Dissertations

9-21-2017

Trolling Twitter

Eric Vorst
ecvorst@gmail.com

Recommended Citation

Trolling Twitter

by

Eric Charles Vorst

M.A. Political Science, University of Missouri – St. Louis, 2015
M.B.A., Lindenwood University, 2000
B.A., English Literature and Philosophy, Central Missouri State University, 1996

A dissertation submitted to the Graduate School at the University of Missouri – St. Louis
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
Political Science

December, 2017

Advisory Committee

David Robertson, Ph.D.
Chair

David Kimball, Ph.D.

Jason Gainous
Ph.D.

Adriano Udani
Ph.D.

Acknowledgements

Any list of acknowledgements must begin with my family, as they are by far the most important thing in my life and have been the primary reason why I have been able to achieve my goals thus far. Words alone are insufficient for expressing the depth of my gratitude to:

My amazing, beautiful, intelligent, and patient wife Chrissy, for supporting and encouraging me in too many ways to count. Chrissy has been a rock-solid foundation throughout, and has kept me motivated, grounded, and optimistic as we have pursued this goal together. I love you baby.

My 6-year-old daughter Cecilia and 4-year-old son William. While they may not fully realize it yet, they have been a powerful motivation for me to become "Dr. Daddy". I pray I can be an example to them regarding the importance of hard work, determination, and a relentless stubbornness to complete every task you begin to the absolute best of your ability, no matter how large or how small.

My parents, Carl and Carole, who for my entire life have encouraged me to be creative, to work hard, and to pursue my dreams – no matter how ambitious they may have seemed. Throughout my life, they've set an example of persistence, charity, and determination to never leave a job unfinished. Dad's brilliance as an inventor and electrical engineer and Mom's excellence as an educator have provided me with the perfect examples to emulate.

My maternal grandparents, Charles and Mildred, who instilled in me from a young age the importance of family, and whose experiences growing up during the Depression instilled a deep respect and gratitude for all the amazing opportunities we have in the United States if we are willing to make sacrifices, to take responsibility for our actions, and to never give up in pursuit of your goals; and my paternal grandparents, Joseph and Lena, whose experiences fleeing 1930's Germany illustrate the lengths people will go to in order to pursue freedom and liberty.

My parents-in-law Frank and Renae, who (in addition to allowing me to marry their daughter) encouraged me to pursue the dream of earning a PhD in 2010, and provided even more support in the form of childcare and assistance with my GRE test preparation courses.

I have also been blessed with incredible guidance and examples of academic excellence from Kenny Thomas, who shepherded me into the graduate program, and David Kimball, who introduced me to the art of writing publishable papers. I owe special thanks to Dave Robertson, who was my first-ever political science professor in 2010, and who helped me shape my writing to be less overweight and more focused, focused, focused.

Finally, I thank everyone who has believed in me, as they have given me the chance to prove them right – just as I must thank everyone who hasn't believed in me, as they have

given me the chance to prove them wrong. Special thanks are due to Sibelius, Rachmaninoff, Mozart, Beethoven, Chopin, and Holst, whose works of musical genius have been and continue to be mainstays in the soundtrack that accompanies my academic pursuits. I thank God for blessing me with talents and for giving me the strength to pursue them: I am blessed to have been born in the greatest nation in the history of civilization. While the United States is not perfect, my fellow Americans and I live in a prosperous nation that offers opportunity to those willing to take risks and make sacrifices, while preserving the right to criticize those within whom we have entrusted power.

Abstract

Political polarization is a defining feature of the contemporary American political landscape. While there is little doubt that elite polarization levels have risen dramatically in recent decades, there is some debate over the existence of a corresponding rise in mass polarization. Recent scholarship on mass polarization has cited evidence related to citizens' positions on public policy issues, party sorting, and geographic polarization; however, questions remain as to the nature and extent of mass polarization in online spaces. Specifically, more needs to be known regarding how expressions of elite polarization influence the formation of polarized communities within social media.

This dissertation examines the question: Does elite polarization contribute to mass polarization in social media? This question is approached in three stages. First, this dissertation tests whether or not a causal link between elite and mass polarization strengthens with temporal proximity to highly politicized and potentially polarizing events over the span of the 2016 Republican presidential primary. Second, this dissertation examines the instant effects of elite polarization by examining a minute-by-minute live stream of reactions on Twitter during the first 2016 presidential debate. Third, this dissertation tests a contemporary theory which claims a presidential candidate's patterns of speech sows the seeds of mass polarization in the form of resentment, fear, or incivility.

This dissertation also employs the use of network analysis tools to measure the extent to which polarized communities form on social media in response to elite cues. The nature of such causal relationships provides insight into the influence polarizing messages by elites may have on mass polarization while taking into consideration the unique characteristics of the social media communications environment. In doing so, this dissertation offers a blueprint for future researchers who seek to better understand how networked technologies shape human interactions.

Table of Contents

Dissertation Chapter 1

Trolling Twitter: Introduction

Eric C. Vorst

"When you give everyone a voice and give people power, the system usually ends up in a really good place." – Mark Zuckerberg

"Social media … is a veritable battleground, where insults fly from the human quiver, damaging lives, destroying self-esteem and a person's sense of self-worth." – Anthony Carmona

"I don't know a twitter from a tweeter but I know it's important." – Hillary Clinton

"I love Twitter.... it's like owning your own newspaper--- without the losses." – Donald Trump

The 2016 presidential race was considered by many to be one of the most divisive, uncivil, and polarizing political races in recent American history. According to recent polling by Zogby Analytics, 68% of Americans viewed the contest between Hillary Clinton and Donald Trump as being "extremely or very uncivil" (PR Newswire, 2016). This represented a more than three-fold increase over Americans' views regarding the "extremely or very uncivil" nature of prior presidