

---

<sup>3</sup> People can find posted content by going to the Facebook pages of news sources directly, which forms another channel for people to get exposed to some posts. However, because this study was focused on secondary gatekeeping and also because direct visits to a Facebook Page contributes a very small portion of the total impressions for a news source, this behavior will not be part of the analysis presented in the remainder of the dissertation.



**Figure 6: Likes and comments are logged in a sidebar on the far right of the Facebook web interface in a smaller font than News Feed posts.**

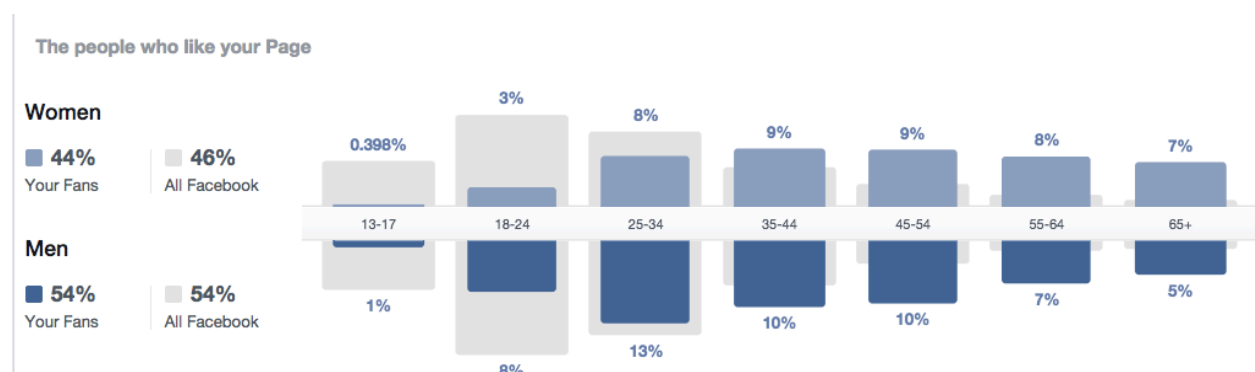
Among the non-sharing behaviors, commenting has one characteristic in common with sharing, which is to allow users to accompany a post with a user-generated message (see Figure 6). When Facebook users comment on posts published by the people or organizations they follow, their comments do not show up on their own Timelines and are not placed on their friends' News Feeds, which are the two main ways to view information on Facebook, but instead are listed in the less prominent sidebar on their friends' Timelines, where the listing then has to be clicked for the comment to be read. Because likes and comments are less prominently displayed than shares and, in the case of comments, require an extra click to be read, they are likely to be accessed by fewer people than shares.

That means in general a shared post on a user's Timeline is highly visible and easily accessed by a user's friends, while his comments on a Facebook Page are accessible but less visible and require the effort of an additional click to view. As a result, although both sharing and

commenting make it possible for users to express their opinions on news items, these differences allow Facebook users to choose between them in managing their relationships with their friends on the platform and helps explain why different stories are associated with different numbers of shares and comments. I will elaborate on users' motives to choose among various activities in the literature review in Chapter 3.

## PBS NewsHour

Founded in 1970, the Public Broadcasting Service (PBS) is a non-profit organization that provides programs to public television stations in the United States. On the PBS website, news stories are assigned from one to three of the following topic labels by its editors: Art, economy, education, health, nation, politics, science and world. These labels function like keywords that help readers grasp a rough sense of a story before making a decision to view it. For instance, an article with the title “In Netherlands, Insurers Compete Over Quality of Care” was labeled as economy, health and politics. The story covered healthcare and therefore was related to health; it discussed the financing of Dutch healthcare and its effect on the Dutch economy; it reported the Dutch government's shake-up of healthcare and therefore politics.



**Figure 7: A screenshot from Facebook that shows the age-gender distribution of PBS NewsHour's fans on Facebook.**

PBS NewsHour started its Facebook Page in 2011 and had attracted over 360 thousand fans as of March of 2015. Among them, 44% were female, slightly below the average level for Facebook (see the first and second columns from the left in Figure 7). Segmented by age and gender, the audience attracted by PBS NewsHour differs considerably from the average Facebook audience. In Figure 7, the grey blocks show the distribution of the overall Facebook audience by age and gender and compare it to the corresponding distribution for PBS NewsHour's Facebook audience. This chart shows that compared to the overall Facebook user base, PBS NewsHour attracts proportionally far fewer fans below the age of 35 and a much larger proportion of fans above this age. These gaps increasingly widen beyond age 35 and below age 24. The fans of NewsHour on Facebook are from large cities in the United States and also outside the US, such as Dhaka of Bangladesh and Lagos of Nigeria.

## **CHAPTER 3: LITERATURE REVIEW**

As was briefly described in Chapter 1, this dissertation focuses on certain roles social media users may play as gatekeepers influencing the flow of news items originating with formal news organizations to individual users of social media. This chapter begins with an overview of the literature on gatekeeping needed to contextualize this research in the stream of ongoing research in this field, with special emphasis on Shoemaker and Vos' (2009) suggestion that research on the influence of news item characteristics other than subject matter on social media users gatekeeping choices would be beneficial. This dissertation responds to this challenge by applying findings from the growing psychology literature on dual process models of decision making to develop hypotheses and research questions concerning Facebook users' responses to news items posted by PBS NewsHour that can be addressed with data available to the author. The remainder of this chapter reviews the studies and findings from the dual process literature that informed this part of my research effort. Because the behaviors measured by dependent variables in the empirical models described in the next chapter may be influenced by factors other than those implicated by dual process theory, literature on factors that influence media consumption and plausibly might also influence social media users' responses to news items posted with social media is reviewed in the next chapter on research methods.

### **The origin of gatekeeping**

The term gatekeeping was introduced by Lewin (1947a), who explored how to change Americans' dietary habits through psychological means by examining how food was produced, distributed and consumed in a social system. Lewin was a trained physicist and he thought social movements could be analyzed mathematically. Adapting concepts from physics, he developed the theory of gatekeeping based on gates, channels, sections, and forces (Shoemaker & Vos,



2009). He used gatekeeping as a metaphor to portray a journey of foodstuffs toward family dining tables. As described by Lewin (1947a) and (Shoemaker & Vos, 2009), before reaching that point, food has to travel through various *channels*. Channels are divided into *sections* with an action point at the beginning of each section. For example, the buying channel includes grocery stores and buying foodstuffs as sections, while seed stores, buying seeds and fertilizers, planting the seeds, and harvesting crops are all gardening channel sections. The buying channel and the gardening channel converge at the kitchen channel, where a chef or a parent selects and cooks food for his customers or family.

The entrance to a channel or a section is called a *gate*, and the individuals or organizations at each gate control the movement of items into or through a channel or a channel section and they are called *gatekeepers* (Lewin, 1951, p. 186). Gatekeepers not only determine which food items to select but also which to reject, which ultimately determines the diet of a family or the menu of a restaurant. More importantly, while moving food items through various channels, gatekeepers make changes to them, such as cutting vegetables into various shapes, preparing meat rare or well done, and frying or baking dough. In addition, certain gatekeepers decide how to present the food. For example, a restaurant may accompany its food with classical music and place it on a table with a tablecloth or pair it instead with pop songs and disposable utensils. The presentation affects how eaters perceive and their appreciation for the food passing through all the channels and all the gates.

Another important factor is *forces*, which affect and influence gatekeepers' decisions. Forces have polarity and they can vary in strength, with stronger positive forces increasing the

likelihood that an item will pass through a gate while this likelihood is diminished the stronger is a negative force. The forces encountered before and after a gate may also vary in polarity. For example, a high price is a force against purchasing a pack of organic blueberries, but once they are bought, the purchaser will make sure his family eats them quickly so that no blueberries will be spoiled or wasted. Lewin believed the concept of force was central to the theory.

### **Mediated communication through the lens of gatekeeping**

Lewin believed that the theoretical framework of gatekeeping could be applied to other scenarios in addition to food consumption and in particular he believed this model “holds not only for food channels but also for the traveling of news items through certain communication channels” (Lewin, 1951, p.187). His ideas inspired communication scholars to examine the flow of information using the gatekeeping model.

When applying gatekeeping theory to communication research, Shoemaker and Vos identify *information* as anything “being moved about in the gatekeeping process” (2009, p. 5). Shoemaker and Vos observe that “information is generally about events” and they call the specific information items aggregated and presented by mass media to the audience *messages* (2009, p. 5). Messages include news, opinion, features, video, and more, and the messages selected for presentation to audiences are called *news items*. In the context of mass media, gates, as decision or action points, may include an editorial meeting, a newsstand, or an online news curator, while gatekeepers can be editors, vendors, or algorithms who/which determine which news items will be allowed to proceed further into an information channel and how they will be modified before passing them on. When a gatekeeper publishes or forwards a news item to the next stage in a channel, Shoemaker and Vos (2009) describe the process using expressions like

“a gatekeeper moves a news item through a gate” or “a news item passes through a gate”. This dissertation employs Shoemaker and Vos’ terminology to describe gatekeeping processes on Facebook.

The first communication scholar to use gatekeeping theory to research the flow of information was David Manning White, who learned about this framework while working as Lewin’s research assistant. White studied a small-city newspaper with a wire editor who was identified under the pseudonym “Mr. Gates”. White realized that Mr. Gates’ news selection process was “highly subjective” (1950, p. 386). White’s study encouraged other scholars to adopt the gatekeeping framework for communication research and many communication models of gatekeeping have since been proposed (e.g., McQuail & Windahl, 1981; Bass, 1969; Westley & MacLean, 1957).

In his study of newsrooms, Bass (1969) identified two important groups of people in the gatekeeping process: News gatherers and news processors. According to Bass, the roles of the former, which includes writers, reporters, bureau chiefs and city editors, involve collecting and reporting information. Those in the second group, which includes copy editors, copyreaders, and translators, modify and integrate the copy into the final product that will be transmitted to an audience. However, even though news audiences sent feedback and signaled their preferences when Bass was writing his paper in 1969, Bass failed to recognize this feedback communication from news audiences to newsrooms. Feedback channels for PBS NewsHour’s Facebook content are indicated by the arrow from the lower box to the upper on in Figure 1 and the channels that

lie entirely within the relationship between Facebook and its users are examined more closely in this chapter's subsections on Facebook users as gatekeepers.

Westley and MacLean (1957) developed a gatekeeping model with two abstract agents, sender A and receiver B, where A and B can be individuals or groups. As examples, sender A might be a politician or a spokesperson for a business, and receiver B might represent potential voters in a city or consumers visiting the same shopping mall. As Sender A has something to tell receiver B, he can do so either through face-to-face communication or by using some form of mediated communication. Examining prior gatekeeping models, Westley and MacLean recognized the missing feedback channels from the receiver to mass media organizations. However, their model allowed for only a few communication channels and could not capture the high degree of interactivity made possible more recently by digital media. Lacy (1989) expanded this model to make it more interactive and he presented a more complete picture of how gatekeeping works in news organizations. Lacy's revision and extension of the gatekeeping model has been applied by later scholars to the now widely adopted digital media. Also, Westley and MacLean omitted communication among receivers, which is a popular and defining feature of social media. However, this phenomenon is nothing new because word of mouth has long been an important channel for the flow of information in human societies (Stephens, 2014, p. 7). Shoemaker and her colleagues included readers as gatekeepers to update the gatekeeping model for the environment of online news and studied news characteristics as forces around gates. (Shoemaker, Johnson, Seo & Wang, 2010) They found that readers in the US and Brazil differed from Chinese readers in the character of the news they shared. For example, Chinese readers preferred to share more positive news than did US readers.

Among gatekeepers in the media ecosystem, newsrooms and their information suppliers, such as government agencies, corporate bureaucracies, and advertisers, remain crucial gatekeepers who shape what a mass audience sees and thinks about. Breed (1955) and Tuchman (1978) observed that it was everyday practice for reporters to rely on “news values” to decide which potential news items to offer to their editors, while news values are developed internally within a newsroom as a set of “criteria of relevance which guide reporters’ choice and construction of newsworthy stories” (Chibnall, 1977, p. 13). On the other hand, Gans (1979a) and Schlesinger (1987) showed that journalists generally worked with second-hand information and their jobs often boiled down to compiling stories. Further, Fishman (2014) argued that government agencies and corporate bureaucracies to a large extent paint the reality perceived by news consumers because these organizations constantly provide newsrooms with updates and in a sense subsidize their information gathering; taking advantage of newsrooms’ common reliance on “beats”, which consist of places for journalists to go to, people to interview, and topics to cover for their media outlets. As a result, “without exception, only formally constituted organizations and groups were the routine subjects of information gathering on beats” (Fishman, 2014, p. 49).

As an organizing framework for these forces of gatekeeping, Shoemaker and Reese (2014) have proposed an umbrella model comprised of five hierarchically ordered levels of influence. From macro to micro, they are social systems, social institutions, organizations, routines, and individuals. At the individual level, individual news consumers can be important influences on other individuals’ reasoning and decision making through the news items they pass on to them.

The last two decades have witnessed the rise of online media together with their audience members' increasingly critical role as gatekeepers, who, for example, may be the first to break a news story or make a story already broken go viral. Against this backdrop, it has become vital for communication scholars to develop an understanding of how consumers perform their gatekeeping roles. This research endeavor can be seen as building on a long history of research on mass media audiences' preferences and reactions to news items that Golding (1981, p.74) described as focusing on the following questions: "Is [the news] important to the audience or will it hold their attention? Is it of known interest, will it be understood, enjoyed, registered, perceived as relevant?"

### **Gatekeepers as decision makers**

Shoemaker and Vos (2009, p. 7) recognized that online "audience members have become active in a secondary gatekeeping process", where the usual mass media process stops and where Bass (1969) ended his model, because, compared to older media, online media provide more opportunities and mechanisms for audience members to interact with news organizations and with each other. The study presented in this dissertation focuses on the individual level of the Shoemaker and Vos (2009) model and analyzes how characteristics of news items other than their subject matter can affect individuals' choices as gatekeepers on Facebook.

Social media users influence the selection and flow of news items in three ways. First, by forwarding and sharing news items, they serve directly as gatekeepers for each other. Second, social media users shape the gatekeeping process indirectly through sharing and other ways they interact with content because when engaging in these activities they provide social media

platforms and news organizations information about their preferences and their predilections to share news items with their friends or in other ways make their opinions of these items known. For example, newsrooms may decide which stories to develop further based in part on the online performance of already published stories as measured by page views, clicks and other content interaction metrics; further, some newsrooms incorporate audience members' selections as components of their content strategies by highlighting stories appearing on "most emailed articles" on the *New York Times* website, "Twitter Trends" or "Facebook Newswire". Third, although not addressed in this study, it is worth noting that Facebook members can serve as their own gatekeepers by subscribing to news organizations' Facebook pages or identifying for Facebook sources from which they do not want to receive posts.

For both news producers and consumers, the gatekeeping process is essentially a series of cognitive judgments on news items that determine whether information items are allowed to pass through a gate (Nisbett & Ross, 1980; Gandy, 1982). At the individual level gatekeeping is a cognitive process executed by news consumers, such as Facebook users, who utilize judgmental heuristics for problem solving and decision making (Shoemaker & Vos, 2009). Applying the theoretical framework of cognitive heuristics proposed by Kahneman, Slovic, and Tversky (1982), Shoemaker and Vos have argued that, when consuming media content, people generally process information without engaging their critical facilities (p. 37), a description that resembles that for one of the two types of cognitive processes that I discuss in the section on the dual process theory below.

## Dual process theory

In the 1970s, Kahneman and Tversky conducted a series of “heuristics and biases” studies and showed that people heavily rely on heuristics established from past experiences and that these heuristics may embed biases that influence performance on current cognitive tasks (Kahneman & Tversky, 1973). This research program interested Evans, who has since immersed himself in investigating why people make irrational decisions when they have the cognitive tools required to make rational ones (Evans, 2013). Later, Evans and Over (1996) proposed dual process theory after being inspired by Reber’s (1993) suggestion that there were implicit and explicit cognitive processes underlying reasoning, judgment and decision-making. For example, generally “ $1 + 1 = ?$ ” can be handled by the implicit cognitive process by an average adult whereas “ $23 \times 46 = ?$ ” has to be calculated using the explicit cognitive system.

According to Evans (2013), dual process theory was proposed as a response to some puzzling questions regarding human cognition and decision-making. As examples, Evans (2013) raises the following questions: Why are people influenced by context in their reasoning when instructed to form a logical argument? Does the demonstration of cognitive biases mean that people are irrational? Is logic the right normative standard by which to judge reasoning based on cognitive biases? Expanding on Evans and Over’s ideas and findings, Stanovich (1999) explored a much wider range of psychological phenomena involving both explicit processing and implicit processing, and he called the two processes System 1 and System 2. Kahneman was “greatly influenced” by Stanovich’s research and adopted Stanovich’s System 1 and System 2 terminology and employed it in his bestseller *Thinking Fast and Slow* (2011, p. 450). There is now a school of scholars that has reached the conclusion that there are two cognitive processes operating in human brains and these scholars have been developing what is now commonly



called dual process theory by expanding on and also criticizing each other's theories and experiments (e.g., Evans & Over, 1996; Kahneman & Tversky, 1973; Reber, 1993; Stanovich, 1999).

While there are variations on dual process theory (Evans, 2009a) and considerable disagreement over the precise nature of the mechanisms involved, it is generally accepted that there are two distinct types of cognitive processes—one intuitive, rapid, and automatic; the other reflective, slow, deliberate and correlated with general intelligence or IQ (Evans, 2010). These processes work so differently that some scholars have used the metaphor “two minds” to characterize them and refer to the two minds hypothesis (Evans, 2010, p. 76). A large and varied set of descriptive terms has been applied to the two mechanisms in the research literature: automatic vs. controlled (Schneider and Schiffrin, 1977), experiential vs. rational (Epstein, 1994), heuristic vs. systematic (Chaiken, 1980), implicit vs. explicit (Reber, 1993), Heuristic and analytic (Evans, 1989), associative vs. rule-based (Sloman, 1996), intuitive vs. analytic (Hammond, 1996), System 1 vs. 2 (Stanovich, 1999), holistic vs. analytic (Nisbett et al., 2001), adaptive unconscious vs. conscious (Wilson, 2002), reflexive vs. reflective (Lieberman, 2003), stimulus-bound vs. higher-order (Toates, 2006), impulsive vs. reflective (Strack & Deutsch, 2004). These varied conceptualizations are closely related and commonly examined across four main dimensions of human cognition: Consciousness (unconscious or conscious, implicit or explicit), evolutionary age (evolutionarily old or new), and characteristics of functionality (pragmatic or logical, parallel or sequential) (Evans, 2013, p. 214).

## **Two Types of processing vs. two processing Systems**

Central to the two minds hypothesis is the distinction between intuition and reflection in human cognition, both of which involve reasoning and decision making. The implicit mind relies on intuition and habit, while the explicit mind considers alternative courses of action when called upon by the implicit mind (Evans, 2003).

Rejecting Stanovich's terminology, Evans argues the dual processes should not be examined as distinct systems because he sees the intuitive and reflective processes relying upon many subsystems, some of which are common to both processes. Evans has suggested that *Type 1* and *Type 2* processing would describe these two cognitive categories more accurately than "System 1 and System 2" or other system-based terminology, because the notion that any "singular systems underlie these two kinds of processing" is wrong (2007). Instead, Evans argues that both cognitive processes involve multiple systems. For example, Type 1 processing draws on innate modular systems, such as vision, perception, attention and language processing. Further, while so-called "System 2" appears to be a singular system because it relies on a central working memory system of limited capacity, it still has to call on various subsystems and these calls are made automatically. Therefore, Evans believes that the proposal of two types of cognitive processing is less problematic than that of two systems, and both Evans and Stanovich have abandoned the System 1 and System 2 terminology and adopted the Type 1 and Type 2 processing terminology (Evans & Stanovich, 2013; Stanovich, 2005).

Type 1 and Type 2 processes are "two distinct ways of thinking, deciding and acting" (Evans, 2010). The rapid Type 1 mind collects and supplies information to the slow Type 2 mind, cueing default behaviors which the Type 2 mind may approve, reject or modify. Type 2 processes are

slow and sequential, because they require access to a single central working memory system that holds and processes information that is newly acquired. Working memory is transient and limited in capacity. Therefore, the functioning of the Type 2 processes is correlated with individual differences in cognitive capacity and ability to resist disruption to working memory processes. By contrast, Type 1 processes do not rely on working memory and operate more rapidly and often in parallel.

Whereas it is widely accepted that Type 2 processing is slow, cognitively expensive and tiring, there is a widely accepted belief that Type 1 processing is entirely unconscious. This view is disputed by Stanovich and Evans who argue that the fast and easy mind has to reason based on at least some knowledge and merely being conscious or unaware of the processing being done does not fully characterize the differences between the two types of processing. Stanovich, (2009c) further points out that humans have to be conscious to be able to do anything other than sleep.

### **The mostly seamless collaboration of the two minds**

We may wonder, in the context of media consumption, whether Type 1 and Type 2 minds take dominance in turn and, if so, when media consumers switch from one mind to the other. However, the fast and slow minds do not work in shifts and the slow mind does not have to be activated by the fast mind, as Kahneman (2011) states. As Evans (2008) points out, there is no singular system to handle cognition alone, in which case, and to the extent it matters, it would be more instructive to employ the Type 1 and 2 processing framework for scholarly examination of media consumption. For this study, I examined responses to PBS NewsHour posts on Facebook through the lens of dual process theory and statistically tested predictions derived from the theory.

In fact, Type 1 and Type 2 processes interact as well as compete with each other every moment we are awake, and during these interactions and competitions they are simultaneously active. One way this happens is when the Type 2 mind overrides the Type 1 mind's initial intuitions and judgments. Intuition and habit, which underlie Type 1 processing, exert a powerful influence on decision-making because they often suggest an immediate answer or a course of action independent of any serious reflection. The suggestions and inferences from Type 1 processing are called default responses and they are default because most of the time we accept the default judgments and permit the default behaviors suggested by Type 1 processing. However, the Type 2 mind has the capacity to override the default responses and replace them with the products of slower and more contemplative thoughts (Evans, 2007, 2008). Our two minds seem to cooperate *seamlessly* most of the time, except when the Type 2 mind becomes alert. In Evans' words, "the conscious person is a fictitious narrative, told with the concepts of a one-mind folk psychology. Only when the conflict between the two minds becomes extreme, does the storytelling break down" (2010, p. 211).

Besides correcting default decisions suggested by Type 1 processing, Type 2 processing also trains the fast mind. Referring to the process of learning to drive, Evans says the Type 2 mind "practices the actions required, intervenes with volitional control of our actions" (2010, p. 201). When we first learn to drive, we gather all our attention and energy for this task because driving is far from an immediate and intuitive cognitive task. As we drive more, this task becomes second nature and we can even make light conversation or sip coffee while processing information about the road and making immediate and proper driving decisions. At this point,

driving tasks have graduated from the reflective Type 2 mind and been handed over to the intuitive Type 1 mind.

### **Dual processing and media use**

People often base their judgments and decisions on intuition without seeking a second opinion from their Type 2 minds, because humans are “cognitive misers” (Stanovich, 2009c), which is what we would expect if the energy required for Type 2 processing is a scarce resource and too limited to be applied to all decisions that have to be made. Facebook users are presumably doing the same in their roles as news consumers and gatekeepers when confronted with an endless stream of texts, pictures, and videos. At the same time, Facebook users need to process other streams of information and make decisions for the other areas of their lives. Thus we would expect them to frequently make decisions on the basis of small amounts of apparently relevant information and ignore other data, using “fast and frugal heuristics”, a strategy in which “people rely on simple rules of thumbs” to make decisions (Marewski, Galesic & Gigerenzer, 2009, p. 121). In sum, we should not be surprised to find that consumer-gatekeepers on Facebook rely heavily on their Type 1 minds for information collection and distribution.

Moreover, with the many options for requesting, posting and responding to information made available by Facebook, Facebook users operate in a complex information environment. In studying other information environments, McKenna and Martin-Smith note that “complexity and chaos in this kind of non-linear and unpredictable situation has redefined decision making” and “personal resources such as time and attention are scarce in such positions” (2005, p. 832). They state that decisions in such situations are not essentially determined by intelligence or education but conditioned upon moderating factors, such as personality and motivation. Even though

members of the PBS NewsHour audience are better educated than the general US population and the articles posted may be informationally dense and targeted towards educated news consumers, they, like everyone else, must still work with limited time and cognitive resources and are still likely to rely heavily on their Type 1 minds to process NewsHour posts on Facebook.

Furthermore, pervasive media multitasking impedes use of the Type 2 mind. Today in developed countries people spend about one quarter of their media consumption time consuming multiple media simultaneously and Facebook is found to be one of those that is frequently consumed while users are engaged with other media products (Koolstra, Ritterfeld & Vorderer, 2009). Because the central working memory system determines the cognitive capacities potentially available for Type 2 processing, any one of the other media consumed would compete for the working memory capacity required for Type 2 processing. As a result, consuming multiple media simultaneously is likely to lead to information loss and misjudgment (Koolstra et al., 2009).

### **Two types of processing and Facebook users' behaviors**

The five Facebook user behaviors examined in this study are preceded by either Type 1 or Type 2 processing on some occasions and by both on other occasions. Type 2 processing occurs when a user stops and thinks before conducting a behavior on Facebook, let it be liking, sharing, commenting, clicking a link or leaving negative feedback. But this is not always the case. For example, likes are presumably the easiest actions to initiate on Facebook and might be thought of as nearly automatic actions that can plausibly be considered as consequences of Type 1 mind reactions to content. By contrast, since comments must be composed, a Facebook user has to involve her Type 2 mind to some degree to complete this action. Shares may require some

amount of cognition but not always, because Facebook users may share a post without clicking the link and reading through it simply because they know their friends well enough to expect them to enjoy it based on the headline alone. On Twitter, for instance, six out of ten stories were shared without being clicked through (Gabiellkov et al., 2016). On the other hand, if people only do a like or a share after reading the linked material, then there is also some delay following initial exposure to the post and further potential for Type 2 mind engagement, because clicking behaviors reflect a decision to engage in a source's content beyond what is revealed in a post.

In addition, negative feedback may require Type 2 mind involvement to complete, because a Facebook user has to employ more cognitive resources to make a decision on how to react to a piece of unfavorable content because, as there are more options to choose from as well as long term consequences to take into account. For example, after seeing objectionable content, a Facebook user may have an immediate impulse to ask Facebook to never send any more posts from the same source again. On the other hand, the user may have liked and benefitted from the posts from the same source in the past and his unsubscription of the Facebook Page would be cutting off those types of posts too. At the same time, if a user criticizes harshly the viewpoint in this post, would the criticism appear aligned with his Facebook friends' views or rather offensive to them? Therefore, negative feedback of the most extreme sort involves tradeoffs that have to be weighed and go beyond a mere immediate response to a specific piece of content unless the user is unaware that his "negative feedback" entails longer-term implications. As such, a Facebook user may only unsubscribe to a news source after concluding on the basis of multiple posts that on balance he no longer benefits from posts from that news source. Based on the above

reasoning, negative feedback should be less automatic and cognitively costlier for Facebook users.

Furthermore, as simple as likes and shares may appear, they may require more cognitive resources to be carried out when they are used to pursue certain types of objectives, as I will discuss in a following section on reasons for Facebook use, including need to belong and need for self-presentation. That means a Facebook user may start an action with an initial Type 1 mind impulse and then delay and engage in more thought about whether to complete this action. In this scenario, the delay marks a transition from Type 1 processing to Type 2 processing. As discussed above, a sharing behavior may turn from a Type 1 process to Type 2 when a Facebook user clicks a link, reads through the story and shares it based on his conclusion that his sharing this story would benefit himself or his Facebook friends for various reasons.

The effect of Type 2 processing should have a bigger moderating effect on negative feedback and comments than for likes and shares and that perhaps the longer term considerations for negative feedback might give rise to further moderation, although it is hard to know whether this involves more cognitive processing than writing a comment. Also it is not clear how likes and shares should be ordered relative to each other in this regard, because various factors, such as visibility and cognitive cost, are involved which makes it nearly impossible to make clear-cut predictions. For example, shares are more visible to other users than likes and comments and all these three are more visible than negative feedback, and the differences in visibility could influence the likelihood that these behaviors occur.



Based on the above discussion, we can see that Facebook users' behaviors can be engaged by one of the two minds or both through transitioning between them and through collaboration. The likelihood of a statistically significant coefficient would be greater the simpler the activity and the less likely there is to be Type 2 mind involvement before completing the action. The magnitude of the effect (independent of statistical significance) should also be greater when an action is more involved with one of the two minds rather than a mix of both, because involvement of the Type 2 mind dilutes the effect of Type 1 mind reactions to cognitively easy stimuli and this dilution varies among the behaviors examined in this study. Therefore, this moderating effect is fairly difficult to isolate and predict. In fact, we will see in the final section that some of the five studied behaviors are positively correlated with an independent variable, such as readability or content topic, hypothesized to influence Facebook users' behavioral responses to posts, while others are negatively correlated with the same IV, possibly because of different levels of involvement by the Type 1 and Type 2 minds.

### **Type 1 processing and cognitive ease**

Because people are facing so many media choices, they must constantly make decisions on media selection and avoidance, and this cognitive task has been examined using the dual process theoretical framework. For example, Strack and Deutsch suggest that media use is affected by the two types of intertwined cognitive processes (2004, p. 209). Because immediate urges arising from the Type 1 mind can override action plans formulated by the Type 2 mind (Evans, 2008; Stanovich, 2011), people may not be able to resist some kinds of media offerings, e.g., cute cat videos, celebrity news and reports on dramatic events, even when they intend to do so. In fact, some people frequently break from their work to check out Facebook under the influence of a strong habit (Ouellette & Wood, 1998; Kim, LaRose & Peng, 2009).

Lang (2000) investigated habitual media use in the context of television watching and found some viewers would develop *orienting responses*, which she described as an “automatic allocation of resources to a medium as a reaction to novel or interesting stimuli such as sound effects, visual complexity, movements, cuts, and zooms, presented in a television program” (Lang, 2000, p. 37). Lang’s observation supports Shoemaker and Cohen’s hypothesis that “human brains are ‘hard-wired’ to prefer information about oddities, threats, and change, and these forms of deviance are found in the news media of countries around the world” (Shoemaker & Cohen, 2006). These same types of information items appear on Facebook and, with so many items to process, one would expect that at least some initial sorting would be handled by the Type 1 mind.

In sum, endless media offerings and strong media use habits lead me to expect that Facebook users’ selections of posts will rely to a substantial degree on Type 1 processing and therefore be influenced by message characteristics that affect the ease with which these messages can be processed. If Facebook users’ selections from the posts that show up in their News Feeds are determined primarily or even substantially by Type 1 processing, it is reasonable to expect that posts with characteristics that facilitate Type 1 processing would be favored in these selections.

As discussed in the subsection on gatekeeping, Shoemaker and Vos (2009) stressed that message characteristics should be explored more because they play a crucial role in gatekeeping processes. Further, they pointed out “we need to progress beyond the categorization of messages (such as human interest, economy, international issues) to develop a number of continuous dimensions on

which messages can be measured, and this will add much to our ability to predict whether and in what form a message passes through a gate” (Shoemaker & Vos, 2009, p. 135). Following their advice, this study examines the relationships between an array of variables, such as text length, readability, and sentiment<sup>4</sup>, hypothesized to facilitate or otherwise induce Type 1 processing and the likelihoods that Facebook users will respond to NewsHour posts with the Facebook enabled actions of share, comment, like, link click and negative feedback.

### **Text length**

The notion that short text helps with information propagation has been around for well over a century. In explaining crowd psychology, Gustave Le Bon (1895) offered three techniques for mobilizing the masses for collective action: concision, repetition and contagion. While the 140-character limit on Twitter may appear insufficient for any meaningful conversation, seen through the Le Bon crowd psychology lens, the concision enforced by the 140-character limit could expedite information diffusion. Szell, Grauwin, & Ratti (2014) found that most messages on Twitter contained 70 to 120 characters, whereas those retweeted 200 times as frequently as the sample average were only about 25 characters in length on average. Focusing on Facebook, Malhotra, Malhotra, and See (2013) found that a shorter message did not influence the number of shares but increased the frequency of likes. In addition to investigating short text on social media, other scholars have also examined full-length articles. For example, Berger and Milkman (2012) found text length of articles on the New York Times website was significantly associated with number of shares via email, with every 1,000 more words related to 77% more shares. Contrary to the negative effect of message length for Facebook messages, the effect of article length on sharing for New York Times articles may be positive because people reading NYT articles are

---

<sup>4</sup> Sentiment is a measurement that quantifies positive and negative expressions of emotions, evaluations, and stances (Wilson, Wiebe & Hoffmann, 2005).

engaged in Type 2 processing when doing so and longer articles may on average be appreciated more for being more informative. On the other hand, short posts are intended to cue readers to the nature of the content and because they are scanning multiple posts they likely are relying more on Type 1 processing for scanning, which should be made more difficult by longer text. In this study, I ask how length of text for NewsHour posts is correlated with the volumes of shares, comments, likes, link clicks and negative feedback elicited by PBS NewsHour's posts on Facebook and expect the relationship to be negative in each case.

### **Readability**

Orwell (1946) stresses that simple words enhance language as an instrument for expressing thought. We can't know whether Orwell had some intuitive analogue to depletion of cognitive resources in mind, but, like text length, text that uses words easier to process should promote Type 1 mind selection of posts for further attention. To measure the ease with which language can be understood, a variety of readability measures have been developed for different purposes, including some developed especially for American English. One is the Flesch Reading Ease score (Flesch, 1948), which is a linear combination of the average number of syllables per word and the average number of words per sentence. Readability is defined as how accessible a given text is to an intended audience (Flesch, 1948). The higher the score is, the easier the text is. A Flesch reading ease score is calculated as follows:

$$Flesch = 206.835 - 1.015 \times \frac{total\ words}{total\ sentences} - 84.6 \times \frac{total\ syllables}{total\ words}$$

**Table 2: Flesch Reading Ease scores (Flesch, 1948)**

Reading Ease Score	Description of style	Representative publication
90–100	Very easy	Comics
80–90	Easy	Pulp-fiction
70–80	Fairly easy	Slick-fiction
60–70	Standard	Digests
50–60	Fairly difficult	Quality
30–50	Difficult	Academic
0–30	Very difficult	Scientific

Flesch scores have been used for a number of academic studies. For example, based on Flesch scores, D'Alessandro, Kingsley & Johnson-West (2001) found online materials for pediatric education were four grades higher than the reading level of the intended audience. On Wikipedia, a knowledge sharing website freely available to everyone, the complexity of the language is also a barrier to comprehension. Measured with the Flesch formula, regular English entries averaged 51 and were considered fairly difficult for a general audience (see Table 2; Lucassen & Schraagen, 2011). In their study of Twitter, Tan, Lee, and Pang (2014) found messages with higher reading ease scores received more retweets. Considered by themselves, the literature just reviewed supports the following hypotheses regarding text length and reading ease:

**H1a: The quantities of shares, comments, likes, link clicks and negative feedback elicited by NewsHour posts are negatively associated with the number of words.<sup>5</sup>**

---

<sup>5</sup> The number of words is calculated as the sum of the three parts of a Facebook post: message, name and description. These elements will be elaborated on in Chapter 4 on research methods.

**H1b: The quantities of shares, comments, likes, link clicks and negative feedback elicited by NewsHour posts are positively associated with their Flesch Reading Ease scores.**

In addition, the associations predicted by the above two hypotheses should be stronger and larger in magnitude for likes and shares than for comments, negative feedback and link clicks.

Although theories and prior research suggest that shorter and simpler language would be more favored by the Type 1 mind, as discussed earlier in the section on Type 1 processing and cognitive ease, there are a number of situations where user behavior on Facebook is preceded by both types of processing. When this occurs, transitions from one mind to the other and collaboration between the two minds makes it more difficult to predict, and formulate hypotheses regarding, the relationships between content characteristics and the Facebook behaviors examined, because both processes are at work and the data I collected can be used only as indirect indicators of the way Facebook users are processing the information in posts. We can, however, look for the moderating effect of co-involvement of one type of processing on the predicted effect of a content characteristic for behavioral responses when the only the other type of mind is involved. As a general matter, if Type 2 processing is a cognitively costly activity, then to the extent that a message has characteristics that predict Type 2 mind involvement, the effects of message characteristics that facilitate Type 1 processing on the likelihood a Facebook activity occurs should be diminished.

For example, if, other things are held constant, we expect shorter text length to increase the likelihood of use for one of the five Facebook behaviors because it makes Type 1 processing

easier, then the positive influence of text length reduction on the likelihood of engaging in that activity should be diminished the larger is the role of Type 2 processing in determining whether that activity occurs. That is, assuming a common metric can be applied, we should expect the measured effect size (a regression coefficient for this study) for message length to decline the bigger is the role of Type 2 processing in determining a user's response to a post and the statistical significance of measured effects should decline as well (as reflected in larger p-values).

### **Emotion and decision making**

Besides intuition and habit, emotions also play an important role in Type 1 processing, and emotional responses often compete with slower and more reflective Type 2 cognitive processes. While intuition and habit may suggest inaccurate or unwise decisions and behaviors that we fail to examine and reject, emotions can take over our thoughts and control our actions from time to time (Evans, 2010, p. 18). In other words, both the Type 1 and Type 2 minds can be dominated by strong emotions (Evans, 2013, p. 166).

All emotions have polarity, either positive or negative, and basic emotions include pleasure (or happiness), anger, fear, disgust and sadness. Although positive and negative emotions may appear to occupy opposite sides of a single emotional spectrum, there is considerable evidence that they are rather independent processes (Russell & Carroll, 1999). Cacioppo and Berntson (1994) have found positive and negative emotions are actually based on different neurological processes. Staats and Eifert (1990) corroborated this hypothesis by identifying two differing control mechanisms in the human brain: One controls approach behavior (positive emotional response) and the other controls avoidance behavior (negative emotional response). They have

also discovered that these distinct mechanisms are co-located with the reward and punishment centers in the human brain.

That emotions are more strongly connected to Type 1 processing than to Type 2 processing has been observed in various research fields. Evidence from psychological experiments by Epstein (1994) support this association. Lieberman (2003) has identified neurological regions where emotions cluster with Type 1 processes. Hassin, Uleman, and Bargh (2005) have specifically attributed emotions to automatic processes while studying social cognition.

Further, the reflective Type 2 mind sometimes may not even correct the misjudgments of the intuitive Type 1 mind but instead seek justification for them, a process called *confabulation* (Evans, 2009b). According to Kahneman, the reflective mind is “more of an apologist for the emotions” coming out of the intuitive mind than “a critic of those emotions” and it is “an endorser rather than an enforcer” (2011, p. 103). That means the reflective mind can be undemanding and help the emotional mind construct coherent stories by suggesting supportive facts and opinions of other people. In fact, while being called to engage in costlier cognitive tasks, a person’s Type 2 mind often restricts itself to the information and arguments that are consistent with that individual’s pre-existing beliefs instead of making an effort to curb his emotions and review his reasoning and decisions.

The PBS NewsHour audience is well-educated and, compared to the average Facebook user, they may have been better trained to reflect on and, when appropriate, override their intuitions and emotions. But as many scholars have recognized, humans are bound by their habits and emotions most of the time and education and intelligence come to the rescue to a very limited



extent. Stanovich (2009b, 2009c) coined the term *dysrationalia* to describe the phenomenon of smart people making unwise decisions, either because they fail to engage their reflective abilities even though they have high general intelligence, or because they have a personality that tends to be influenced by peers who may not be good thinkers.

### ***Framing***

Kahneman describes dysrationalia as “frame-bound” rather than “reality-bound” (2011, p. 367). Specifically, he has shown that colorful words can be used to sway people’s emotions and eventually decisions. For instance, people may react differently to “Italy won” and “France lost”, although logically the two phrases describe the same result for a game between Italy and France. Kahneman offers this as another example illustrating the difference between the two minds and points out that even trained professionals can be influenced by the emotions elicited by different framings. In one study, for example, physicians were asked to choose between two radiation treatments, one with a survival rate of 77% and the other with a mortality rate of 23%; the survey result showed that the majority of the participants preferred the one with the survival frame, although the two statements are logically equivalent (McNeil, Pauker, Sox Jr, & Tversky, 1982). Media professionals know especially well the power of framing, because it comes into play in generating attention for news (Scheufele, 2006). Therefore, newsrooms try to give a “spin” to stories, while “taking into account their organizational and modality constraints, professional judgments, and certain judgments about the audience” (Neuman, Just & Crigler, 1992). Indeed, people click more sentimentally charged stories when they read news online (Hensing, Flaounas, & Cristianini, 2013).

People not only respond differently to content that elicits different emotions (positive or negative), but also to the intensity of the emotions elicited. In their study of news content, Berger and Milkman (2012) manually content analyzed 6,956 articles collected from the home page of the *New York Times* website and examined the frequency of email shares. They hypothesized that higher levels of emotional arousal would be associated with more email shares after controlling for other factors like position on the front page and word count. Their findings supported their hypotheses, in line with prior studies of interpersonal communications conducted in offline settings (e.g., Rimé, Mesquita, Boca & Philippot, 1991; Peters, Kashima & Clark, 2009). Berger and Milkman's study was carried out in 2008, when the investigators focused on news websites rather than social media and on messages narrowcasted via email rather than broadcast via social media. Later, Stieglitz and Dang-Xuan (2013) found that emotionally charged tweets tended to be retweeted more often and more quickly than emotionally neutral ones.

### ***Measuring the emotional content of text***

Hu and Liu (2004) compiled a dictionary of polarized words using online reviews of products and created and tested a measure for the emotional content of words using sentiment analysis. Sentiment analysis is one application of natural language processing (NLP), in which the essential purpose is “to identify how sentiments are expressed in texts and whether the expressions indicate positive (favorable) or negative (unfavorable) opinions toward the subject” (Nasukawa & Yi, 2003, p. 70). Hu and Liu called their measurement a polarity score, which indicates how polarized a word is either in the positive (e.g., beautiful) or the negative (e.g., painful) direction. Although Hu and Liu's opinion words were extracted from product reviews, testing has shown their polarity scores to be similar when they appear in news content and they

have been employed in various media studies (e.g., Chowdhury, Routh & Chakrabarti, 2014; Zhang, Chen, Härdle, Bommers, 2015). This approach has been implemented and integrated in qdap (Qualitative Data Analysis Program) by Lu and Shulman (2008) and is freely available for the R environment.

To evaluate the polarity of a sentence, qdap first identifies polarized words based on Hu and Liu's dictionary, extracts a context of four words before and two words after a polarized word, and examines whether they shift the polarity and strength of the word's sentiment. Four types of modifiers are recognized. Negators, such as no and not, flip the polarity; amplifiers, such as very and extremely, increase the valence; de-amplifiers, such as somewhat and kind of, decrease the valence; and neutral words do not affect the direction or valence of polarized words. For each sentence, polarized words and the four types of aforementioned modification words are detected and together they determine the polarity score of a sentence using the formula proposed and validated by Hu and Liu (2004). If there are both negative and positive polarized words in one sentence, they will cancel each other's polarity. For example, the following sentence generates a zero polarity score: "I love vanilla ice cream but hate chocolate ice cream." This sentence has no modifiers while "love" and "hate" are polarized words and cancel each other's polarity. From the perspective of Hu and Liu (2004), this sentence in whole expresses neither positive or negative polarity. For each post on Facebook, I calculated the average polarity score for all the included sentences in the post and used the average of their scores as the polarity measurement for the post.<sup>6</sup>

---

<sup>6</sup> "polarity {qdap}". Inside-R. from <http://www.inside-r.org/packages/cran/qdap/docs/polarity>

## **Media avoidance**

Strongly emotional content not only elicits news consumption but may also lead to avoidance of some media content, which Woltman Elpers, Wedel and Pieters (2003) demonstrate is an integral part of media use. Woltman et al. exposed a group of subjects to a running television program and found some of them decided to avoid the program due to specific cognitive or emotional reactions evoked by its content. Communication researchers have interpreted these avoidance behaviors as possible responses to negative emotions the content might elicit (Schmidt-Atzert, 1995; cited in Fahr & Böckling, 2009) and argue they serve as a protective mechanism that helps people protect themselves from excessive psychological damage (Updegraff, Gable & Taylor, 2004).

Sending negative feedback is a specific type of media avoidance on Facebook. Facebook negative feedback features three user activities: Hide this post, hide all the posts from this Facebook Page, and unfollow this Facebook Page. These behaviors demonstrate increasingly strong desire to avoid content from a specific source.

## **Negativity bias**

As discussed earlier, News coverage is jointly determined by newsrooms, audiences, and advertisers because news topics are selected for coverage based on newsrooms' assessments of the audiences they will attract along with advertisers' willingness to pay for access to the audiences attracted by different types of content and (perhaps) audience members' willingness to pay for the coverage (Giddens, 1984; Hamilton, 2004). Mass media systematically report negative events with a higher frequency than their occurrence in real life, as discovered in the cases of employment, inflation, and interest rates (Soroka, 2012). The underlying reason, as

Trussler and Soroka (2014) discovered, is because people prefer to read negative news. Trussler and Soroka attribute this preference to *negativity bias*, which posits that negative events attract more attention from audiences than positive ones (Rozin & Royzman, 2001). According to (Pratto & John, 1991), it is much easier for people to identify negative stimuli and they are quicker to recognize negative words (Dijksterhuis & Aarts, 2003), because humans evolved to attend to negative signals critical to survival during the hunter-gatherer era. In modern societies, people still seek negative news reports for information crucial to avoiding potential risks.

However, to maintain good social relationships, negative news is often ignored while positive news is shared in everyday interactions (Maynard, 2003). This practice is not only observed in offline settings but also in an online setting like Facebook, which is in line with the findings of the studies on the need to belong discussed earlier. For these reasons, people may use the “comment” function to express their thoughts and feelings on negative stories in relatively impersonal space and not as part of their more personal interactions with Facebook friends. It is important to control for polarity and valence in examining factors that influence the five dependent variables in this study. Therefore, polarity measures are included in the regression equations described in the next chapter.

The literature reviewed above provides plausible reasons for expecting user behaviors on Facebook to increase with polarity due to fast Type 1 processing, such as framing (Scheufele, 2006). At the same time, it is also plausible to expect these behaviors to decline with polarity, as posited by negativity bias theory for instance (Trussler & Soroka, 2014). While the effect of polarity on these behavioral measures is an interesting topic, the research literature does not

support an unambiguous prediction. Therefore, I am addressing this topic with a research question rather than a hypothesis.

**RQ1: How do the quantities of shares, comments, likes, link clicks and negative feedback elicited by NewsHour posts vary with polarity of a post?**

### **News preferences**

Communication scholars believe that the motivations for news consumption are often explained by the anticipated benefits of post-exposure applications of information gained from the news to be consumed. “[A] message has instrumental utility for the receiver when it provides him with a helpful input for responding to everyday environmental stimuli or for defending personal predispositions; he may need information to keep abreast of governmental actions, to guide his consumer decision-making, or to reinforce his political preferences” (Atkin, 1973, p. 205). While this observation should also apply to news items posted on social media, as with other media, we should expect news items posted on social media to address a variety of user needs, including for entertainment. While tabloids constantly post content on entertainment and lifestyle on Facebook, as expected, national newspapers also include softer content in the mix with their hard news offerings to better fit with the social climate on the platform (e.g., Rainie & Smith, 2012; Pew Research, 2012). For instance, the *Washington Post* has posted several videos of animals taken at the Smithsonian’s National Zoo and a post of pandas rolling in snow has gained more shares and likes than news about most politicians in Washington. The *New York Times* and other major news outlets are also not laggards in employing cuteness.

The explanation for why major news producers like the *New York Times* and the *Washington Post* provide a medley of different types of news content may be provided by “the duality of structure” hypothesis, which describes “how people use the resources offered by the media

environment to enact their preferences and, in doing so, shape the very structures within which they operate” (Giddens, 1984, p. 24). This phenomenon is observed for both linear media, such as radio and television, and non-linear media, such as DVRs and online media services (Webster, 2009, p. 221). From this perspective, editorial choices to an extent reflect the preferences of audience members who want to consume multiple types of content. For example, although people claimed to regularly read political news, Nielsen log data has revealed that they in fact read more about sports and entertainment than about public affairs (Tewksbury, 2003). Focusing on 11 major news websites including BBC, NPR, and Yahoo! News, Hensing, Flaounas, and Cristianini (2013) similarly found that financial and political news topics were visited less frequently than softer news topics. Shifting focus from news sites to social media, Hargittai and Litt (2011) found that on Twitter young adults were attracted to celebrity and entertainment news but showed evident dislike for science, politics, local, national and international news.

Other scholars, however, discerned quite different patterns of online news readership. Neuberger et al. (1998) conducted one of the earliest studies by examining the online presences of 81 daily newspapers in Germany and found that readers of regional newspapers preferred local news whereas readers of national dailies preferred political and business news. Similarly, D'Haenens, Jankowski, and Heuvelman (2004) found that online readers preferred international news to sports. Expanding online readership research to the realm of social media, Bastos (2014) found that readers of *the Guardian* and *the New York Times* preferred to share opinion pieces along with national, local, and world news as opposed to pieces on sports and entertainment. Moreover, Bastos found that social media users preferred to share different types of news topics using

different social media vehicles, with Twitter preferred more for political news, StumbleUpon and Delicious for science and technology, and Pinterest for fashion, arts, lifestyle and entertainment.

In fact, social media users' preferences for news items can be affected by factors beyond their intrinsic interests because Facebook, Twitter and the like are shifting news consumption from a solitary activity to a social activity. As such, it is also plausible that Facebook users read and share news not only to satisfy their own needs for information, as posited by Atkin (1973), but also but also to facilitate more social interactions with other users.

### **Reasons for Facebook use**

Facebook has created an environment to facilitate users' creation, consumption and sharing of various types of content. Vernuccio describes such environments from the perspective of social media users as "[p]latforms of digital communication that continually appear in their interactive environment, underlining their participative and collaborative social characteristics" (2014, p. 213). Prior studies have recognized a variety of reasons and motivations for Facebook use. For example, Sheldon, Abad, and Hirsch (2011) examined Facebook use from the perspective of self-determination theory (Ryan & Deci, 2000), which posits that humans seek to satisfy three innate psychological needs: competence, autonomy, and relatedness. When these needs are satisfied, self-motivation and mental health are both higher, and when not, motivation and well-being may be impaired. Applying this framework, Sheldon et al. found Facebook users coped with disconnection in real life by making satisfactory connections on the social media platform. In addition, media use has been studied using uses and gratifications theory (Lasswell, 1948), which assumes media consumers are not passive but actively use media to achieve their own gratifications (Levy & Windahl, 1985). Applying this framework, Diddi and LaRose (2006)



discovered two gratifications that predicted college students' online news consumption: one was a need for in-depth stories and local news and the other was a desire to escape reality and pass time. In the same vein, Lee and Ma (2012) investigated the motivations behind news sharing on Facebook and identified three dimensions (information, socializing, and status seeking) through a factor analysis of the survey responses. In particular, the information dimension refers to contributing "relevant and timely information" and facilitates future information seeking through reciprocated sharing activities; the socializing dimension refers to a desire to "develop and maintain relationships with acquaintances" and maintain a connection to a virtual community through news sharing; and the status seeking dimension describes as a goal to "attain status among peers" and boost self-esteem and confidence (Rubin & Hewstone, 1998).

### ***Need to belong***

Based on their meta-analysis of 42 empirical studies of Facebook, Nadkarni and Hofmann (2012) identified two major motivations for Facebook use: a need to belong and a need for self-presentation. The need to belong refers to an intrinsic desire to connect with others and obtain social acceptance, and the need for self-presentation involves an ongoing process of impression management. These two motivations can act independently or jointly and are also shaped by other factors, such as cultural background, demographics, and personality traits. With regard to interactions with news publishers on Facebook, these two motivations are likely to be closely related to sharing behavior because shared posts are displayed on Facebook users' Timelines and thus expose their interests, thoughts and ideas to their friends, which may improve or undermine the users' relationships with their friends. By contrast, comments on news posts on Facebook Pages are not displayed on Timelines or News Feeds, where Facebook users browse information, and thus are not immediately visible to Facebook users' friends. Due to the different spaces

where comments and shared posts are displayed, a user may feel more comfortable commenting on a controversial post on a Facebook Page than sharing it on his Timeline.

Research on sharing political news has shown that political news is more likely to cause disputes among friends than soft news, such as art and entertainment, and thus is less likely to be shared on Facebook. For instance, when asked “have you ever decided NOT to post a political comment or link on a social networking site because you were worried it might upset or offend someone?” 77% out of 1,047 participants answered they had decided not to post something for the reasons stated; when asked about how they would respond to posted opinions they did not agree with, 66% of surveyed users stated that they would typically ignore the posts without doing anything on social media (Rainie & Smith, 2012). Additionally, according to Pew Research, 58% of Facebook users and 59% of Twitter users felt somewhat unwilling to discuss the Snowden-NSA case on the two platforms.<sup>7</sup>

### *Need for self-presentation*

Besides a sense of belonging, people are also concerned with their self-presentations on social media. Zhao, Grasmuck, and Martin (2008) called the identity constructed on Facebook a “social product”, which was far from an individual characteristic or an expression of something innate in a person. They attributed this outcome to the diminishment of problems that may hinder use of face-to-face communication, such as appearance and shyness (McKenna, Green, & Gleason, 2002) and research has found that online users may “stretch the truth a bit” (Yurchisin, Watchravesringkan, & McCabe, 2005). In other words, people are able to present themselves on social media as having socially desirable qualities, including being popular, well rounded, and

---

<sup>7</sup> Anderson M. and Caumont A. (September 24, 2014). "How social media is reshaping news." Pew Research. from <http://www.pewresearch.org/fact-tank/2014/09/24/how-social-media-is-reshaping-news/>

thoughtful, and some of them may not in fact possess the qualities they claim (Zhao et al., 2008). Focusing on another aspect of idealized identities, Peluchette and Karl (2010) investigated how students managed their public images on Facebook. They found that students who posted provocative information perceived themselves as portraying a “sexually appealing, wild, or offensive image” and those who did not do so considered themselves portraying a hardworking image. In fact, self-presentation on the Internet had been studied prior to the popularity of Facebook. For example, personal website owners may follow the motto “we are what we post”, because they can manage their online images by presenting brand logos and products as they wish, as opposed to a real life with financial constraints that preclude consumption of those brands (Schau & Gilly, 2003). On dating websites, users tend to carefully select their photos and some of them even heavily retouch them, thereby misleading other users (Hancock & Toma, 2009).

In addition to crafting overly flattering profiles, creating and sharing content is another strategy to construct ideal identities on Facebook, in a similar way that brand logos were exploited by personal websites owners. For example, complex circuits and charts on *Wired*’s Facebook Page are often highly shared. For the *New York Times*, posts that are smart (e.g., latest scientific findings) and hip (e.g., new diets and exercise routines reported to enhance health) are among the most shared on Facebook. Besides providing relevant articles to friends, sharing stories on Facebook may also improve self-presentation through “conspicuous consumption” of media content to demonstrate taste and social class. Conspicuous consumption refers to public display of one’s consumption to impress others with one’s social status and economic power (Veblen, 2000). Particularly, Shipman (2004) noted individuals conspicuously consume cultural products

to show off the education that enables them to appreciate such highly sophisticated culture products, which tend to be unavailable or inaccessible to the majority. For instance, art and science stories require more education to appreciate and comprehend and therefore are seen as evidence of sophistication.

Indeed, people with different preferences and social reference groups may share, like, and comment on different types of content. For example, sharing science reporting may serve the self-presentation needs of someone whose Facebook friends value this kind of knowledge, while sharing entertainment news may enhance self-presentation and social connections for someone whose friends are more interested in what is happening in popular culture. Thus, there is good reason to believe that different news topics may influence the dependent variables, and thus posts' topics should be included as controls for this analysis even though, because I have no independent measures of individual users' content preferences, there are also no clear-cut predictions related to self-presentation that can be tested with my dataset. However, the topic controls can be seen as addressing the following research question.

**RQ2: How do the quantities of shares, comments, likes, link clicks and negative feedback elicited by NewsHour posts vary with news topics?**

## **CHAPTER 4: RESEARCH METHODS**

### **Data collection**

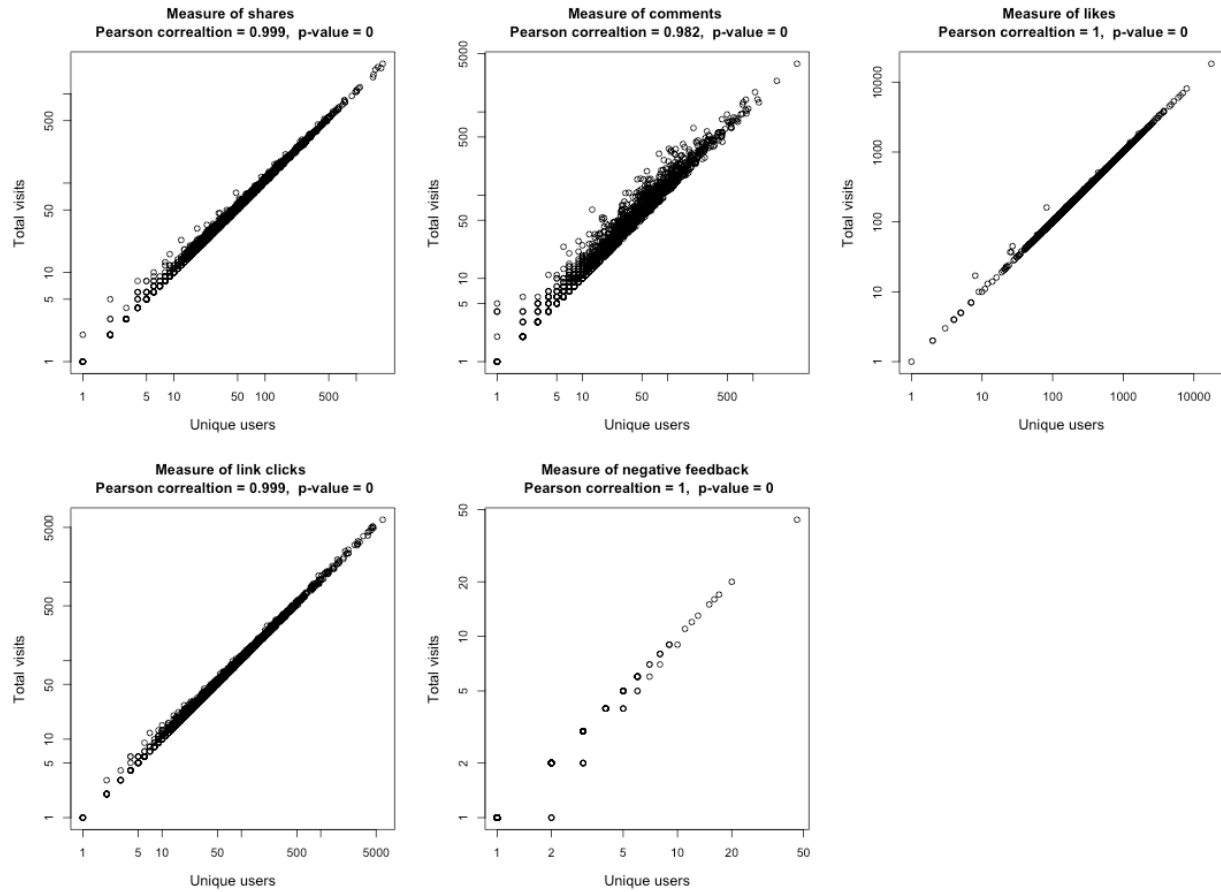
PBS NewsHour publishes posts on its “Facebook Page” and interacts with its audience through this venue as well. The interactions between Facebook users and PBS NewsHour are documented and reported by a built-in tool called “Facebook Insights”, through which I collected the data for this study. Facebook Page and Facebook Insights and the types of data collected through them were described in Chapter 2. This chapter presents the empirical model employed to test the hypotheses and research questions presented in Chapter 3.

### **Conceptual model and empirical model**

I employed activity measures reported in Facebook Insights to test the hypotheses and address the research questions set out in the preceding chapter. The dependent variables are activity measures for shares, comments, likes, link clicks, and negative feedback and they are modeled as counts. As discussed in the Chapter 2 sections on “Facebook Insights and user behaviors”, Facebook Insights activities are measured and reported in two ways, “unique users” and “total visits”. The total visits metric focuses on sessions (an uninterrupted connection between a user and a server) and considers each session an opportunity to expose that user to some content. The metric of unique users focuses on people as individuals and considers each user an opportunity to generate exposure to some content. Although the two metrics differ in some ways, they are highly correlated across various measurements of user behaviors, including all five dependent variables. Certainly, “unique users” and “total visits” have subtle differences. For example, a Facebook post may appear interesting to a niche audience and attract a small number of unique users who have each contributed multiple comments on several days while the conversation was advanced and consequently generated a number of total visits whose count is several times the

count for unique users. On the other hand, a post reporting a football game score may be read by a lot of unique users who will not come back to this post after the first visit. In this case, the count for unique users is nearly the same as the count for total visits.

For my sample, these two measures of users' behaviors are highly correlated (see Figure 8) and they generated regression coefficients with nearly identical significance and signs for each of my five dependent variables. Therefore, I am reporting the results for "unique users" only. I have explored the ratio of total visits to unique users and found they are informative in ways not directly related to my dissertation research. For example, on a snowy day, the number of unique users that are attracted by the New York Times may be similar to what it is on a sunny day, but the total visits on a snowy day could be twice that on a sunny day. That means during bad weather the same body of users would visit their favorite website more often. However, I will pursue these observations further in a future study. Taking the example in Table 1 on Facebook Insights and user behaviors, the quantity of shares for that Facebook post is 816, calculated with the number of unique users.



**Figure 8: The scatterplots for five user behaviors measured as total visits and unique users. The correlation values show they are highly correlated.**

The independent variables are measures of the following content attributes: news topic, cognitive ease, and sentiment. Two variables, number of words and reading ease, are included as measures of factors influencing cognitive ease, while sentiment is measured as the average polarity for all the sentences in a post. I also included a polarity dummy, which takes a value of one when polarity is negative and a value of zero when polarity is positive (no posts in this sample had a zero polarity score). This dummy variable was used to capture the changes in the mean of a dependent variable when polarity changes from positive to negative. (Studenmund, 2010). Its interaction with polarity (polarity dummy times average polarity) was introduced to the model as

a slope dummy (also called a spline variable) to capture any difference in the counts at which user behaviors would change in response to changes in the absolute value of polarity when polarity was negative or positive. To control for other factors that might also affect responses to content, I included number of questions, number of fans, post type, the hour, weekday and month of a post and whether an item was posted on a holiday as control variables. I elaborate on the reasons for including each of these variables in the section on control variables below.

**Table 3: The hypothesized relationships between dependent and independent variables.**

IV\DV	Share	Comment	Like	Click	Negative feedback	RQ & H
<b>Number of words</b>	–	–	–	–	–	H1a
<b>Reading ease</b>	+	+	+	+	+	H1b
<b>Polarity</b>	?	?	?	?	?	RQ1
<b>News topics</b>	?	?	?	?	?	RQ2

**Table 4: List of dependent, independent, and control variables. Reference category lists the default for the categorical variables.**

	Variable	Type	Reference Category	Operational Definition
<b>Dependent variables</b>	<b>Share</b>	Continuous	--	Number of unique users who shared a post.
	<b>Comment</b>	Continuous	--	Number of unique users who commented on a post.
	<b>Like</b>	Continuous	--	Number of unique users who liked a post.
	<b>Link clicks</b>	Continuous	--	Number of unique users who clicked a link in a post.
	<b>Negative feedback</b>	Continuous	--	Number of unique users who left negative feedback on a post. Negative feedback includes “hide this post”, “hide all the posts from this Page”, and “unlike this Page”.
<b>Independent variables</b>	<b>News topic</b>	Categorical	Uncategorized posts	Eight news topics were covered by PBS NewsHour: Arts, economy, education, health, nation, politics, science, and world. The home page and section front pages were labeled with no new topic.



**Table 4 (cont'd)**

	<b>Words</b>	Continuous	--	Number of words in a post combining three parts: Message, name and description.
	<b>Reading ease</b>	Continuous	--	The Flesch reading ease score. The higher the score, the easier is the text to read.
	<b>Average polarity</b>	Continuous	--	It measures the sentiment score of a post, which averages all the sentences in the post and reflects the general level of sentiment of a post on Facebook.
	<b>Polarity dummy</b>	Dummy	--	The polarity dummy takes a value of one when polarity is negative and it takes a value of zero when polarity is positive. There are no zero polarity posts in the sample. Its coefficient is the change in the intercept when polarity changes from positive to negative.
	<b>Slope dummy</b>	Interaction of polarity dummy & average polarity	--	Polarity dummy times average polarity. It is used to estimate differences in size and direction of effects that changes in positive and negative polarity scores have on user behaviors.
	<b>Impressions</b>	Continuous	--	The number of unique users exposed to a post from a Facebook Page, whether the post is clicked or not.
<b>Control variables</b>	<b>Questions</b>	Continuous	--	Number of questions embedded in a post.
	<b>Fans</b>	Continuous	--	Daily total of users who “liked” PBS NewsHour on Facebook. It increased monotonically over the period of data collection.
	<b>Post type</b>	Categorical	Text	Four types that were used by PBS NewsHour: Text, image, link and video.
	<b>Hour</b>	Categorical	Midnight through 8 am	The hour a post was posted.
	<b>Weekday</b>	Categorical	Sunday	Day of week a post was posted.
	<b>Month</b>	Categorical	December	The month a post was posted.
	<b>Holiday</b>	Categorical	Non-holidays	Categorical variable with value of one if the post was posted during a holiday
	<b>Trend</b>	Continuous	--	Number of days from the start of the course of the study.

Based on the literature reviewed in chapter 3, I have hypothesized that each of the dependent variables to be functions of the independent variables (number of words, reading ease, news topics and average polarity) and the control variables listed in Table 4. The basic conceptual

model is that a behavior is a function of the independent variables, or  $B = f(\text{number of words, reading ease, news topics, average polarity})$ , where  $B$  can be a count for shares, comments, likes, link clicks or negative feedback. Table 3 in this chapter presents the summary of the predicted effects of the independent variables on the dependent variables.

Equation [1] is the empirical model used to estimate this relationship, which was estimated using the generalized linear model (GLM). Descriptions and operational definitions for the variables are provided in Table 4.

$$\begin{aligned}
 \text{Measure of User behavior} = & \beta_0 + \beta_1 \text{Number of Words} + \beta_2 \text{Reading ease} + \\
 & \beta_3 \text{Arts} + \beta_4 \text{Economy} + \beta_5 \text{Education} + \beta_6 \text{Health} + \\
 & \beta_7 \text{Nation} + \beta_8 \text{Politics} + \beta_9 \text{Science} + \beta_{10} \text{World} + \\
 & \beta_{11} \text{Average polarity} + \beta_{12} \text{Polarity dummy} + \beta_{13} \text{Slope dummy} + \\
 & \beta_{14} \text{Photo} + \beta_{15} \text{Link} + \beta_{16} \text{Video} + \\
 & \beta_{17} \text{Hour9} + \beta_{18} \text{Hour10} + \beta_{19} \text{Hour11} + \beta_{20} \text{Hour12} + \beta_{21} \text{Hour13} + \beta_{22} \text{Hour14} + \\
 & \beta_{23} \text{Hour15} + \beta_{24} \text{Hour16} + \beta_{25} \text{Hour17} + \beta_{26} \text{Hour18} + \beta_{27} \text{Hour19} + \beta_{28} \text{Hour20} + \\
 & \beta_{29} \text{Hour21} + \beta_{30} \text{Hour22} + \beta_{31} \text{Hour23} + \\
 & \beta_{32} \text{Monday} + \beta_{33} \text{Tuesday} + \beta_{34} \text{Wednesday} + \\
 & \beta_{35} \text{Thursday} + \beta_{36} \text{Friday} + \beta_{37} \text{Saturday} + \\
 & \beta_{38} \text{January} + \beta_{39} \text{February} + \beta_{40} \text{March} + \beta_{41} \text{April} + \beta_{42} \text{May} + \beta_{43} \text{June} + \\
 & \beta_{44} \text{July} + \beta_{45} \text{August} + \beta_{46} \text{September} + \beta_{47} \text{October} + \beta_{48} \text{November} + \\
 & \beta_{49} \text{Holiday} + \beta_{50} \text{Questions} + \beta_{51} \text{Fans} + \beta_{52} \text{Trend} + \varepsilon. \quad [1]
 \end{aligned}$$

Slope dummy is an interaction term between average polarity and polarity dummy. When average polarity was negative or positive, polarity dummy took the value of one or zero,

respectively. If user behaviors change at differing counts for changes in negative polarity compared to changes in positive polarity, the difference will be reflected in  $\beta_{13}$  as follows:

When polarity is negative and *polarity dummy* = 1,

$$\Delta \text{ user behaviors} / \Delta \text{ average polarity} = \beta_{11} + \beta_{13} \quad [2]$$

When polarity is positive and *polarity dummy* = 0,

$$\Delta \text{ user behaviors} / \Delta \text{ average polarity} = \beta_{11} \quad [3]$$

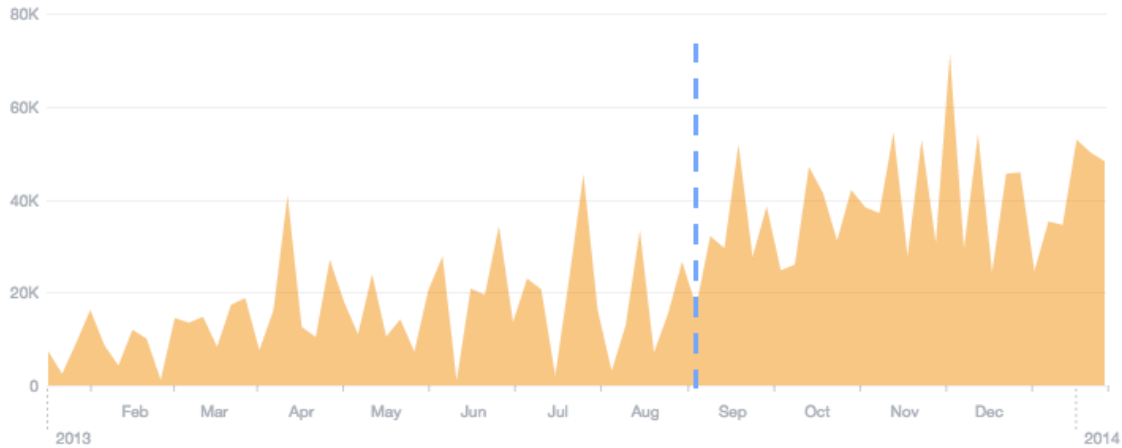
### **Independent variables: data collection and variable construction**

The beauty of testing theories with Facebook is that data on behaviors measured by Facebook can be easily collected by researchers because Facebook provides the Graph API and Facebook Query Language (FQL) for data retrieval. With permission from PBS NewsHour, I retrieved all posts from its Facebook Page from July 16, 2011 (the first NewsHour post) to October 6, 2014. For each post, I collected for the sample period the total counts for unique users and total visits reported in Facebook Insights for shares, comments, likes, link clicks, and negative feedback. A user engaged in the same activity multiple times added one to the total count for unique users but increased the count for total visits by the number of times the user engaged in that activity. For example, a user may have a debate with other users over a news story posted on the Facebook Page of PBS NewsHour and leave six comments on this post. In this case, his comments generate one count in the number of unique users and six counts in the number of total visits because he came back five times to add more comments as rebuttals to other users' comments, thereby generating six visits in total.

Both measures have their uses. For example, advertisers may be interested in the frequency with which individual users are exposed to their content as well as in the number of people exposed.

Because individual users function as gatekeepers in generating exposures from other users to content they (the original users) have encountered, measures based on unique users make the most sense for this study. Therefore, for this study, I investigated user behavior from the perspective of the individual user and employed the count of unique users for each of the posts in the sample. Each post was considered as one case for this study and the entire sample of posts was examined to identify the associations between user behaviors and content attributes. For example, to examine the effects of the independent variables on shares, for all posts in my sample share counts per unique user were placed on the left side of Equation 1 as the dependent variable, and measures for attributes of posts (text length, news topic, etc.) were included on the right side as independent variables.

I selected September 1, 2013 as the starting date for the data set used for the regression analysis for two reasons. First, PBS NewsHour did not begin publishing new material on Facebook on a daily basis until April 2013 and it was not until August of 2013 that its rate of publishing stabilized at around its current rate of about five posts per day. One would expect the number of people interacting with NewsHour content to increase with the number of posts per day. Second, although Facebook updates its algorithms quite frequently, unlike the more typical updates, the update between August and September 2013 was observed in reports by Adweek (Cohen, 2013) and ComScore (Shields, 2013) to have been followed by substantial increases or decreases in traffic counts for a number of websites, such as HuffPost, BuzzFeed, and Business Insider. As can be seen in Figure 6, the trend for NewsHour's count for impressions for unique users does appear to have jumped to a new and higher level at the beginning of September 2013 and the average has remained relatively stable since then.

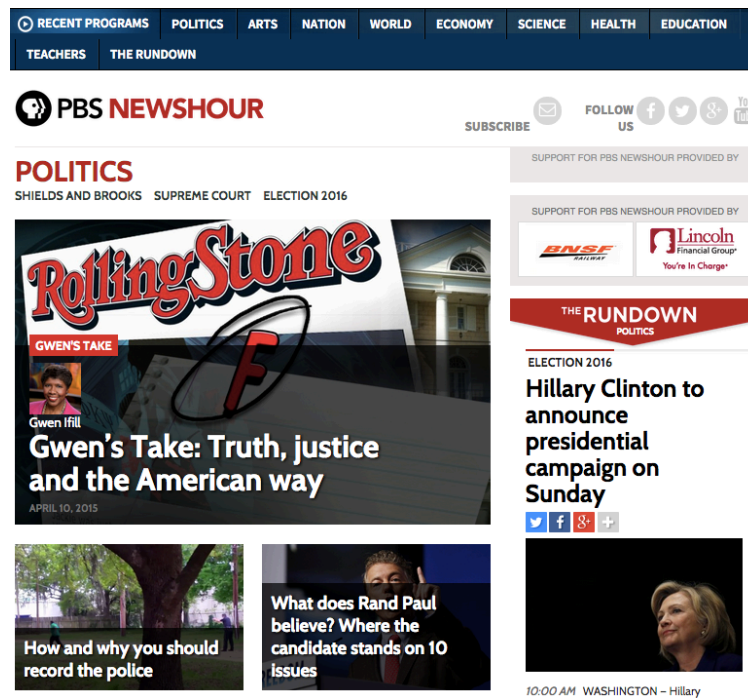


**Figure 9: A screenshot of Facebook Insights that shows the daily impressions for unique users of PBS NewsHour.**

As I observed no major turning points other than the one between August and September of 2013, I assumed that no algorithmic changes subsequently introduced by Facebook substantially affected PBS NewsHour’s readership on Facebook. Hence, I focused on the posts published after the major algorithmic update, examined the 1,778 posted between September 1, 2013 and October 6, 2014 and eliminated the posts with no text to generate the final sample of 1,734.

PBS assigns one or more of eight news topics—art, economy, education, health, nation, politics, science and world—to its posted articles and I extracted these news topics for use as article descriptors. Home page and section front pages are assigned with no news topic by PBS editors (see the politics section as an example in Figure 10) and for this analysis they were labeled as uncategorized. Among 1,734 posts under investigation, 16% were linked to home pages or section front pages, 70% were labeled with one of the eight topics, 13% with two topics, and 2% with three topics. I experimented with a control variable that measures how many topics were

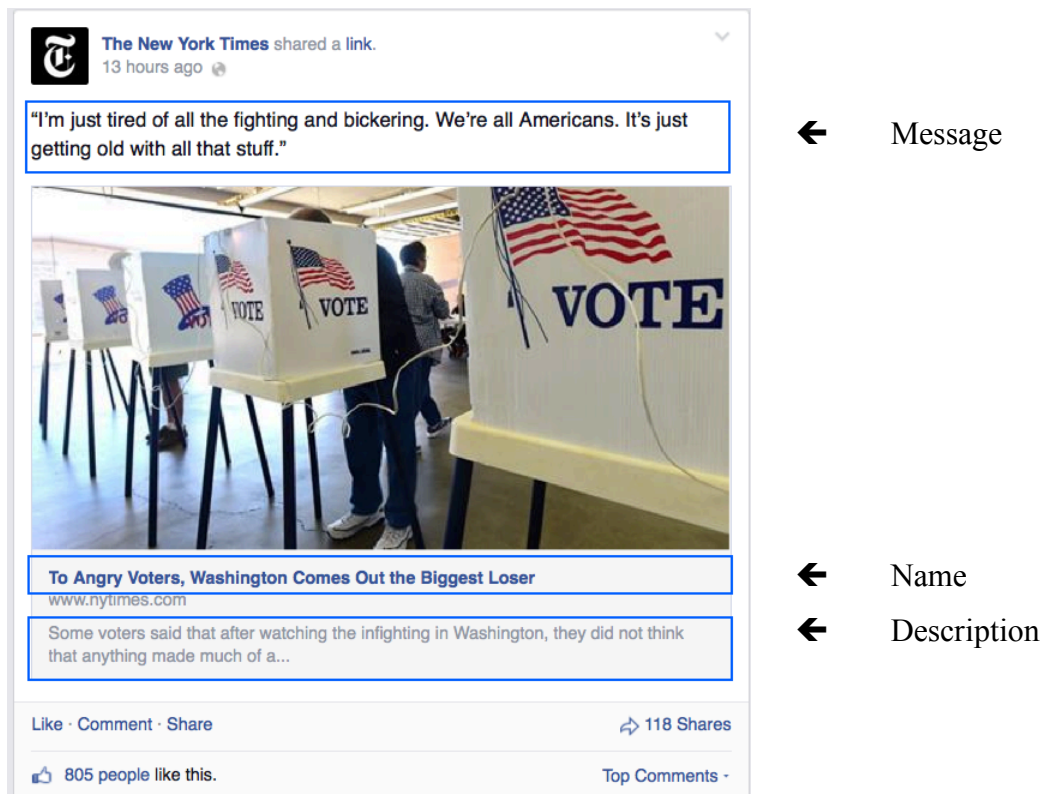
assigned to a post and found no significant relationship with any user behavior at all. For this reason, I excluded this variable from the final model.



**Figure 10:** A screenshot of the political section front page, which was assigned no topic label by PBS editors

While it would have been nice to use a formal content analysis to categorize the posts, to economize on the limited time and resources available to pursue this study I chose to adopt the categories that were already assigned by PBS NewsHour. Use of industry-recognized categories has been common in economic studies of media for decades (e.g., Wildman & Robinson, 1995; Chang & Ki, 2005; Litman & Kohl, 1989; Lee, 2006). The standard justification for this practice is that these categories have acquired meaning within the industry and that the firms classifying their own product have an incentive to conform to the rest of the industry's expectations in doing so. For this reason, NewsHour has a strong financial incentive to select categories that will have

meaning to its advertisers and to be consistent in applying their classification scheme over time. The fact that audience members may also see articles' category assignments and use them as an aid for finding articles that appeal to them is further reason for NewsHour to be consistent over time in the way it classifies its content and by itself is a reason for including NewsHour's topic assignments as control variables.



**Figure 11: A screenshot of a post on Facebook showing three blocks as message, name and description.**

Each post on Facebook has three parts: message, name and description (see Figure 11). When a “link” is posted on Facebook, the “name” and “description” elements are automatically extracted by Facebook from the linked web page. The name is generally the posted article's title and a truncated version of the first paragraph of an article is generally used for its description. Social

media editors in various newsrooms have the liberty to edit all the three elements in their posts on Facebook but sometimes leave them in the form automatically extracted by Facebook. A “message” serves as a blurb or a teaser intended to give the audience the flavor of the full content. This field will not be automatically filled in by Facebook and will be left blank if a social media editor does not write any words in it. Some editors make the effort to craft compelling “messages” to attract more attention to their posts, while other editors simply use the first paragraph in the web page reached through the posted link as the description.

It is hard to tell which element of a post will attract more attention or trigger more interactions. The differences in font size, font color, text length and the different types of information conveyed all affect information processing. The “message” is at the top, above the other two parts, but the “name” is displayed in a larger font than the other two parts, and, while the “description” is shown in a smaller and grayer font, it can be longer than the other two. I combined the three elements of a post, measured each combined text unit with the qdap package for R, and included text length, readability as measured by its Flesch Reading Ease score and sentiment (average polarity) as independent variables in my analysis.

### **Control variables**

In addition to the aforementioned independent variables, I included a number of control variables: number of questions in a post, daily number of fans for PBS NewsHour, post type, and posting time marked in the time zone where PBS NewsHour is located, i.e., Eastern Time or Eastern Daylight Time. With regard to posting time, I examined several aspects, including hour of the day, day of the week, month, year, and whether it was a holiday or not. Content providers can post questions on Facebook and Facebook users respond to questions by leaving comments as



illustrated by Figure 12. Questions are invitations to respond and if, for a post with multiple questions, some readers respond to one of the questions and other readers respond to other questions, then the number of comments will increase with the number of questions. For this reason, I have included number of questions in a post as a control variable. In addition, I retrieved the daily number of fans of PBS NewsHour for the period for which posts were sampled (September 1, 2013 to October 6, 2014) and included it as a control variable. The fan count increased monotonically during this period. I included number of fans because the increased fan count may have been in part a response to more compelling content and a better-operated page, both of which could generate more fan interactions with posts.



**Figure 12: A screenshot of questions raised in a PBS NewsHour Facebook post and fans' comment responses.**

Facebook supports a variety of formats and mechanisms for presenting content. While there are other options, PBS NewsHour used only text, photos, links to external websites, and videos during the period for which data was collected. According to Bastos (2014), photos and videos, have rendered Facebook a primarily visual medium. Prior research suggests that different presentational modalities may elicit different types of responses from users. Cvijikj, Spiegler, & Michahelles (2011) studied the Facebook Pages of 14 brands including Walmart, Coca-Cola and Disney and found posts with photos and videos to be associated with more likes and comments

than posts without photos or videos. At the opening of his book *Thinking Fast and Slow*, Kahneman states that, for the fast type of cognitive processing, visual stimuli attract more attention and gives an example for which visual materials elicited more responses than pure text (Kahneman, 2011). Therefore, photos and video posts would be expected to elicit more user behaviors than text posts. In fact, Facebook suggests its users will better respond to photos and videos, because they are colorful and elicit more user interaction (Facebook, 2015).

Because the amounts and types of competing activities vary across a day, the time of day a post appears on Facebook may also influence how much attention it will attract from users. Although media advance and evolve, people manage to settle down with fairly stable schedules for consuming their selections for media formats and content. Webster and Phalen (1997) summarized prior studies focused on audience availability. Among them, Barnett et al. (1991) found time spent watching television followed a seasonal pattern in the United States. It was lowest in July and negatively correlated with the amount of daylight and positively with the amount of precipitation. Webster and Lichty (1991) reported television viewing peaked in early evening, or prime time, on a daily basis. Further, Webster (2014) reported hourly variation in media use across different media formats, with computers, television and radio competing for people's attention. While computer use was high during office hours, television viewing was high in the evening and radio listening was high during the morning and evening drive times. Also, people tend to go to the cinema on holidays, and watch more television in the winter when weather is colder and less during the summer (Webster, 2014).

Classical conditioning suggests that media use can be stimulated by daily activities that are often paired with media consumption. For example, coffee aroma in the morning may provoke a desire to read a newspaper and dead time waiting for a class to start or boredom at work may cause people to reflexively turn to social media (Whiting & Williams, 2013). To control for variation in media use across different hours of the day, days of the week and seasons, I included variables for the time of day (Eastern Standard Time), day of the week, and month of the year. Unfortunately, I did not have the information for the time zones from which user interactions originated. To allow for the possibility that user behavior patterns were influenced by unobserved changes in the online and media environments during the period covered by the data set, I included a trend variable, which, for each day in the sample period, was calculated as the number of days from the first day of the period sampled to that day.

## Empirical distributions for variables

The distributions of the independent and dependent variables are shown in Figure 13 and Figure 14. The pie charts are for categorical variables and the histograms are for continuous variables. A correlation matrix with correlations between the independent variables and dependent variables is provided in Table 9.

**Table 5: Distribution by Eastern Time hour of the day.**

Hour	0-8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
# Posts	9	20	84	238	150	183	175	162	150	172	150	101	67	35	28	10

**Table 6: Distribution by day of the week.**

Weekday	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
# Posts	266	239	234	250	243	241	261

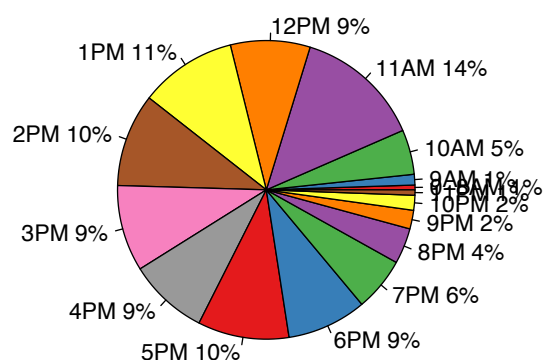
**Table 7: Distribution by month of the year.**

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
#Posts	124	108	126	118	129	135	183	170	291	129	107	114

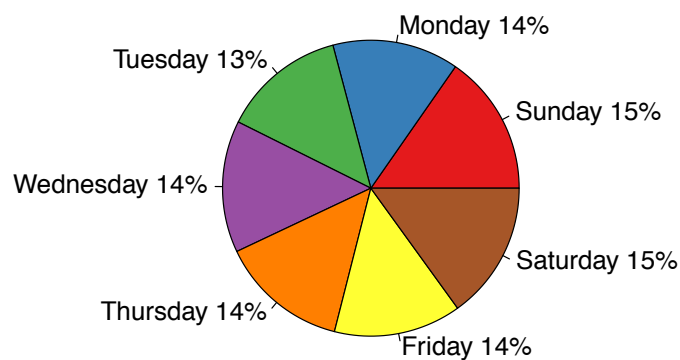
**Table 8: Distribution by content category.**

Content category	Arts	Economy	Education	Health	Nation	Politics	Science	World
# Posts	218	137	88	152	421	182	212	328

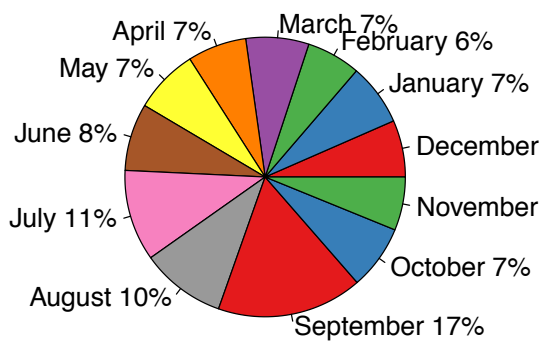
**Posts across a day**



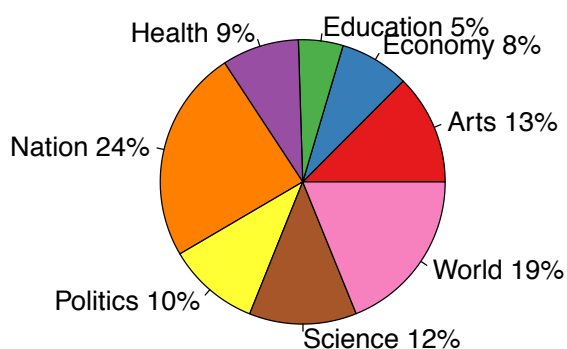
**Posts across a week**



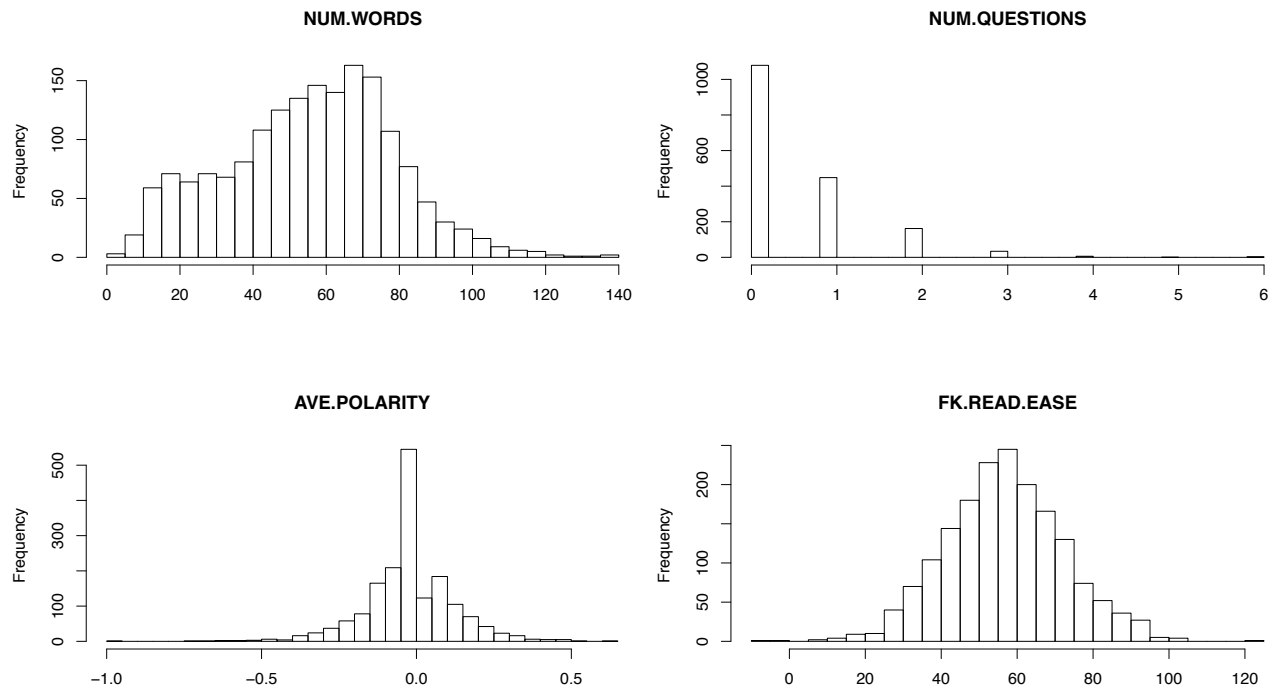
**Posts across a year**



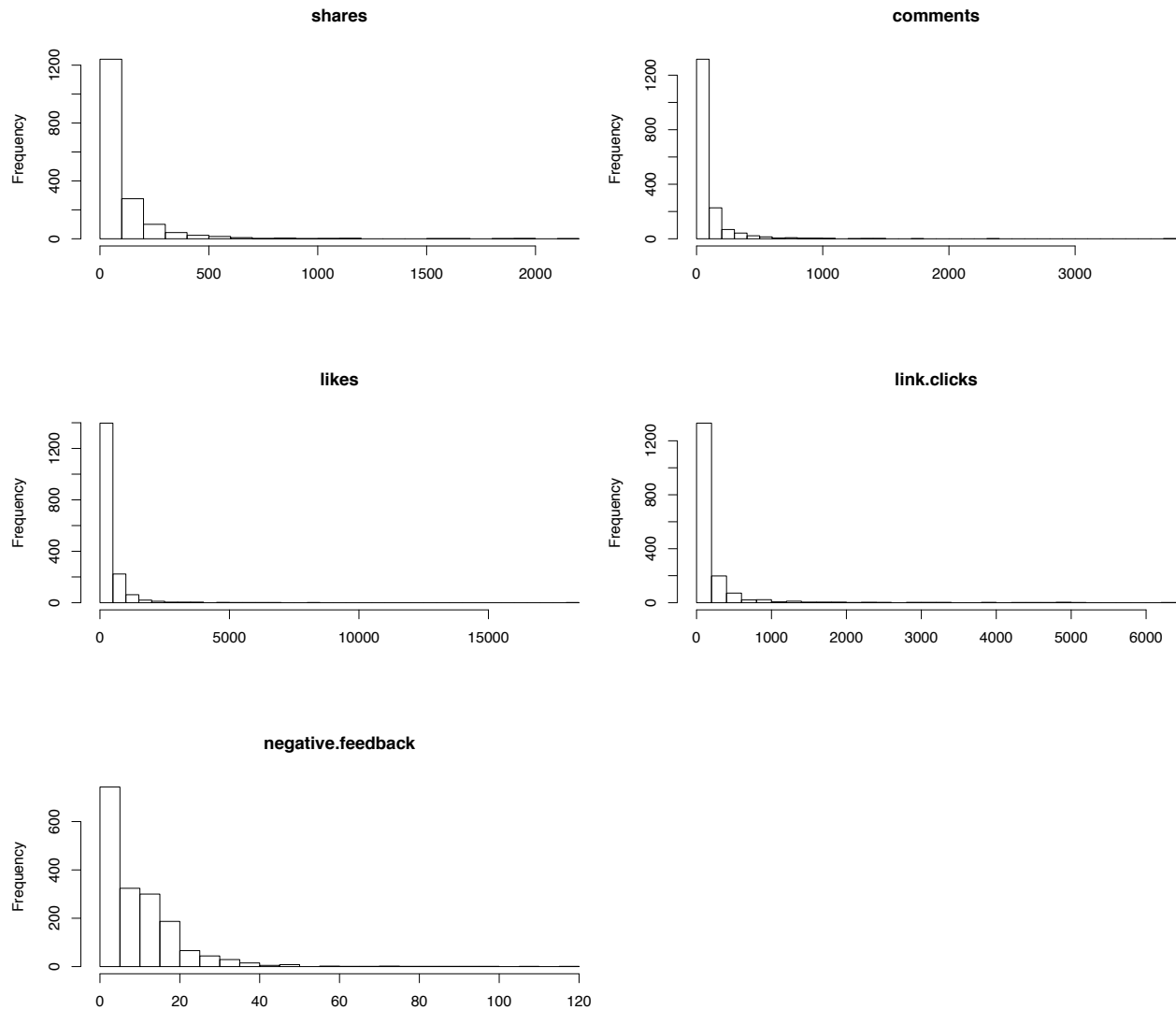
**Posts across sections**



**Figure 13: Percentage distribution of posts across a day (top left), a week (top right), a year (bottom left) and across sections (bottom right)**



**Figure 14: Distributions of independent variables in histograms. Each observation is a post that was posted by PBS NewsHour on Facebook.**



**Figure 15: Distributions of dependent variables (five user behaviors) in histograms.** Each observation is a post that was posted by PBS NewsHour on Facebook. The x-axis shows how many users behaviors, e.g., shares, comments, etc., a given post receives. The y-axis shows how many posts receive one, two, or more of each user behavior. The histograms show that the majority of the posts elicited very a few user behaviors and only a few posts elicited a large number of user behaviors.



**Table 9: Pearson correlation matrix for variables with p-value listed in parentheses. Due to the large sample size ( $N=1,734$ ), the correlations tend to be significant.<sup>8</sup>**

	Share	Com- ment	Like	Link click	Negativ e feedbac k	Words	Ques- tions	Reading ease	Average polarity	Fans
Share	1.00	0.55 (0.00)	0.77 (0.00)	0.45 (0.00)	0.40 (0.00)	0.00 (0.97)	-0.06 (0.01)	0.01 (0.77)	0.02 (0.52)	0.05 (0.05)
Comment	0.55 (0.00)	1.00	0.48 (0.00)	0.28 (0.00)	0.43 (0.00)	-0.01 (0.64)	0.12 (0.00)	0.01 (0.56)	-0.01 (0.77)	0.09 (0.00)
Like	0.77 (0.00)	0.48 (0.00)	1.00	0.35 (0.00)	0.42 (0.00)	-0.01 (0.70)	-0.08 (0.00)	0.00 (0.99)	0.07 (0.00)	0.04 (0.09)
Link click	0.45 (0.00)	0.28 (0.00)	0.35 (0.00)	1.00	0.25 (0.00)	0.08 (0.00)	0.01 (0.64)	-0.01 (0.60)	0.01 (0.68)	0.05 (0.03)
Negative feedback	0.40 (0.00)	0.43 (0.00)	0.42 (0.00)	0.25 (0.00)	1.00	-0.02 (0.32)	-0.01 (0.72)	0.01 (0.82)	0.01 (0.77)	0.15 (0.00)
Words	0.00 (0.97)	-0.01 (0.64)	-0.01 (0.70)	0.08 (0.00)	-0.02 (0.32)	1.00	0.06 (0.01)	-0.30 (0.00)	-0.02 (0.32)	0.49 (0.00)
Questions	-0.06 (0.01)	0.12 (0.00)	-0.08 (0.00)	0.01 (0.64)	-0.01 (0.72)	0.06 (0.01)	1.00	0.19 (0.00)	0.08 (0.00)	0.09 (0.00)
Reading ease	0.01 (0.77)	0.01 (0.56)	0.00 (0.99)	-0.01 (0.60)	0.01 (0.82)	-0.30 (0.00)	0.19 (0.00)	1.00	0.11 (0.00)	-0.15 (0.00)
Average polarity	0.02 (0.52)	-0.01 (0.77)	0.07 (0.00)	0.01 (0.68)	0.01 (0.77)	-0.02 (0.32)	0.08 (0.00)	0.11 (0.00)	1.00	0.00 (0.86)
Fans	0.05 (0.05)	0.09 (0.00)	0.04 (0.09)	0.05 (0.03)	0.15 (0.00)	0.49 (0.00)	0.09 (0.00)	-0.15 (0.00)	0.00 (0.86)	1.00

## Modeling overdispersed count data

The dependent variables in this study are all counts and their distributions are far from normal. Gardner, Mulvey, and Shaw (1995) discussed various statistical attributes of count data and suggested using either Poisson or negative binomial distributions for modeling them. Further, they argued that it is inappropriate to consolidate count data to create categorical variables,

<sup>8</sup> The categorical independent variables and control variables were excluded from this correlation matrix. This matrix shows no IVs are highly correlated and collinearity is not a concern. Further, VIF (variance inflation factor) tests did not identify high collinearity among the categorical variables either. VIF calculates how much variance of an independent variable can be explained by the other independent variables. The more explained variance, the higher the collinearity. A VIF of 1 indicates a complete absence of collinearity while any value over 10 signals a problematic amount of collinearity (James, Witten, Hastie & Tibshirani, 2013, p. 101). In this study, all the VIFs were under 10 so collinearity is not a concern.

because consolidation wastes rich information and different cut-off points may lead to different results. Gardner et al. (1995) provided empirical evidence that models based on the Poisson, especially quasi-Poisson (QP) and negative binomial (NB) distributions are better choices than ordinary least square (OLS) for modeling count data, particularly because count data used in social science are often overdispersed.

Dispersion is a metric used to measure the relationship between mean and variance. When mean and variance are equal, the data strictly follow the Poisson distribution, when variance exceeds the mean, the data are overdispersed and when variance is less than the mean, underdispersed. O'Hara and Kotze (2010) recommended against using log-transforms of count data, because they perform poorly. Before reaching this conclusion, they generated a dependent variable from a negative binomial distribution and an independent variable in an association with the dependent variable and estimated the association between the DV and IV with QP regression, NB regression and OLS regression after a log-transformation. They found OLS regression generated larger error terms than the other two models except when the dispersion was small and that they also may predict negative values even though counts cannot be negative.

To handle dependent variables that do not follow normal distributions, Nelder and Wedderburn (1972) proposed the generalized linear model (GLM), a more flexible model that generalizes linear regression by allowing dependent variables to have any arbitrarily selected distribution. They achieve this goal by transforming the distributions of dependent variables to normal distributions with a formula called a “link function”, the inverse of which is used to transform a linear estimate of the dependent variable as a function of the independent variables into the GLM

estimate. A variety of methods, including maximum likelihood estimation, can be used to estimate the linear relationship between the dependent variable and the independent variables. Using this technique, GLM does not have to model a dependent variable as a linear combination of independent variables as OLS does. Mathematically, OLS models have the form  $E(Y) = \mu = X\beta$ , while GLM models have the form  $E(Y) = \mu = g^{-1}(X\beta)$ , where  $g$  is the link function. For the Poisson and negative binomial distributions,  $g$  is the logarithm function. The link function that employs the logarithmic formula is analogous to transforming count data for OLS, but GLM generates smaller errors and predicts responses within a variable's realistic boundaries. For example, a count variable will not be predicted to have a negative value (Nelder & Wedderburn, 1972). For the Poisson and negative binomial distributions, GLM is analyzed through  $E(Y) = \log \exp(X\beta)$ .

The next question is which model to choose between QP regression and NB regression. In fact, count data in social science often violate the assumption of a Poisson distribution in two ways. One is overdispersion, as discussed above, and the other is not being “memoryless”. Memorylessness means that the occurrence of an event does not change the probabilities assigned to subsequent occurrences of the event. For example, if a car engine is memoryless, the probability that it will last 100,000 miles without a breakdown as a new engine and the probability that it will last another 100,000 miles without a breakdown after the first 100,000 miles are the same. That means that with respect to breakdowns a car engine has no “memory” of what it has done before. In reality, a car engine tends not to be memoryless. Mathematically, a memoryless system is denoted as  $\Pr(X > m + n \mid X > m) = \Pr(n)$ , which means the conditional probability that at least  $n$  more occurrences will take place after the event has occurred more than

$m$  times is equal to the probability that at least  $n$  more occurrences will take place after the event has occurred for any positive  $k \neq m$  times. Being memoryless also means the occurrence of one individual case will not change the probabilities for other cases to occur (Student, 1919; Feller, 1943).

However, this assumption does not hold for some real-life cases, like the car engine example given above. In analyzing violent incidents, Gardner et al. (1995) noticed that when people behaved violently, they would be more controlled than normal later and therefore their probabilities of violence went down for a period. In the case of my study, one more “like” of a post may make Facebook’s sorting algorithms rank this post higher and show it to more people than it would have had it not received this like, which, in turn, could increase the number of subsequent likes. In their study of the accidents of machinists, Bliss and Fisher noted: “If each machinist had had the same initial probability of being involved in an accident but if this probability were increased (or decreased) by his having an accident, contagion would be present and a negative binomial distribution could result” (1953, p. 188). When a system, such as Facebook’s sorting algorithms, is not memoryless, the probabilities of serial events can be estimated with negative binomial models (Lawless, 1987; McCullagh & Nelder, 1989; Land, McCall, & Nagin, 1996).

Besides modeling non-memoryless data better than QP regression, NB regression is also better for dealing with overdispersed data than QP regression. Bliss and Fisher compared a variety of models that could be applied to overdispersed data and concluded that “the negative binomial is the most widely adaptable and generally useful of those that have been proposed so far” (1953, p.

196). Therefore, NB regression generally outperforms QP regression in terms of dealing with both overdispersion and non-memorylessness.

The differences in estimation between QP models and NB models stem from how they model the relationship between the mean and variance. While QP regression models the variance as a fixed multiple of the mean (Equation [4]), NB regression models the variance as a quadratic function of the mean (Equation [5]).

$$E(Y) = \mu \quad [4]$$

$$var(Y) = v_{QP} = \theta\mu \quad [5]$$

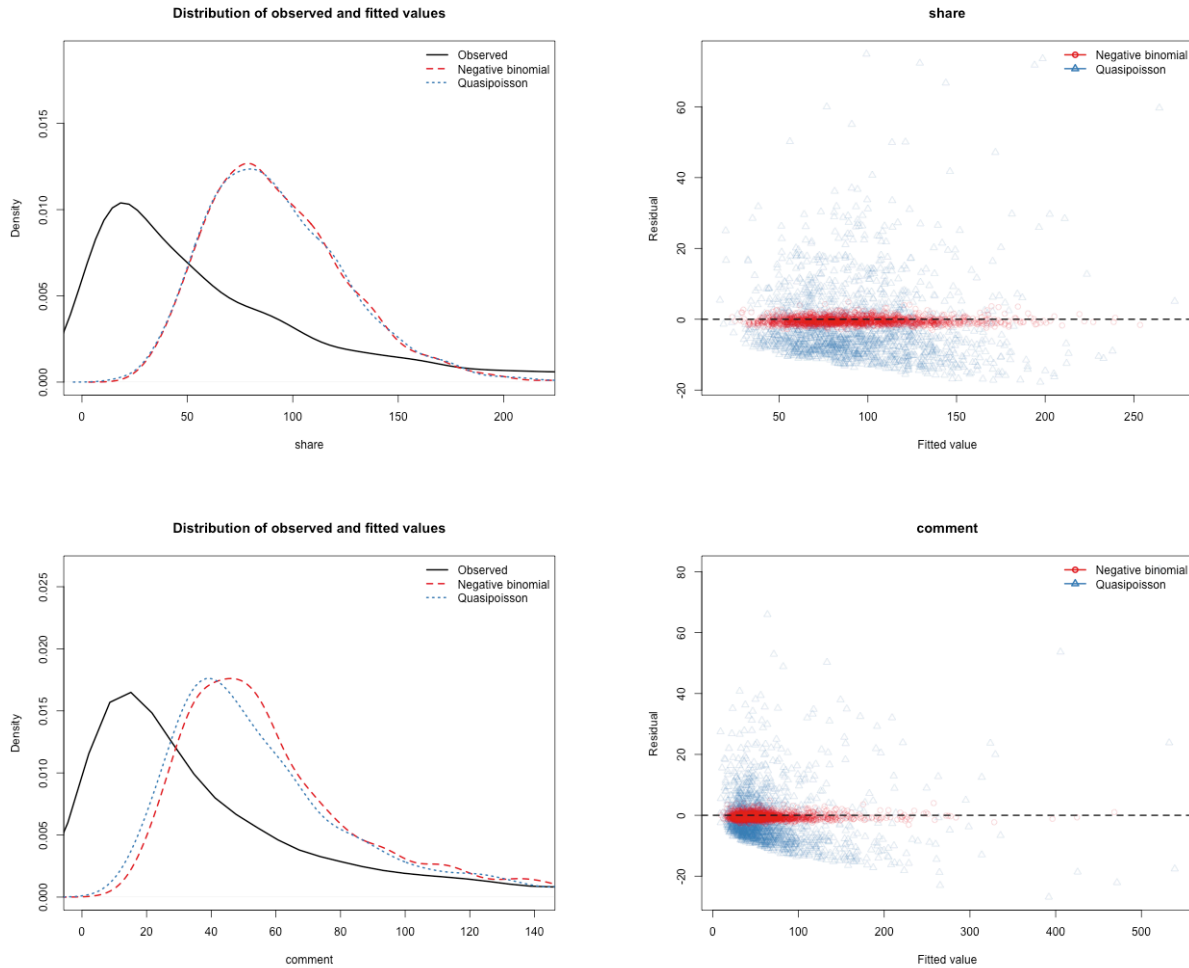
$$var(Y) = v_{NB} = \mu + \kappa\mu^2 \quad [6]$$

A trend variable was included in the model to take into account unobserved factors contributing to the overall growth that were not immediately related to the other independent and control variables. Although trend and number of fans are correlated at .78 ( $p < .001$ ), I decided to keep both because I hoped to do as much as possible to control for factors that might influence my results, acknowledging that neither the coefficient for fans nor the coefficient for trend have a simple interpretation. For this study, it was best to sacrifice clear interpretations for these two coefficients for more confidence that the controls are doing their intended job as neither of these two variables was of primary interest for this study. The trend variable,  $d$ , was calculated as the number of days from the start of the focused time period, ranging from 1 to 401.

To determine the more appropriate model for the data at hand, I employed R to analyze the count data using both QP and NB. To do this, I created two types of charts to compare their

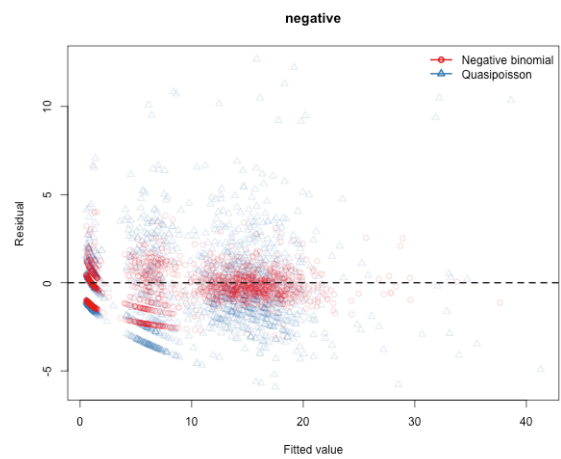
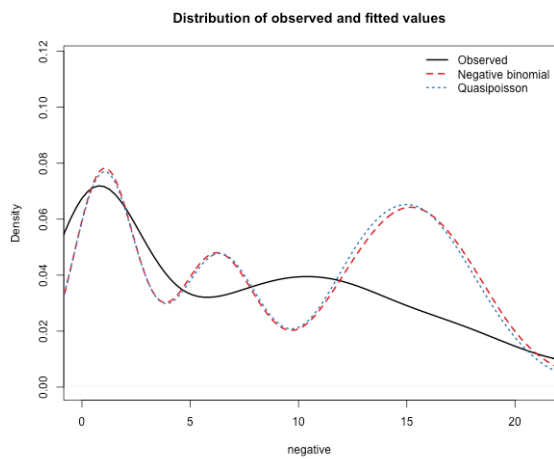
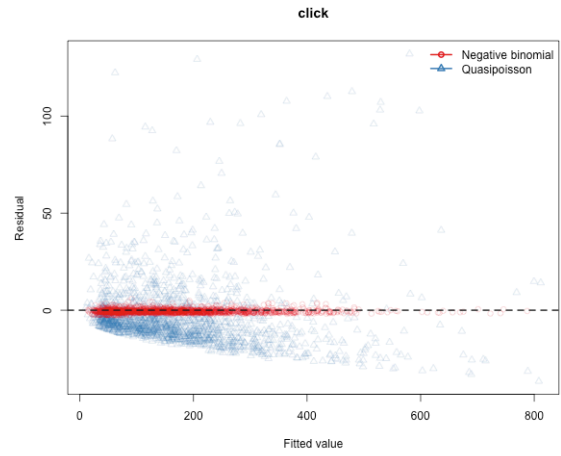
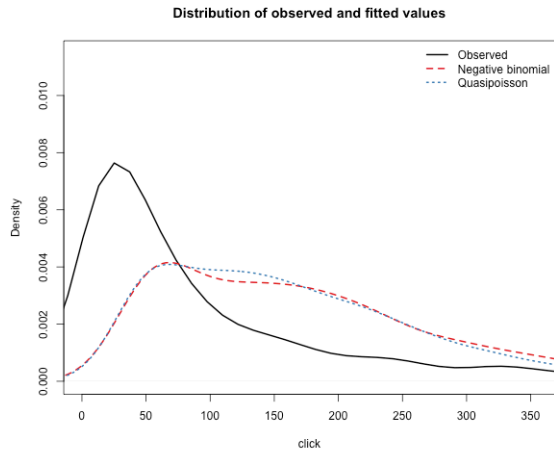
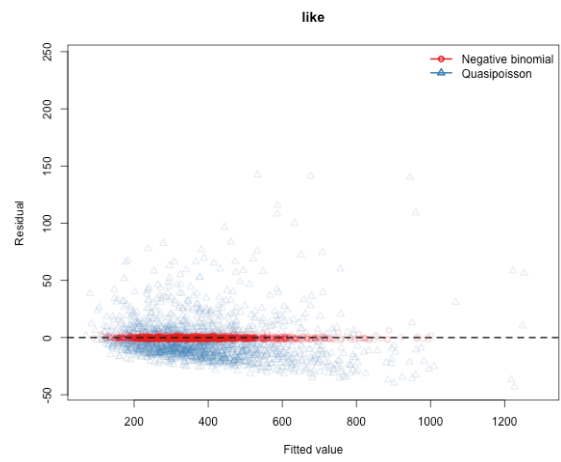
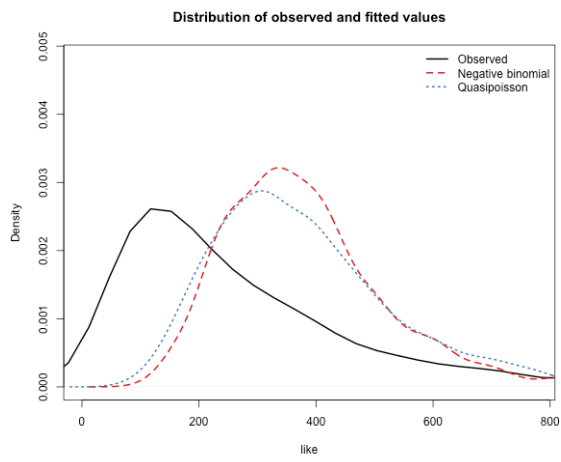
performance. One superimposed the distributions of the observed and fitted values for the two models; the other plotted the two models' residuals against their fitted values (see Figure 16). The conclusion from both comparisons was the same: the NB model works better than the QP model.

Bliss and Fisher (1953) assert that the most convincing test of the utility of the negative binomial is to compare observed values against fitted values as computed from the relevant sample for the NB and any other model under consideration as an alternative as computed from the sample. I have run both the regression with both NB and QP and illustrated the comparisons across all the five dependent variables in Figure 16. In Figure 16, the left panel shows the fitted values against observations while the right panel shows the residuals against the fitted values. In terms of fitted vs. observed values, neither NB or QP regression show consistently superior results to the other, but NB regression does have a narrower band of residuals against the fitted values. These comparisons indicate that NB regression works slightly better than QP regression for this study, which is consistent with prior empirical findings. In a variety of disciplines, researchers have concluded that NB models more closely approximate the true probability distributions for individual observations and higher estimates of uncertainties than do QP models (e.g., Anscombe, 1948; Bliss & Fisher, 1953; Gardner et al., 1995; Ver Hoef & Boveng, 2007). In this study, fitted values are consistent across the NB and QP model specifications. But because the former generates smaller residuals, it is superior to the latter in this regard. Therefore, I conclude that NB regression is a superior alternative to QP regression for this study and I report the results generated by NB regressions.



**Figure 16: Comparison between negative binomial regression and quasi-Poisson regression. The charts on the left show the distributions of the observed and fitted values; those on the right plot the residuals against the fitted values. Among the five behaviors, negative binomial regression explained variation in the current sample better than quasi-Poisson regression, as it has a narrower band of residuals against the fitted values and closely follows the distributions of the observed values except in the cases of “comment” and “like” where QP regression follows the true distribution nearly as well as NB regression.**

**Figure 16 (cont'd)**





## CHAPTER 5: RESULTS AND DISCUSSION

I organized the estimated coefficients in a matrix to compare how each DV responded to the variance of each IV and how each IV contributed to the variance of each DV (Table 10). In the matrix, each column represents a DV (i.e., behavior) as fitted by all the IVs, while each row lists an IV explaining different DVs. Background colors indicate association directions and effect sizes. Red denotes positive correlations and blue negative. Significant coefficients ( $p < .05$ ) are highlighted in yellow. For each category, the saturation indicates the effect size (but not the significance level), which means the darker the color the larger the effect size. This way the colors help illustrate the trends in effect levels and significance for the categorical IVs. For example, in the first column, “share”, the colors become less and less red from January to November, which implies that the frequency of shares was lower during those months than for December. The coefficients are exponentials of the log-odds coefficients. For example, the coefficient for "share" and "Friday" is 1.1982. As the base for this categorical variable is Sunday, Friday received 1.1982 as many shares as on Sunday, or 20 percent more shares, than Sunday.

For categorical IVs, the reference categories are as follows:

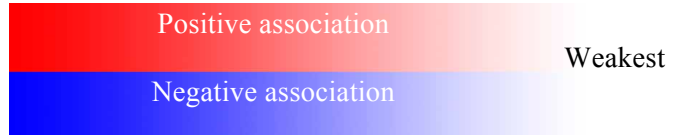
- Hour: midnight throughout 8 am combined;
- Week: Sunday;
- Month: December;
- Holiday: non-holiday;
- Post type: “text”;
- News section: uncategorized web pages.

**Table 10: Exponentiated coefficients of NB regressions with p-values in parentheses. Each column represents a user behavior (dependent variable) and each row represents an independent variable. Background colors indicate association directions and effect sizes. Red denotes positive correlations and blue negative. For each category, the saturation indicates the effect size (note not the significant level).**

\* Significant at the 0.10 level; \*\* significant at the 0.05 level; \*\*\* significant at the 0.01 level; \*\*\*\* significant at the 0.001 level; \*\*\*\*\* significant at the 0.0001 level.

**Legend:**

Strongest



Variable	Share	Comment	Like	Link click	Negative feedback
<b>(Intercept)</b>	24.1019***** (0)	41.0233***** (0)	328.0480***** (0)	27.5543***** (0)	6.7167***** (0)
<b>Number of words</b>	1.0001 (0.676)	0.9994** (0.04)	0.9999 (0.805)	0.9990*** (0.002)	0.9997 (0.13)
<b>Reading ease</b>	1.0024 (0.189)	0.9989 (0.552)	0.9998 (0.881)	1.0031 (0.14)	1.0011 (0.449)
<b>Positive polarity</b>	2.1470** (0.04)	1.2792 (0.513)	2.7206***** (0.001)	1.9529 (0.117)	1.3137 (0.36)
<b>Negative polarity</b>	1.2121 (0.109)	3.0507 (0.523)	0.6892 (0.155)	1.6468 (0.17)	1.2953 (0.567)
<b>Polarity dummy</b>	1.0459 (0.56)	0.9786 (0.781)	0.9807 (0.762)	1.0223 (0.802)	1.0330 (0.6)
<b>Slope dummy</b>	0.3843* (0.069)	0.2563*** (0.01)	0.5333 (0.154)	0.3109* (0.053)	0.5877 (0.207)
<b>Topic: arts</b>	1.1506 (0.106)	0.7671*** (0.003)	1.0976 (0.2)	1.1546 (0.148)	1.0274 (0.697)
<b>Topic: economy</b>	1.1485 (0.181)	0.7970** (0.03)	0.7546***** (0.001)	1.8687***** (0)	1.0337 (0.688)
<b>Topic: education</b>	1.3009** (0.034)	0.9350 (0.591)	1.1306 (0.237)	0.8527 (0.262)	0.9046 (0.307)
<b>Topic: health</b>	1.1874* (0.077)	1.1518 (0.15)	0.8552* (0.055)	1.4792***** (0)	1.0030 (0.969)

<b>Table 10 (cont'd)</b>					
<b>Topic: nation</b>	1.2294*** (0.002)	1.3826***** (0)	1.2150**** (0.001)	1.2699*** (0.002)	1.1018* (0.071)
<b>Topic: politics</b>	1.0285 (0.763)	2.2144***** (0)	1.0147 (0.852)	1.1208 (0.287)	1.0165 (0.823)
<b>Topic: science</b>	1.6033***** (0)	1.1271 (0.175)	1.1088 (0.158)	1.4797***** (0)	1.0556 (0.439)
<b>Topic: world</b>	1.0146 (0.847)	0.9955 (0.953)	0.9114 (0.141)	1.1367 (0.137)	0.8520*** (0.008)
<b>Post type: photo</b>	1.5217** (0.022)	1.2284 (0.264)	1.4800*** (0.01)	0.8624 (0.479)	0.9983 (0.99)
<b>Post type: link</b>	1.0695 (0.734)	1.0772 (0.709)	1.1908 (0.29)	3.9515 (0)	0.9724 (0.853)
<b>Post type: video</b>	1.1899 (0.505)	0.8827 (0.636)	1.1650 (0.484)	1.1004 (0.749)	1.5348** (0.038)
<b>Holiday</b>	0.8408 (0.366)	0.9570 (0.821)	1.3817** (0.044)	1.0624 (0.783)	1.4673** (0.014)
<b>Monday</b>	0.8289* (0.061)	0.7671*** (0.009)	0.7491**** (0.001)	0.8179* (0.08)	0.7994*** (0.005)
<b>Tuesday</b>	1.0430 (0.672)	0.9063 (0.327)	1.0336 (0.691)	1.0116 (0.919)	0.9276 (0.345)
<b>Wednesday</b>	0.9563 (0.65)	0.8307* (0.062)	0.9429 (0.475)	0.8183* (0.075)	0.8726* (0.084)
<b>Thursday</b>	1.0419 (0.678)	0.8753 (0.182)	0.9887 (0.891)	0.8522 (0.158)	0.8408 (0.028)
<b>Friday</b>	1.2200** (0.045)	1.0518 (0.614)	1.1776** (0.049)	1.1349 (0.266)	0.9016 (0.192)
<b>Saturday</b>	0.8588 (0.105)	0.8026** (0.02)	0.9493 (0.508)	0.7154*** (0.002)	0.9136 (0.233)
<b>9 -10AM</b>	1.0550 (0.903)	1.2863 (0.569)	0.5798 (0.137)	2.2564 (0.104)	1.1278 (0.72)
<b>10-11 AM</b>	0.9485 (0.89)	1.1567 (0.705)	0.6590 (0.191)	1.5217 (0.336)	1.0764 (0.797)
<b>11-12 AM</b>	1.3264 (0.442)	1.4333 (0.333)	0.8426 (0.578)	1.6496 (0.235)	1.0763 (0.79)
<b>12-1 PM</b>	1.2003 (0.623)	1.2893 (0.498)	0.8206 (0.524)	1.6756 (0.225)	1.0203 (0.942)

<b>Table 10 (cont'd)</b>					
<b>1-2 PM</b>	1.1823 (0.651)	1.4116 (0.357)	0.7565 (0.368)	1.5582 (0.296)	1.0037 (0.989)
<b>2-3 PM</b>	1.2707 (0.517)	1.3316 (0.444)	0.7847 (0.434)	1.7571 (0.184)	1.1903 (0.53)
<b>3-4 PM</b>	1.4865 (0.284)	1.2282 (0.583)	0.8877 (0.701)	1.5707 (0.287)	0.9931 (0.98)
<b>4-5 PM</b>	1.1827 (0.651)	1.2802 (0.51)	0.7690 (0.398)	1.8268 (0.157)	1.0888 (0.76)
<b>5-6 PM</b>	1.2607 (0.531)	1.4191 (0.348)	0.8626 (0.633)	2.1853* (0.065)	1.1387 (0.639)
<b>6-7 PM</b>	1.1109 (0.777)	1.2269 (0.586)	0.7671 (0.393)	2.2464* (0.057)	1.0855 (0.769)
<b>7-8 PM</b>	1.0385 (0.92)	1.0587 (0.881)	0.6957 (0.25)	1.9539 (0.121)	1.0827 (0.779)
<b>8-9 PM</b>	0.7624 (0.48)	0.6158 (0.212)	0.5730* (0.083)	1.0226 (0.959)	0.7802 (0.39)
<b>9-10 PM</b>	0.6489 (0.287)	0.9082 (0.814)	0.5923 (0.123)	0.9603 (0.931)	1.1396 (0.668)
<b>10-11 PM</b>	0.7514 (0.49)	0.8012 (0.597)	0.6516 (0.216)	0.7743 (0.59)	0.9249 (0.802)
<b>11-12 PM</b>	0.5473 (0.226)	1.2028 (0.713)	1.1003 (0.818)	1.5535 (0.439)	0.8998 (0.783)
<b>January</b>	0.8565 (0.286)	0.8838 (0.398)	1.0311 (0.801)	0.9059 (0.552)	0.1875***** (0)
<b>February</b>	0.8001 (0.144)	0.9521 (0.75)	0.9585 (0.74)	1.1066 (0.562)	0.1831***** (0)
<b>March</b>	0.7547 (0.062)	0.6450*** (0.004)	0.8339 (0.15)	1.1838 (0.328)	0.1368***** (0)
<b>April</b>	1.0543 (0.746)	0.5677*** (0.001)	0.8308 (0.174)	1.4979** (0.03)	0.1110***** (0)
<b>May</b>	0.6270*** (0.01)	0.2967***** (0)	0.4941***** (0)	0.8296 (0.368)	0.8277 (0.169)
<b>June</b>	0.6961* (0.053)	0.3077***** (0)	0.4471***** (0)	0.6945* (0.089)	1.0750 (0.607)
<b>July</b>	0.7375 (0.11)	0.3745***** (0)	0.5357***** (0)	1.4074 (0.118)	2.5239***** (0)

<b>Table 10 (cont'd)</b>					
<b>August</b>	0.6824* (0.06)	0.3253***** (0)	0.4063***** (0)	0.9329 (0.765)	2.3845***** (0)
<b>September</b>	0.8674 (0.475)	0.6187** (0.017)	0.7001** (0.033)	1.0640 (0.786)	2.4222***** (0)
<b>October</b>	0.9090 (0.547)	0.7707 (0.103)	0.8223 (0.14)	1.2336 (0.248)	2.4566***** (0)
<b>November</b>	1.0095 (0.949)	0.6947** (0.014)	0.7683** (0.032)	1.0174 (0.919)	1.8418***** (0)
<b>Number of questions</b>	0.8718***** (0)	1.2372***** (0)	0.8070***** (0)	0.9943 (0.884)	0.9741 (0.337)
<b>Fans</b>	1.0000** (0.046)	1.0000***** (0)	1.0000***** (0)	1.0000* (0.084)	1.0000 (0.726)
<b>Trend</b>	0.9935 (0.957)	0.7497** (0.018)	0.8301* (0.067)	1.2073 (0.176)	0.9880 (0.891)

The coefficients here have been exponentiated, which means they indicate the ratios of effect sizes compared to their reference categories (categorical variables) or for each unit of increase (continuous variables). Therefore, a coefficient over 1 (less than 1) indicates a positive correlation (negative correlation) with a DV, whereas a coefficient of 1 indicates no relationship between an IV and a DV. For example, the coefficient for number of words was 0.999 for link clicks, which translates to each additional word in a post contributing to a link clicks count 0.999 times what it would have been without that word, or a 0.1% reduction in the number of link clicks. On the other hand, the number of likes on holidays is 1.3817 the number of likes on other dates (or approximately 38% more), holding other things constant. It should be noted that while the coefficient for “positive polarity” is exponentiated  $\beta_{11}$ , the exponentiated sum of  $\beta_{11}$  and  $\beta_{13}$  should be thought of as the coefficient for “negative polarity”.

How do the coefficient values reported in Table 10 relate to this dissertation's hypotheses and research questions? The rest of this chapter is devoted to this discussion.

## Type 1 processing and cognitive ease

Table 11: coefficients for IVs used to test H1a and H1b.

Variable	Share	Comment	Like	Link click	Negative feedback
Number of words	1.0001 (0.676)	0.9994** (0.04)	0.9999 (0.805)	0.9990*** (0.002)	0.9997 (0.13)
Reading ease	1.0024 (0.189)	0.9989 (0.552)	0.9998 (0.881)	1.0031 (0.14)	1.0011 (0.449)

The directional predictions of H1a and H1b were supported for some behaviors and rejected for the others, while the relative magnitude predictions for coefficient size and significance were not supported. The two hypotheses were not fully supported, quite possibly due to the same complicating factors that likely affected the signs and significance of the coefficients for the variables included in this study as discussed later in this chapter.

In particular, H1a and H1b were based on the theory that Type 1 processing favors cognitive ease, focused on text length and readability and also on the assumption that Type 1 processing would elicit more user activities on Facebook as Type 2 processing slows down actions. In particular, H1a predicted that the quantities for all of the five user behaviors elicited by NewsHour posts would be negatively associated with number of words. Table 11, which replicates the relevant rows of Table 10, shows that for four of the five behaviors the direction of the correlation was as predicted and that for two, comments and link clicks ( $p=.004$  and  $p=.002$ ), the empirical correlation was statistically significant while it fell only a little short of statistical significance for negative feedback ( $p=.130$ ). However, the effect sizes were very not large because the coefficients were close to one. Only the direction for shares was not correctly predicted, but this coefficient estimate was far from significant. Based on the same theory, H1b

focused on reading ease and predicted that the quantities of the five behaviors elicited by NewsHour posts would be positively associated with their Flesch Reading Ease scores, but none of the estimated coefficients were statistically significant and for two (comments and likes) the direction of the correlation was opposite what was predicted. However, the positive correlations with more shares ( $p=.189$ ) and more link clicks ( $p=.14$ ) were statistically strong enough to merit mention, suggesting that the easier to read was a post's text, the more shares and more link clicks it would generate.

Although the H1a directional prediction regarding text length was supported to a considerable extent while the H1b regarding directional prediction regarding readability was mostly rejected, I would not reach an immediate conclusion that dual process theory does not hold on Facebook or with regard to news content. My hesitation lies in the gap between actions observable in the data and the unobservable cognitive process posited by dual process theory. I attempted to bridge the existing gap with the assumption that the easier the text, the more Type 1 processing and the more user activities. However, the results showed that harder text, as measured by readability, was correlated with more shares and link clicks. From the perspective of dual process theory, a possible explanation could be that harder text will slow down people's browsing on Facebook but slower browsing could increase or reduce user behaviors on Facebook. On the one hand, slower browsing by itself means that processing a post is more time consuming, which would leave less time and reduce the likelihood for a reader to initiate a reaction. On the other hand, if people spend more time reading a post, they may have a higher likelihood to share a link because they think they understand what the linked article is all about. Further, if harder text



communicates more about what the linked article is all about, the reduced uncertainty might encourage more sharing and clicks on links.

In addition, NewsHour readers are a special crowd who might actually prefer more challenging materials because they are better educated. 55% of them hold a college degree or above compared to 27% of the general US population.<sup>9</sup> If the degree of the challenge is positively correlated with the difficulty level of the post, then this would be a positive indicator that the linked material was of the type sought. Further, beyond Type 1 and Type 2 processing, a lot of other processes are simultaneously involved with content consumption on Facebook, and the control variables were attempts to control for these other considerations. Although my ability to control for other factors that affect content consumption was limited, the empirical limitations of those attempts and the challenge this poses to empirical investigations of this type could possibly lead to interesting and meaningful future research, such as finding ways to more clearly disentangle the effects of the various factors (things like negativity bias or desire to manage one's image) that influence responses to posts on social media from the effects of the psychological processes associated with dual processing.

The not fully supported hypotheses regarding cognitive ease might also be related to a biased sample introduced by PBS NewsHour. In fact, Facebook fans of PBS NewsHour may not be very sensitive to harder text because NewsHour is known for fairly sophisticated material and it attracts more sophisticated readers looking for challenging content. If so, this result may not indicate that the PBS audience does not prefer cognitive ease as other media consumers do, but

---

<sup>9</sup> Pew Research. (2010). The state of news media: PBS. Retrieved from <http://www.stateofthemedias.org/2010/network-tv-summary-essay/pbs/>

that this crowd, compared to the general population, may come to the challenge with larger vocabularies and more practice dealing with more complex expressions of ideas than most people in the broader population of readers covered by Flesh Reading Ease scores, so NewsHour words and expressions that may seem difficult for the average reader may not be challenging to them. There is also the possibility that if NewsHour fans are not representative of fans on Facebook in general, hypotheses based on studies of a broader population may not generalize to the NewsHour audience.

### **Dual process theory tested with Facebook data**

To test dual process theory using Facebook behavioral data, I used text length and reading ease as proxies for the amount of cognitive resources consumed in evaluating a post, because engagement of Type 1 and Type 2 processing can be distinguished by the amount of cognitive resources required to evaluate a post, particularly more cognitive resources for Type 2 processing and less for Type 1. Therefore, I hypothesized that longer and harder text would engage more Type 2 processing. In other words, text length should be negatively and reading ease should be positively correlated with user behaviors on Facebook, as posited in H1a and H2b. Negative and positive correlations are reflected in the coefficient values, and particularly a coefficient above one indicates positive correlation, below one negative and exactly one no correlation. The coefficient values show that H1a and H1b were partially supported. Regarding text length, only the shares coefficient indicated a positive relationship, as opposed to the negative relationship predicted by the hypothesis, while the coefficients for the other behaviors supported the hypothesis. That said, the coefficient for shares was insignificant and the estimated effect size was close to zero ( $\beta=1.0001$ ,  $p\text{-value}=0.676$ ), which suggests that Type 1 and 2 minds may both moderate the sharing behavior, which as I discussed earlier, would be the case if people

either share after reading through the entire post or do so after a quick skim. In the aggregate, use of the two strategies may canceled the individual effects of the two minds on the sharing behavior and rendered the text length factor nearly irrelevant because some people may rely primarily on one strategy while other people generally employ the other when responding to different posts and scenarios. Although likes was negatively correlated with text length, thereby supporting H1a, like shares, it was insignificant and its effect size was negligible ( $\beta=.9999$ ,  $p\text{-value}=0.805$ ). Its irrelevance as a whole with text length may be due to the same cancelling scenarios for sharing. As for reading ease, comments and likes did not support the hypothesis H1b as they were negatively correlated with this IV. However, they were insignificant, like text length, indicating user behaviors were processed by both Type 1 and Type 2 minds whose effects were not separated in this study.

Partial rejection of H1a and H1b does not mean that dual process theory is invalid when applied to social media. When people respond to social media content, a lot of other processes are simultaneously involved in addition to Type 1 and Type 2 processes. In fact, the control variables in this study were attempts to control for these other considerations, but the empirical limitations of those attempts and the challenge this poses to empirical investigations of this type were beyond the scope of this study. These limitations and challenges, which could not all be anticipated at the beginning of this study, could be addressed more directly in future through research designed to more clearly disentangle the effects of the various factors. Examples include studies of whether people share posts before reading them or not, studies of the effect of negativity bias on the types of cognitive processes employed in response to posts with different

types of content, and studies of the effects of self-image management on responses to social media content with different characteristics.

### **Framing, media avoidance and negativity bias**

**Table 12: coefficients for IVs used to answer RQ1.**

Variable	Share	Comment	Like	Link click	Negative feedback
Positive polarity	2.1470** (0.04)	1.2792 (0.513)	2.7206**** (0.001)	1.9529 (0.117)	1.3137 (0.36)
Negative polarity	1.2121 (0.109)	3.0507 (0.523)	0.6892 (0.155)	1.6468 (0.17)	1.2953 (0.567)

RQ1 focused on sentiment measured with average polarity scores and it investigated how the quantities of the five user behaviors elicited by NewsHour posts would vary with the polarity of a post? Based on prior studies on framing, the relationships between polarity and user behaviors should vary even though their directions cannot be predicted from the same studies because prior research found both positive and negative polarity in text could be associated with more user behaviors (e.g., Berger & Milkman, 2012). At the same time, negativity bias theory suggests negative polarity in media content should elicit more user reactions because people are more attuned to threats than potential positive developments, whereas media avoidance theory suggests people will avoid content with negative polarity to maintain their mental well-being. This dissertation study found that positive polarity was positively correlated with all the five studied user behaviors, among which it was significantly correlated with shares ( $p=.004$ ) and likes ( $p=.001$ ) and its correlation with link clicks was also worth mentioning ( $p=.117$ ). That means when the polarity was positive, the more polarized was the text the more user behaviors were elicited, which is consistent with prior findings on positive polarity (Berger & Milkman, 2012).

The row for negative polarity in Table 12 shows that a one unit increase in negative polarity was correlated with fewer likes and more of the other four user behaviors. The positive correlations between negative polarity and all the examined user behaviors except for likes is consistent with negativity bias theory if Facebook users are sensitive to bad news and also feel compelled to share it with friends, which likely would be the case if associated with threats. (Trussler & Soroka, 2014). While none of the five behaviors' coefficients were statistically significant at the standard five and 10 percent levels, the significance levels for shares ( $p=.109$ ) and link clicks ( $p=.170$ ) are high enough to merit notice. Because “like” is generally associated with something positive, it is possible that “liking” a negatively polarized Facebook post may create cognitive dissonance, which would explain why negative polarity was negatively correlated with likes at a level of significance high enough ( $p=.155$ ) to suggest that this relationship might merit further study.

### News preferences and Facebook use

**Table 13: coefficients for IVs used to answer RQ2.**

Variable	Share	Comment	Like	Link click	Negative feedback
<b>Topic: arts</b>	1.1506 (0.106)	0.7671*** (0.003)	1.0976 (0.2)	1.1546 (0.148)	1.0274 (0.697)
<b>Topic: economy</b>	1.1485 (0.181)	0.7970** (0.03)	0.7546**** (0.001)	1.8687***** (0)	1.0337 (0.688)
<b>Topic: education</b>	1.3009** (0.034)	0.9350 (0.591)	1.1306 (0.237)	0.8527 (0.262)	0.9046 (0.307)
<b>Topic: health</b>	1.1874* (0.077)	1.1518 (0.15)	0.8552* (0.055)	1.4792***** (0)	1.0030 (0.969)
<b>Topic: nation</b>	1.2294*** (0.002)	1.3826***** (0)	1.2150***** (0.001)	1.2699*** (0.002)	1.1018* (0.071)

<b>Table 13 (cont'd)</b>					
<b>Topic: politics</b>	1.0285 (0.763)	2.2144***** (0)	1.0147 (0.852)	1.1208 (0.287)	1.0165 (0.823)
<b>Topic: science</b>	1.6033***** (0)	1.1271 (0.175)	1.1088 (0.158)	1.4797***** (0)	1.0556 (0.439)
<b>Topic: world</b>	1.0146 (0.847)	0.9955 (0.953)	0.9114 (0.141)	1.1367 (0.137)	0.8520*** (0.008)

RQ2 investigated how the quantities of shares, comments, likes, link clicks and negative feedback elicited by NewsHour posts would vary with news topics. In particular, political posts were found to be associated with more comments compared to the home page and the section fronts of PBS NewsHour. Prior research found political stories were less likely to be shared on Facebook than stories on other news topics, possibly because, as argued earlier, sharing of political news is inconsistent with attainment of two gratifications from news sharing, belonging and self-presentation (Rainie & Smith, 2012; Pew Research, 2014). However, the opposite relationship was observed for this dataset although it was far from significant. Different methods of data collection may be one factor contributing to this difference in findings. Whereas my data were a record of user behaviors in a natural setting, prior research primarily relied on self-reports of users' intentions or proclivities to share political news. Referring to reliance on the Type 1 mind most of the time, Kahneman states that even scholars "observe and theorize about our own social behavior in much the same way as we attempt to perceive and understand the behavior of others" (2011, p. 228). If people are mostly guided by their Type 1 minds, Facebook users may not be aware that they do not avoid sharing political and other controversial stories even though they think they should.

## Post types and time variables

Table 14: coefficients for the post type variables.

Variable	Share	Comment	Like	Link click	Negative feedback
Post type: photo	1.5217** (0.022)	1.2284 (0.264)	1.4800*** (0.01)	0.8624 (0.479)	0.9983 (0.99)
Post type: link	1.0695 (0.734)	1.0772 (0.709)	1.1908 (0.29)	3.9515 (0)	0.9724 (0.853)
Post type: video	1.1899 (0.505)	0.8827 (0.636)	1.1650 (0.484)	1.1004 (0.749)	1.5348** (0.038)

Among the four post types (text, photo, link and video), videos were significantly and positively correlated with negative feedback, while photos were significantly and positively correlated with shares and likes. A possible reason for more user reactions for visual materials (photos and videos) than text ones is that the Type 1 mind processes visual material much faster than text and therefore visual messages on Facebook more quickly and frequently elicited reactions from users.

Various temporal control variables appeared related to user behaviors, which is interesting in its own right and may be worth future research for a better understanding. However, these variables mark PBS NewsHour's posting times rather than the times at which users respond, and users' exposures to the posts should generally somewhat lag behind their posting times. This is one challenge Shoemaker and Vos (2009) identify for studies of online media because online media are available anytime anywhere while media consumption activities are scattered across time zones.

**Table 15: coefficients for control variables regarding hours.**

Variable	Share	Comment	Like	Link click	Negative feedback
<b>9 -10AM</b>	1.0550 (0.903)	1.2863 (0.569)	0.5798 (0.137)	2.2564 (0.104)	1.1278 (0.72)
<b>10-11 AM</b>	0.9485 (0.89)	1.1567 (0.705)	0.6590 (0.191)	1.5217 (0.336)	1.0764 (0.797)
<b>11-12 AM</b>	1.3264 (0.442)	1.4333 (0.333)	0.8426 (0.578)	1.6496 (0.235)	1.0763 (0.79)
<b>12-1 PM</b>	1.2003 (0.623)	1.2893 (0.498)	0.8206 (0.524)	1.6756 (0.225)	1.0203 (0.942)
<b>1-2 PM</b>	1.1823 (0.651)	1.4116 (0.357)	0.7565 (0.368)	1.5582 (0.296)	1.0037 (0.989)
<b>2-3 PM</b>	1.2707 (0.517)	1.3316 (0.444)	0.7847 (0.434)	1.7571 (0.184)	1.1903 (0.53)
<b>3-4 PM</b>	1.4865 (0.284)	1.2282 (0.583)	0.8877 (0.701)	1.5707 (0.287)	0.9931 (0.98)
<b>4-5 PM</b>	1.1827 (0.651)	1.2802 (0.51)	0.7690 (0.398)	1.8268 (0.157)	1.0888 (0.76)
<b>5-6 PM</b>	1.2607 (0.531)	1.4191 (0.348)	0.8626 (0.633)	2.1853* (0.065)	1.1387 (0.639)
<b>6-7 PM</b>	1.1109 (0.777)	1.2269 (0.586)	0.7671 (0.393)	2.2464* (0.057)	1.0855 (0.769)
<b>7-8 PM</b>	1.0385 (0.92)	1.0587 (0.881)	0.6957 (0.25)	1.9539 (0.121)	1.0827 (0.779)
<b>8-9 PM</b>	0.7624 (0.48)	0.6158 (0.212)	0.5730* (0.083)	1.0226 (0.959)	0.7802 (0.39)
<b>9-10 PM</b>	0.6489 (0.287)	0.9082 (0.814)	0.5923 (0.123)	0.9603 (0.931)	1.1396 (0.668)
<b>10-11 PM</b>	0.7514 (0.49)	0.8012 (0.597)	0.6516 (0.216)	0.7743 (0.59)	0.9249 (0.802)
<b>11-12 PM</b>	0.5473 (0.226)	1.2028 (0.713)	1.1003 (0.818)	1.5535 (0.439)	0.8998 (0.783)

Concerning times when units of content were posted, while few individual cells approached statistical significance, the general pattern is consistent with prior findings for other media for



hour, weekday, month, and holiday. In particular, hourly patterns were observed in the interactions between PBS NewsHour and its fans on Facebook and four out of the five DVs shared similarities with the findings for competing media use described by Webster (2014), in particular computer use, which declines during prime time when television use goes up. I would expect to observe some hourly pattern of Facebook consumption that reflects the rhythm of everyday life, as other scholars have found in other media use because people tend to develop habitual and therefore predictable patterns of media consumption (Rosenstein & Grant, 1997; Webster & Phalen, 1997). However, the coefficients for hourly controls were mostly insignificant and showed no definite patterns across a day. One possibly reason for this observation is because the audience of PBS NewsHour lives across four time zones in the United States and the time differences might smear the patterns of Facebook use to the extent they exist. Because I had no data indicating the local time of each user behavior, this problem could not be addressed in this study.

**Table 16: coefficients for control variables regarding day of week.**

Variable	Share	Comment	Like	Link click	Negative feedback
<b>Monday</b>	0.8289* (0.061)	0.7671*** (0.009)	0.7491**** (0.001)	0.8179* (0.08)	0.7994*** (0.005)
<b>Tuesday</b>	1.0430 (0.672)	0.9063 (0.327)	1.0336 (0.691)	1.0116 (0.919)	0.9276 (0.345)
<b>Wednesday</b>	0.9563 (0.65)	0.8307* (0.062)	0.9429 (0.475)	0.8183* (0.075)	0.8726* (0.084)
<b>Thursday</b>	1.0419 (0.678)	0.8753 (0.182)	0.9887 (0.891)	0.8522 (0.158)	0.8408 (0.028)
<b>Friday</b>	1.2200** (0.045)	1.0518 (0.614)	1.1776** (0.049)	1.1349 (0.266)	0.9016 (0.192)
<b>Saturday</b>	0.8588 (0.105)	0.8026** (0.02)	0.9493 (0.508)	0.7154*** (0.002)	0.9136 (0.233)

Regarding the effect of the day of the week, compared to the reference category, Sunday, people seemed more active on Friday in terms of sharing and liking, possibly because Friday primes people for more social events for the coming weekend. By contrast, on Saturday, people were significantly less engaged on Facebook in terms of sharing, commenting, link clicking and leaving negative feedback. Compared to Sunday, weekdays, especially Monday, Wednesday, and Thursday, saw significantly less user activity for shares, comments, likes, link clicks, and negative feedback. The last point also means that there was more negative feedback from users on Sundays. Day of the week is clearly an important variable for studies of human behaviors, including media consumption patterns, but further investigation into lifestyle and work related factors that are likely contributing to the shapes of the patterns observed is beyond the scope of this study.

**Table 17: coefficients for control variables regarding holiday.**

Variable	Share	Comment	Like	Link click	Negative feedback
Holiday	0.8408 (0.366)	0.9570 (0.821)	1.3817** (0.044)	1.0624 (0.783)	1.4673** (0.014)

On holidays, people commented significantly more, liked more and left more negative feedback than non-holidays. People committed more likes and comments, possibly because they have more time to engage in social behaviors. People consume media differently on holidays and non-holidays, as reported by prior research. For example, people tend to go to the cinema more on holidays than working days (Webster, 2014).

**Table 18: coefficients for control variables regarding month.**

Variable	Share	Comment	Like	Link click	Negative feedback
<b>January</b>	0.8565 (0.286)	0.8838 (0.398)	1.0311 (0.801)	0.9059 (0.552)	0.1875***** (0)
<b>February</b>	0.8001 (0.144)	0.9521 (0.75)	0.9585 (0.74)	1.1066 (0.562)	0.1831***** (0)
<b>March</b>	0.7547 (0.062)	0.6450*** (0.004)	0.8339 (0.15)	1.1838 (0.328)	0.1368***** (0)
<b>April</b>	1.0543 (0.746)	0.5677*** (0.001)	0.8308 (0.174)	1.4979** (0.03)	0.1110***** (0)
<b>May</b>	0.6270*** (0.01)	0.2967***** (0)	0.4941***** (0)	0.8296 (0.368)	0.8277 (0.169)
<b>June</b>	0.6961* (0.053)	0.3077***** (0)	0.4471***** (0)	0.6945* (0.089)	1.0750 (0.607)
<b>July</b>	0.7375 (0.11)	0.3745***** (0)	0.5357***** (0)	1.4074 (0.118)	2.5239***** (0)
<b>August</b>	0.6824* (0.06)	0.3253***** (0)	0.4063***** (0)	0.9329 (0.765)	2.3845***** (0)
<b>September</b>	0.8674 (0.475)	0.6187** (0.017)	0.7001** (0.033)	1.0640 (0.786)	2.4222***** (0)
<b>October</b>	0.9090 (0.547)	0.7707 (0.103)	0.8223 (0.14)	1.2336 (0.248)	2.4566***** (0)
<b>November</b>	1.0095 (0.949)	0.6947** (0.014)	0.7683** (0.032)	1.0174 (0.919)	1.8418***** (0)

With December as the benchmark, fewer shares were observed in March and comments reached their nadir during May throughout August, likes and link clicks hit their bottoms in June, climbed slightly during the next several months, and reached their peak in December, and these four behaviors seemed to trend lower in the colder weather than in warmer weather, which is the opposite of the television viewing pattern. As discussed earlier, on a daily basis, user behaviors on Facebook also showed a contrasting trend against television viewing during prime time. As such, the opposing monthly patterns observed on Facebook against television viewing could be

seen as another piece of evidence that various media compete for people’s attention (Webster, 2014). On the other hand, negative feedback appeared in a different pattern, reaching its bottom in April and its peak in July and staying near the peak level throughout October. December throughout January was the holiday season, and people were happier and more social (Cunningham, 1979), so more shares and likes during this period may have reflected the mood of the season. In this case, bad news might have seemed less tolerable, leading people to “hide” it, because the negative feelings engendered by negative news may be inconsistent with the type of mood people are trying to cultivate during holidays.

The above findings related to temporal variables suggest that newsrooms could strategize their content publications on Facebook and other social media to cater to people’s needs and moods, according to Gidden’s (1984) the “duality of structure” hypothesis. By investigating user behavioral data, content providers can learn about users’ preferences and habits and in turn adjust their content offerings to gain more attention, as Webster (2009) proposes.

### Other control variables

**Table 19: coefficients for the remaining control variables.**

Variable	Share	Comment	Like	Link click	Negative feedback
<b>Number of questions</b>	0.8718***** (0)	1.2372***** (0)	0.8070***** (0)	0.9943 (0.884)	0.9741 (0.337)
<b>Fans</b>	1.0000** (0.046)	1.0000***** (0)	1.0000***** (0)	1.0000* (0.084)	1.0000 (0.726)
<b>Trend</b>	0.9935 (0.957)	0.7497** (0.018)	0.8301* (0.067)	1.2073 (0.176)	0.9880 (0.891)

Number of questions was positively and significantly correlated with comments ( $p < 0.0001$ ) and negatively and significantly correlated with likes and shares ( $p < 0.0001$ ). It had been expected that number of questions would be positively correlated with number of comments, because questions are invitations to respond and multiple questions in a post may should elicit more answers and thus more reply comments as some readers respond to one of the questions and other readers respond to others. However, it had not been expected that likes and shares would be negatively and significantly correlated with number of questions. That likes and shares fell as comments increased is consistent, however, with Lin et al.'s (2014) finding that during the 2012 U.S. presidential campaign, Twitter users retweeted more but replied less during eight major campaign events compared to their behaviors during the four days preceding each of the four presidential debates, which were used to establish a baseline. Lin et al. explained this finding as a consequence of social media users restricting their time and energy to a single Twitter behavior rather than engaging in two or more at the same time. When a debate was held or another major campaign event occurred, Twitter users focused on sharing information with a larger audience rather than interacting with their friends on the platform. Similarly, Facebook likes and shares may have been displaced by the increase in comments elicited by including more questions in posts.

Number of fans was correlated with the trend variable and therefore likely picked up some of the effects of external factors that I could not control for that changed over time. Although number of fans of PBS NewsHour on Facebook was statistically significantly correlated with all the user behaviors, its effect sizes were exactly 1 for each of them, meaning a change in number of fans was not correlated with a change in quantities of user behaviors. This hard-to-interpret

coefficient might be a result of including both the trend variable and number of fans, which were highly correlated, in the model, which undermined the interpretability of both coefficients. Even though including both of these variables made it difficult to say much with confidence about the meanings of their coefficients, these were not variables of primary interest for this study and more unobserved factors could be controlled for by including both of them.

### **Secondary Gatekeeping on Facebook**

This study started with an inquiry from PBS NewsHour, asking if I could help it reach a wider audience through its Facebook postings. In essence, it was about how to engage its fans, or secondary gatekeepers, to propagate more posts on Facebook, or how to go viral by using a buzzword. Although PBS had access to online readership data provided by Facebook and further information and usage statistics from media analytics services, such as Google Analytics, the editorial staff still had a difficult time understanding the raw data and drawing implications from them. For example, they were unable to compare how differently stories performed across sections, such as international news and national news, because rigorous statistical controls are required to determine whether there was a meaningful difference in the way that Facebook users interacted with international and national news stories. The statistical analysis I conducted for this dissertation showed that in general national stories generated higher counts for the five user behaviors than did international stories.

Especially, the larger amount of shares helped national news reach a wider audience through secondary gatekeeping process and this study was to explain how secondary gatekeeping unfolded on Facebook. In Chapters 1 and 2, I followed up on the Shoemaker and Vos's (2009) suggestion that content characteristics be investigated as factors influencing secondary

gatekeeping processes and I explored their suggestion in this study by adopting dual process theory and including a sizable set of control variables. This analysis helped PBS NewsHour staff see how various post characteristics they controlled affected the volumes of user responses to stories from each of the story types. For example, volumes of responses could be influenced by making language simpler or more complex. As further illustration, Table 10 shows that the coefficient for positive polarity is 2.7206 with number of likes as the DV at the 0.001 significance level. Therefore, a 10% increase in positive polarity would be correlated with a 27% ( $10\% \times 2.7206$ ) increase in the number of likes with a 99.9% confidence interval. This increase would be a considerable gain for PBS NewsHour. As discussed in Chapter 2, sharing is an important force driving information flow on Facebook. On the other hand, the non-sharing user behaviors (e.g., comment, like, link click, and negative feedback) do not directly boost secondary gatekeeping on Facebook, comments and likes do contribute directly to Facebook information flows, but all four contribute indirectly because the Facebook sorting algorithm take them into account in determining what information from other sources users see on their News Feeds, which post on the top, and which to suppress.

Beyond the suggestion from Shoemaker and Vos, some control variables' coefficients revealed interesting patterns that seemed to identify choices made by NewsHour that apparently could boost or impede secondary gatekeeping. For example, this study showed that Facebook users tended to respond with more negative feedback to the posts published by PBS NewsHour on Facebook on Sunday than on other days, especially Monday, Wednesday and Thursday. Looking at the posts receiving the highest volumes of negative feedback, I could see they were often summaries of important events around the world for the past week, such as regional conflicts,



disasters, and death tolls. As the PBS staff considers it crucial to cover these events on Facebook, no matter whether they are pleasant or not, I suggested to them that the higher volume of negative responses on Sundays may be caused by discomfort among the audience members who decided to hide these unpleasant posts to maintain their good moods on their day off. As such, PBS NewsHour would see more of Facebook's negative feedback behaviors on Sundays as they post a week's worth of unpleasant stories on this day. I also suggested that they could experiment with posting a review of last week the first thing Monday morning instead.

As the first attempt to understand user responses to message characteristics and their implications for gatekeeping, this study provided PBS NewsHour and hopefully other newsrooms a glance into the association between user behaviors and media content. Sharing behavior is the major driving force for secondary gatekeeping and my results suggest that a variety of factors contributed to this activity, not only message characteristics but also time variables. In terms of message characteristics, positive sentiment was positively correlated with shares with a larger effect size and a higher significance level than the other content related variables, followed by certain news topics (art and science), post type (photo as opposed to text), and reading ease. Although time variables were included as controls rather than focal interests for this study, Friday did stand out as highly correlated with more sharing activities. That said, a number of IVs were not significantly correlated with the sharing outcome or in the direction as predicted, which may reflect an inherent difficulty in isolating and estimating the effects of message characteristics and other factors when there are so many other plausible confounding variables that must be controlled for, including the strong likelihood that the audiences attracted

to different news sites differ considerably in what they are looking for and how they are impacted by variation in various message characteristics.

Knowing how the characteristics of posts examined above are correlated with the amounts of user behaviors elicited by their Facebook posts, the choice remained in the hands of the PBS staff to decide what stories to post on Facebook and what language to use to optimize online performance given their own specific goals. For example, they could choose to make their Facebook posts shorter and the language more polarized (emotionally charged) and expect to elicit more user behaviors from their audience, as indicated by the analysis in this dissertation. Or they could practice journalism as they believe it should be practiced without regard for how many likes their Facebook posts generate and choose to use language that according to their journalistic standards is most appropriate, but with an understanding of why their Facebook posts might underperform compared to other publishers' posts. After learning about a number of factors, such as readability and polarity in text, that were related to user behaviors elicited by their Facebook posts, the PBS staff felt they were better informed when making a decision about posting on Facebook and had more realistic expectations regarding the responses to individual Facebook posts than before when all they relied on were raw data. Like NewsHour, other online news services could benefit from an analysis of this kind.

## **CHAPTER 6: LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH**

This study attempted to answer some research questions and test some hypotheses regarding news consumption on Facebook. The nature of the challenge ahead for further exploring the implications of message characteristics for secondary gatekeeping is clearly illustrated by my findings and the study's limitations. Also, while my answers to the RQs are informative, there is much to be done to separate out the various factors and forces that may contribute to these relationships, as my review of relevant theory and empirical studies shows. The coefficients for the controls for my study also point to a number of other interesting questions that might be addressed by researchers in the future. This would include why some types of stories generated more of the activities measured than others. This should be of direct interest to news organizations.

There were technical difficulties that I was unable to overcome that future research might address. First, this research was unable to isolate important social and technical factors that may boost or impede user behaviors. As discussed earlier, there are two kinds of gatekeeping process on Facebook, sorting algorithms and user interactions, and without more detailed data, their effects cannot be separated. For instance, it is hard to determine whether a popular post on the President is due to more shares from the users or a higher ranking from Facebook. Second, selection bias cannot be ruled out for this study because some users may ignore a particular post mainly because they have seen it elsewhere but not because they are not interested in it. It boils down to a limitation due to the sampling process, which was not randomized, but based on a convenience sample. Third, the data span covers only 13 months, so some temporal variables may need further validation.

Fourth, this research is focused on only one news publisher and its practices on one social media platform, which covers a very special group of people. Because gatekeepers are socially situated, various audiences are bound and biased by their own perceptions of reality (Entman, 2007). McKenna & Martin-Smith, 2005 observed that people may engage their Type 1 and Type 2 minds differently and make different decisions when passing news messages through gates, because gatekeepers differ in gender, race, religion, education, and other demographic factors. This observation would apply to social media users as well in their roles as gatekeepers. In addition, Facebook, as well as Twitter, are especially appealing to younger, better educated, and more politically engaged people than those that might be drawn randomly from the general population (Andersen, 2003; Rainie et al., 2012). In particular, the fans of NewsHour may be a fairly sophisticated crowd compared to the U.S. population. The fact that NewsHour fans are not typical Facebook users suggests that caution should be exercised in applying its findings to a wider audience. As such, to reach more generalizable conclusions, we need to replicate this study with data from more news publishers and on more social media platforms.

Moreover, there are a number of variables that this study did not examine and future studies may have a chance to discover interesting patterns by incorporating them. First, the attributes of posts were limited to text only and image attributes were omitted in this study. In fact, variation in image attributes could have substantial effects on readers' reactions to news stories, and future researchers can include image attributes by either coding images manually or using computer vision techniques. With advanced statistical tools, future researchers may also be able to uncover to some extent how black box sorting algorithms work. By controlling for algorithm effects,

future researchers should be able to construct more refined descriptions of how users respond to different types of posts on social media.

## **REFERENCES**

## REFERENCES

- The State of the News Media 2013. (2013). Retrieved from <http://stateofthemedias.org/2013/overview-5/>
- “Facebook and Zynga to end close relationship. (November 30, 2012). Retrieved from <http://www.bbc.com/news/technology-20554441>
- Allen, C. (2005). Discovering "Joe Six Pack" Content in Television News: The Hidden History of Audience Research, News Consultants, and the Warner Class Model. *Journal of Broadcasting & Electronic Media*, 49(4), 363-382.
- Andersen, R. (2003). Do newspapers enlighten preferences? Personal ideology, party choice and the electoral cycle: The United Kingdom, 1992-1997. *Canadian Journal of Political Science/Revue canadienne de science politique*, 36(03), 601-619.
- Anderson, D. R., Lorch, E. P., Field, D. E., Collins, P. A., & Nathan, J. G. (1986). Television viewing at home: Age trends in visual attention and time with TV. *Child Development*, 1024-1033.
- Anderson, M., & Caumont, A. (September 24, 2014). How social media is reshaping news. Retrieved from <http://www.pewresearch.org/fact-tank/2014/09/24/how-social-media-is-reshaping-news/>
- Atkin, C. (1973). Instrumental utilities and information seeking. In P. Clarke (Ed.), *New models for mass communication research*, Oxford, England: Sage.
- Backstrom, L. (August 6, 2013). News Feed FYI: A Window Into News Feed. Retrieved from <https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed>
- Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). *The role of social networks in information diffusion*. Paper presented at the Proceedings of the 21st international conference on World Wide Web.
- Bales, R. F., Strodtbeck, F. L., Mills, T. M., & Roseborough, M. E. (1951). Channels of communication in small groups. *American sociological review*, 461-468.
- Barnett, G. A., Chang, H.-j., Fink, E. L., & Richards, W. D. (1991). Seasonality in television viewing: A mathematical model of cultural processes. *Communication Research*, 18(6), 755-772.
- Bass, A. Z. (1969). Refining the “gatekeeper” concept: A UN radio case study. *Journalism & Mass Communication Quarterly*, 46(1), 69-72.
- Bastos, M. T. (2015). Shares, pins, and tweets: News readership from daily papers to social media. *Journalism Studies*, 16(3), 305-325.

- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology letters*, 2(3), 412-414. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1686213/pdf/rsbl20060509.pdf>
- Beam, R. A. (1998). What it means to be a market-oriented newspaper. *Newspaper Research Journal*, 19(3), 2.
- Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological review*, 74(3), 183.
- Bennett, W. L. (1988). *News: The politics of illusion* (2nd Ed.). New York: Longman.
- Bennett, W. L. (1996). *News: The politics of illusion* (3rd Ed.). White Plains, NY: Longman.
- Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of marketing research*, 49(2), 192-205.
- Berkowitz, D., Allen, C., & Beeson, D. (1996). Exploring newsroom views about consultants in local TV: The effect of work roles and socialization. *Journal of Broadcasting & Electronic Media*, 40(4), 447-459.
- Bhat, S., Bevans, M., & Sengupta, S. (2002). Measuring users' web activity to evaluate and enhance advertising effectiveness. *Journal of Advertising*, 31(3), 97-106.
- Bourdieu, P. (1984). *Distinction: A social critique of the judgment of taste*: Harvard University Press.
- Brandtzaeg, P. B., & Haugstveit, I. M. (2014). Facebook likes: a study of liking practices for humanitarian causes. *International Journal of Web Based Communities*, 10(3), 258-279.
- Breed, W. (1955). Newspaper 'opinion leaders' and processes of standardization. *Journalism & Mass Communication Quarterly*, 32(3), 277-328.
- Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: a critical review, with emphasis on the separability of positive and negative substrates. *Psychological bulletin*, 115(3), 401.
- Carlson, M. (2014). When news sites go native: Redefining the advertising—editorial divide in response to native advertising. *Journalism*, 1464884914545441.
- Carr, D. (October 26, 2014). Facebook Offers Life Raft, but Publishers Are Wary. Retrieved from <http://www.nytimes.com/2014/10/27/business/media/facebook-offers-life-raft-but-publishers-are-wary.html>
- Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y., & Moon, S. (2007). *I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system*. Paper presented at the Proceedings of the 7th ACM SIGCOMM conference on Internet measurement.



- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752.
- Chang, B. H., & Ki, E. J. (2005). Devising a practical model for predicting theatrical movie success: Focusing on the experience good property. *Journal of Media Economics*, 18(4), 247-269.
- Chibnall, S. (1977). *Law and Order News: An analysis of crime reporting in the British Press* (London. UK: Tavistock Publications).
- Chowdhury, S. G., Routh, S., & Chakrabarti, S. (2014). News Analytics and Sentiment Analysis to Predict Stock Price Trends. *Int. J. Comput. Sci. Inform. Technol*, 5(3), 3595-3604.
- Christofides, E., Muise, A., & Desmarais, S. (2009). Information disclosure and control on Facebook: are they two sides of the same coin or two different processes? *CyberPsychology & Behavior*, 12(3), 341-345.
- Csikszentmihalyi, M., & Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience* (Vol. 41): HarperPerennial New York.
- Cunningham, M. R. (1979). Weather, mood, and helping behavior: Quasi experiments with the sunshine samaritan. *Journal of Personality and Social Psychology*, 37(11), 1947.
- Cvijikj, I. P., Spiegler, E. D., & Michahelles, F. (2011). *The effect of post type, category and posting day on user interaction level on Facebook*. Paper presented at the Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on.
- D'Alessandro, D. M., Kingsley, P., & Johnson-West, J. (2001). The readability of pediatric patient education materials on the World Wide Web. *Archives of pediatrics & adolescent medicine*, 155(7), 807-812.
- d'Haenens, L., Jankowski, N., & Heuvelman, A. (2004). News in online and print newspapers: Differences in reader consumption and recall. *New Media & Society*, 6(3), 363-382.
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, 11(4), 227-268.
- Denis, M., & Windahl, S. (1981). *Communication Models for the study of mass communication*. New York: Londres, Longman Group.
- Diddi, A., & LaRose, R. (2006). Getting hooked on news: Uses and gratifications and the formation of news habits among college students in an Internet environment. *Journal of Broadcasting & Electronic Media*, 50(2), 193-210.
- Dijksterhuis, A., & Aarts, H. (2003). On Wildebeests and Humans, The Preferential Detection of Negative Stimuli. *Psychological Science*, 14(1), 14-18.

- Donohue, G. A., Tichenor, P. J., & Olien, C. N. (1972). Gatekeeping: Mass media systems and information control. *Current perspectives in mass communication research, 1*.
- Duggan, M., & Smith, A. (2013). Social media update. Washington, DC: Pew Research Center.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science, 11*(1), 19-23.
- Entman, R. M. (2007). Framing bias: Media in the distribution of power. *Journal of Communication, 57*(1), 163-173.
- Epstein, E. J. (1973). News form Nowhere: Television and the News: New York: Random House.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American psychologist, 49*(8), 709.
- Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., . . . Sandvig, C. (2015). *"I always assumed that I wasn't really that close to [her]": Reasoning about invisible algorithms in the news feed.*
- Evans, J. B. S. T., & Over, D. E. (1996). Rationality and reasoning. Hove: Psychology Press.
- Evans, J. S. B. T. (1989). *Bias in human reasoning: Causes and consequences*: Lawrence Erlbaum Associates, Inc.
- Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in cognitive sciences, 7*(10), 454-459.
- Evans, J. S. B. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning, 13*(4), 321-339.
- Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol., 59*, 255-278.
- Evans, J. S. B. T. (2009a). How many dual-process theories do we need? One, two, or many?
- Evans, J. S. B. T. (2009b). Introspection, confabulation, and dual-process theory. *Behavioral and brain sciences, 32*(02), 142-143.
- Evans, J. S. B. T. (2010). Thinking Twice. Two Minds in One Brain. Oxford: Oxford University Press.
- Evans, J. S. B. T. (2013). *Reasoning, rationality and dual processes*. New York: Psychology Press.
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science, 8*(3), 223-241.

- Evarts, H. (2014). NY Times Taps Prof. Wiggins as Chief Data Scientist. Retrieved from <http://engineering.columbia.edu/ny-times-taps-prof-wiggins-chief-data-scientist>
- Facebook. (2013). Reach more people right from your Page. Retrieved from <https://www.facebook.com/business/boosted-posts>
- Facebook. (2013). What's the difference between impressions and reach? Retrieved from <https://www.facebook.com/help/274400362581037>
- Facebook. (2014). What is a Facebook Page? Retrieved from <https://www.facebook.com/help/174987089221178>
- Facebook. (2015). More Support For Small Businesses: Educational Events and Live Chat. Retrieved from <https://www.facebook.com/business/news/small-business-support>
- Fahr, A., & Böckling, T. (2009). Media choice as avoidance behavior: Avoidance motivations during television use *Media choice: A theoretical and empirical overview* (pp. 185-202). New York Routledge.
- Feller, W. (1943). On a general class of "contagious" distributions. *The Annals of mathematical statistics*, 14(4), 389-400.
- Fishman, M. (2014). *Manufacturing the news*: University of Texas Press.
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3), 221.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*: MIT press.
- Gabrielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social Clicks: What and Who Gets Read on Twitter? *ACM SIGMETRICS/IFIP Performance 2016*.
- Gandy, O. H. (1982). *Beyond agenda setting: Information subsidies and public policy*. Norwood, NJ: Ablex Publishing Corporation.
- Gannes, L. (July 18, 2011). Zynga and Facebook Exclusivity Goes Far Beyond Credits. Retrieved from [http://allthingsd.com/20110718/zynga-and-facebook-exclusivity-goes-far-beyond-credits/?mod=ATD\\_iphone](http://allthingsd.com/20110718/zynga-and-facebook-exclusivity-goes-far-beyond-credits/?mod=ATD_iphone)
- Gans, H. J. (1979). *Deciding what's news: A study of CBS evening news, NBC nightly news, Newsweek, and Time*: Northwestern University Press.
- Gans, H. J. (1979). The messages behind the news. *Columbia Journalism Review*, 17(1), 40-45.
- Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological bulletin*, 118(3), 392.
- Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*: Univ of California Press.

- Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.
- Golding, P. (1981). The missing dimensions: News media and the management of social change. *Mass Media and Social Change*. Beverly Hills: Sage, 1981, 63-81.
- Griffit, W., & Veitch, R. (1971). Hot and crowded: Influence of population density and temperature on interpersonal affective behavior. *Journal of Personality and Social Psychology*, 17(1), 92.
- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, 108(4), 814.
- Hamilton, J. (2004). *All the news that's fit to sell: How the market transforms information into news*. Princeton University Press.
- Hamilton, K., Karahalios, K., Sandvig, C., & Eslami, M. (2014). *A path to understanding the effects of algorithm awareness*. Paper presented at the CHI'14 Extended Abstracts on Human Factors in Computing Systems.
- Hammond, K. R. (1996). *The psychology of Egon Brunswik*. New York: Oxford University Press.
- Hancock, J. T., & Toma, C. L. (2009). Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication*, 59(2), 367-386.
- Hanoch, Y., & Vitouch, O. (2004). When less is more information, emotional arousal and the ecological reframing of the Yerkes-Dodson law. *Theory & Psychology*, 14(4), 427-452.
- Hargittai, E., & Litt, E. (2011). The tweet smell of celebrity success: Explaining variation in Twitter adoption among a diverse group of young adults. *New Media & Society*, 13(5), 824-842.
- Hartmann, T. (2009). A brief introduction to media choice. In T. Hartmann (Ed.), *Media choice: A theoretical and empirical overview*. New York: Routledge.
- Hassin, R. R., Uleman, J. S., & Bargh, J. A. (2005). *The new unconscious: Social cognition and social neuroscience*. New York: Oxford University Press.
- Hensinger, E., Flaounas, I., & Cristianini, N. (2013). Modelling and explaining online news preferences *Pattern Recognition-Applications and Methods* (pp. 65-77): Springer.
- Hu, M., & Liu, B. (2004). *Mining and summarizing customer reviews*. Paper presented at the Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining.
- Jacobs, R. N. (1996). Producing the news, producing the crisis: narrativity, television and news work. *Media, Culture & Society*, 18(3), 373-397.

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*: Springer.
- Joshi, Y. V., Ma, L., Rand, W. M., & Louiqa Raschid. (2013). *Building the B[r]and: Understanding How Social Media Drives Consumer Engagement and Sales*. Retrieved from <http://www.msi.org/reports/building-the-brand-understanding-how-social-media-drives-consumer-engagemen/>
- Kahneman, D. (2011) *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics & biases* Cambridge University Press New York.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review*, 80(4), 237.
- Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition*, 11(2), 143-157.
- Kim, J., LaRose, R., & Peng, W. (2009). Loneliness as the cause and the effect of problematic Internet use: The relationship between Internet use and psychological well-being. *CyberPsychology & Behavior*, 12(4), 451-455.
- Koolstra, C. M., Ritterfeld, U., & Vorderer, P. (2009). Media choice despite multitasking. *Media choice: A theoretical and empirical overview*, 234-246.
- Lacy, S. (1989). A model of demand for news: Impact of competition on newspaper content. *Journalism and Mass Communication Quarterly*, 66(1), 40.
- Land, K. C., McCall, P. L., & Nagin, D. S. (1996). A comparison of Poisson, negative binomial, and semiparametric mixed Poisson regression models with empirical applications to criminal careers data. *Sociological Methods & Research*, 24(4), 387-442.
- Lang, A. (2000). The information processing of mediated messages: A framework for communication research. *Journal of Communication*, 50(1), 46-70.
- LaRose, R. (2010). The Problem of Media Habits. *Communication Theory*, 20(2), 194–222.
- Lasswell, H. D. (1948). The structure and function of communication in society. *The communication of ideas*, 37, 215-228.
- Lawless, J. F. (1987). Negative binomial and mixed Poisson regression. *Canadian Journal of Statistics*, 15(3), 209-225.
- Le Bon, G. (1895). *The Crowd: A Study of the Popular Mind*. Mineola, NY: Dover Publications.
- Lee, C. S., & Ma, L. (2012). News sharing in social media: The effect of gratifications and prior experience. *Computers in Human Behavior*, 28(2), 331-339.

- Lee, F. L. (2006). Cultural discount and cross-culture predictability: Examining the box office performance of American movies in Hong Kong. *Journal of Media Economics*, 19(4), 259-278.
- Lessig, L. (1999). *Code and other laws of cyberspace*: Basic books.
- Levy, M. R., & Windahl, S. (1985). The concept of audience activity. *Media gratifications research: Current perspectives*, 109-122.
- Levy, M. R., & Windahl, S. (1985). The concept of audience activity. *Media gratifications research: Current perspectives*, 109-122.
- Lewin, K. (1943). Forces behind food habits and methods of change. *Bulletin of the national Research Council*, 108, 35-65.
- Lewin, K. (1947a). Frontiers in group dynamics II. Channels of group life; social planning and action research. *Human relations*, 1(2), 143-153.
- Lewin, K. (1947b). Frontiers in group dynamics: Concept, method and reality in science; social equilibria and social change. *Human relations*, 1(2), 5-40.
- Lewin, K. (1951). Field theory in social science.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22.
- Lieberman, M. D. (2003). Reflective and reflexive judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. R. Williams, & W. v. Hippel (Eds.), *Social judgments: Implicit and explicit processes* (pp. 44). New York: Cambridge University Press.
- Lin, Y.-R., Keegan, B., Margolin, D., & Lazer, D. (2014). Rising tides or rising stars? Dynamics of shared attention on Twitter during media events. *PloS one*, 9(5), e94093. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4031071/pdf/pone.0094093.pdf>
- Litman, B. R., & Kohl, L. S. (1989). Predicting financial success of motion pictures: The '80s experience. *Journal of Media Economics*, 2(2), 35-50.
- Lu, C.-J., & Shulman, S. W. (2008). Rigor and flexibility in computer-based qualitative research: Introducing the Coding Analysis Toolkit. *International Journal of Multiple Research Approaches*, 2(1), 105-117.
- Lu, C.-J., & Shulman, S. W. (2008). Rigor and flexibility in computer-based qualitative research: Introducing the Coding Analysis Toolkit. *International Journal of Multiple Research Approaches*, 2(1), 105-117.
- Lucassen, T., & Schraagen, J. M. (2011). Evaluating WikiTrust: A trust support tool for Wikipedia. *First Monday*, 16(5).

- Malhotra, A., Malhotra, C. K., & See, A. (2013). How to create brand engagement on Facebook? *MIT Sloan Management Review*, 54(2), 18-20.
- Malhotra, A., Malhotra, C. K., & See, A. (2013). How to create brand engagement on Facebook? *MIT Sloan Management Review*, 54(2), 18-20.
- Marewski, J. N., Galesic, M., & Gigerenzer, G. (2009). Fast and frugal media choices. *Media choice: A theoretical and empirical overview*, 107-128.
- Masicampo, E. J., & Baumeister, R. F. (2008). Toward a physiology of dual-process reasoning and judgment: Lemonade, willpower, and expensive rule-based analysis. *Psychological Science*, 19(3), 255-260.
- Maynard, D. W. (2003). Bad news, good news: Conversational order in everyday talk and clinical settings. University of Chicago Press.
- McCullough, P., & Nelder, J. A. (1989). Generalized linear models: London: Chapman & Hall.
- McKenna, K. Y. A., Green, A. S., & Gleason, M. E. J. (2002). Relationship formation on the Internet: What's the big attraction? *Journal of social issues*, 58(1), 9-31.
- McKenna, R. J., & Martin-Smith, B. (2005). Decision making as a simplification process: new conceptual perspectives. *Management Decision*, 43(6), 821-836.
- McManus, J. H. (1994). *Market-driven journalism: Let the citizen beware?* Sage Publications Thousand Oaks, CA.
- McNeil, B. J., Pauker, S. G., Sox Jr, H. C., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *The New England journal of medicine*, 306(21), 1259-1262.
- McQuail, D., & Windahl, S. (1981). Communication Models for the study of mass communication: Londres, Longman Group.
- Moriarty, S. E., & Everett, S.-L. (1994). Commercial breaks: A viewing behavior study. *Journalism & Mass Communication Quarterly*, 71(2), 346-355.
- Nadkarni, A., & Hofmann, S. G. (2012). Why do people use Facebook? *Personality and individual differences*, 52(3), 243-249.
- Nasukawa, T., & Yi, J. (2003, October). Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd international conference on Knowledge capture* (pp. 70-77). ACM.
- Nelder, J. A., & Baker, R. (1972). Generalized linear models. *Encyclopedia of Statistical Sciences*.

- Nelder, J. A., & Baker, R. J. (1972). Generalized linear models. *Encyclopedia of Statistical Sciences*.
- Neuberger, C., Tonnemacher, J., Biebl, M., & Duck, A. (1998). Online—the future of newspapers? Germany's dailies on the world wide web. *Journal of Computer- Mediated Communication*, 4(1), 0-0.
- Neuman, W. R., Just, M. R., & Crigler., A. N. (1992). *Common Knowledge: News and the Construction of Political Meaning*. Chicago, IL: University Of Chicago Press.
- Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychological review*, 108(2), 291.
- Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment.
- O'Hara, R. B., & Kotze, D. J. (2010). Do not log- transform count data. *Methods in Ecology and Evolution*, 1(2), 118-122.
- O'hara, R. B., & Kotze, D. J. (2010). Do not log- transform count data. *Methods in Ecology and Evolution*, 1(2), 118-122.
- Orwell, G. (1946). Politics and the English Language.
- Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. *Psychological bulletin*, 124(1), 54.
- Peluchette, J., & Karl, K. (2009). Examining students' intended image on Facebook: "What were they thinking?!" *Journal of Education for Business*, 85(1), 30-37.
- Peters, K., Kashima, Y., & Clark, A. (2009). Talking about others: Emotionality and the dissemination of social information. *European Journal of Social Psychology*, 39(2), 207-222.
- Pratto, F., & Bargh, J. A. (1991). Stereotyping based on apparently individuating information: Trait and global components of sex stereotypes under attention overload. *Journal of Experimental Social Psychology*, 27(1), 26-47.
- Pratto, F., & John, O. P. (2005). Automatic Vigilance: The Attention-Grabbing Power of negative Social Information. *Social cognition: key readings*, 250.
- Putnam, R. D. (1993). What makes democracy work? *National Civic Review*, 82(2), 101-107.
- Rainie, L., & Smith, A. (2012). Politics on social networking sites. *Politics*.
- Rainie, L., Smith, A., Schlozman, K. L., Brady, H., & Verba, S. (2012). Social media and political engagement. *Pew Internet & American Life Project*.



- Ratkiewicz, J., Fortunato, S., Flammini, A., Menczer, F., & Vespignani, A. (2010). Characterizing and modeling the dynamics of online popularity. *Physical review letters*, 105(15), 158701.
- Reber, A. S. (1993). *Implicit learning and knowledge: An essay on the cognitive unconscious*: Oxford Univ Press, New York.
- Rime, B., Mesquita, B., Boca, S., & Philippot, P. (1991). Beyond the emotional event: Six studies on the social sharing of emotion. *Cognition & Emotion*, 5(5-6), 435-465.
- Robinson, J. P. (1976). Interpersonal Influence in Election Campaigns Two Step-flow Hypotheses. *Public Opinion Quarterly*, 40(3), 304-319.
- Rosenstein, A. W., & Grant, A. E. (1997). Reconceptualizing the role of habit: A new model of television audience activity. *Journal of Broadcasting & Electronic Media*, 41(3), 324-344.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296-320.
- Rubin, M., & Hewstone, M. (1998). Social identity theory's self-esteem hypothesis: A review and some suggestions for clarification. *Personality and Social Psychology Review*, 2(1), 40-62. Retrieved from <http://psr.sagepub.com/content/2/1/40.full.pdf>
- Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological bulletin*, 125(1), 3.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1), 68.
- Samuels, R. (2009). The magical number two, plus or minus: Dual-process theory as a theory of cognitive kinds. In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 129-146). Oxford: Oxford University Press.
- Sasseen, J., Olmstead, K., & Mitchell, A. (2013). Digital: As Mobile Grows Rapidly, the Pressures on News Intensify. Retrieved from <http://stateofthemedias.org/2013/digital-as-mobile-grows-rapidly-the-pressures-on-news-intensify/#social-media-a-critical-tool-for-news-discovery>
- Schau, H. J., & Gilly, M. C. (2003). We are what we post? Self- presentation in personal web space. *Journal of Consumer Research*, 30(3), 385-404.
- Scheufele, D. A. (2006). Framing as a theory of media effects. *Journal of Communication*, 49(1), 103-122.
- Schlesinger, P. (1987). *Putting 'reality' together: BBC news* (Vol. 980): Taylor & Francis.
- Schmidt-Atzert, L., Hommers, W., & Heß, M. (1995). Der IST 70: Eine Analyse und Neubewertung. *DIAGNOSTICA-GOTTINGEN-*, 41, 108-130.

- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, 84(1), 1.
- Sheldon, K. M., Abad, N., & Hinsch, C. (2011). A two-process view of Facebook use and relatedness need-satisfaction: disconnection drives use, and connection rewards it.
- Shipman, A. (2004). Lauding the leisure class: Symbolic content and conspicuous consumption. *Review of Social Economy*, 62(3), 277-289.
- Shivers, J. S. (1979). The origin of man, culture, and leisure. *Leisure: Emergence and expansion*, 3-44.
- Shoemaker, P. J. (1997). A new gatekeeping model. *Social meanings of news: A text-reader*, 57-62.
- Shoemaker, P. J., & Cohen, A. A. (2006). *News around the world: Practitioners, content and the public*. New York: Routledge.
- Shoemaker, P. J., Johnson, P. R., Seo, H., & Wang, X. (2010). Readers as gatekeepers of online news: Brazil, China, and the United States.
- Shoemaker, P. J., & Reese, S. D. (1996). *Mediating the Message: Theories of influences on mass media content* (2nd ed.). White Plains, NY: Longman.
- Shoemaker, P. J., & Vos, T. P. (2009). *Gatekeeping Theory* New York: Routledge.
- Simon, H. (1971). Computers, communications and the public interest. *Computers, communications, and the public interest. Johns Hopkins Press, Baltimore*, 40-41.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological bulletin*, 119(1), 3.
- Somaiya, R. (October 26, 2014). How Facebook Is Changing the Way Its Users Consume Journalism. Retrieved from <http://www.nytimes.com/2014/10/27/business/media/how-facebook-is-changing-the-way-its-users-consume-journalism.html>
- Somaiya, R., Isaac, M., & Goel, V. (March 23, 2015). Facebook May Host News Sites' Content. Retrieved from <http://nyti.ms/1Htw6rO>
- Song, S. Y., & Wildman, S. S. Evolution of Strategy and Commercial Relationships for Social Media Platforms: The Case of YouTube. In M. Friedrichsen & W. Mühl-Benninghaus (Eds.), *Handbook of Social Media Management: Media Business and Innovation*. Heidelberg, Germany: Springer.
- Soroka, S. N. (2012). The gatekeeping function: Distributions of information in media and the real world. *The Journal of Politics*, 74(02), 514-528.

- Staats, A. W., & Eifert, G. H. (1990). The paradigmatic behaviorism theory of emotions: Basis for unification. *Clinical Psychology Review*, 10(5), 539-566.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*: Psychology Press.
- Stanovich, K. E. (2005). *The robot's rebellion: Finding meaning in the age of Darwin*: University of Chicago Press.
- Stanovich, K. E. (2009a). Is it time for a tri-process theory: Distinguishing the reflective, algorithmic, and autonomous minds. In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55-88). Oxford: Oxford University Press.
- Stanovich, K. E. (2009b). The thinking that IQ tests miss. *Scientific American*, November/December, 33-39.
- Stanovich, K. E. (2009c). *What intelligence tests miss: The psychology of rational thought*: Yale University Press.
- Stanovich, K. E. (2011). *Rationality and the reflective mind*: Oxford University Press.
- Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and brain sciences*, 23(05), 701-717.
- Stephens, M. (2014). *Beyond news: The future of journalism*: Columbia University Press.
- Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4), 217-248.
- Studenmund, A. H. (2010). *Using Econometrics: A Practical Guide* (6th ed.). Boston, MA: Addison-Wesley.
- Student. (1919). An explanation of deviations from Poisson's law in practice. *Biometrika*, 211-215.
- Sumpter, R. S. (2000). Daily newspaper editors' audience construction routines: A case study. *Critical Studies in Media Communication*, 17(3), 334-346.
- Swant, Marty (June 29, 2016). Facebook is Changing its News Feed Algorithm to Focus Less on Publishers' Content. Will Emphasize posts from family and friends.” *Adweek*. Retrieved from <http://www.adweek.com/news/technology/facebook-changing-its-news-feed-algorithm-focus-less-publishers-content-172310>
- Szell, M., Grauwin, S., & Ratti, C. (2014). Contraction of online response to major events. *PloS one*, 9(2), e89052. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3935844/pdf/pone.0089052.pdf>

- Tan, C., Lee, L., & Pang, B. (2014). The effect of wording on message propagation: Topic-and author-controlled natural experiments on Twitter. *arXiv preprint arXiv:1405.1438*.
- Tan, C., Lee, L., & Pang, B. (2014). The effect of wording on message propagation: Topic-and author-controlled natural experiments on Twitter. *arXiv preprint arXiv:1405.1438*.
- Tewksbury, D. (2003). What do Americans really want to know? Tracking the behavior of news readers on the Internet. *Journal of Communication*, 53(4), 694-710.
- Tewksbury, D. (2003). What do Americans really want to know? Tracking the behavior of news readers on the Internet. *Journal of Communication*, 53(4), 694-710.
- Thompson, C. J., & Hirschman, E. C. (1995). Understanding the socialized body: a poststructuralist analysis of consumers' self-conceptions, body images, and self-care practices. *Journal of Consumer Research*, 139-153.
- Toates, F. (2006). A model of the hierarchy of behaviour, cognition, and consciousness. *Consciousness and cognition*, 15(1), 75-118.
- Trussler, M., & Soroka, S. (2014). Consumer Demand for Cynical and Negative News Frames. *The International Journal of Press/Politics*, 19(3), 360-379.
- Tuchman, G. (1978). Making news: A study in the construction of reality.
- Updegraff, J. A., Gable, S. L., & Taylor, S. E. (2004). What makes experiences satisfying? The interaction of approach-avoidance motivations and emotions in well-being. *Journal of Personality and Social Psychology*, 86(3), 496.
- Veblen, T. (2000). *The theory of the leisure class: An economic study in the evolution of institutions*: Verlag Wirtschaft u. Finanzen.
- Vernuccio, M. (2014). Communicating Corporate Brands Through Social Media An Exploratory Study. *International Journal of Business Communication*, 51(3), 211-233.
- Vroom, V. H. (1964). Work and motivation.
- Wang, X.-T. (2006). Emotions within reason: Resolving conflicts in risk preference. *Cognition and Emotion*, 20(8), 1132-1152.
- Weber, S. (2004). *The success of open source* (Vol. 368): Cambridge Univ Press.
- Webster, J. G. (2009). The role of structure in media choice. *Media choice: A theoretical and empirical overview*, 221-233.
- Webster, J. G., & Lichty, L. W. (1991). *Ratings analysis: Theory and practice*. Hillsdale, NJ: Lawrence R=Erlbaum Associates.
- Webster, J. G., & Phalen, P. F. (1997). *The Mass Audience: Rediscovering the Dominant Model* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- Westley, B. H., & MacLean, M. S. (1957). A conceptual model for communications research. *Journalism & Mass Communication Quarterly*, 34(1), 31-38.
- Whiting, A., & Williams, D. (2013). Why people use social media: a uses and gratifications approach. *Qualitative Market Research: An International Journal*, 16(4), 362-369.
- Wihbey, J. (March 31, 2014). What's new in digital and social media research, March 2014: From gatekeeping and filter bubbles to virality and sharing. Retrieved from <http://journalistsresource.org/studies/society/news-media/digital-social-media-research-march-2014-gatekeeping-filter-bubbles-virality-sharing>
- Wildman, S. S., & Robinson, K. S. (1995). Networking Programming and Off-Network Syndication Profits: Strategic Links and Implications for Television Policy. *Journal of media Economics*, 8(2), 27-48.
- Wilson, T. (2003). Knowing when to ask: Introspection and the adaptive unconscious. *Journal of Consciousness Studies*, 10(9-10), 131-140.
- Wilson, T., Wiebe, J., & Hoffmann, P. (2005, October). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 347-354). Association for Computational Linguistics.
- Woltman Elpers, J. L. C. M., Wedel, M., & Pieters, R. G. M. (2003). Why do consumers stop viewing television commercials? Two experiments on the influence of moment-to-moment entertainment and information value. *Journal of Marketing Research*, 40(4), 437-453.
- Yurchisin, J., Watchravesringkan, K., & McCabe, D. B. (2005). An exploration of identity re-creation in the context of internet dating. *Social Behavior and Personality: an international journal*, 33(8), 735-750.
- Zhang, J., Chen, C., Härdle, W. K., & Bommers, E. (2015). *Distillation of news flow into analysis of stock reactions*. Retrieved from
- Zhao, S., Grasmuck, S., & Martin, J. (2008). Identity construction on Facebook: Digital empowerment in anchored relationships. *Computers in Human Behavior*, 24(5), 1816-1836.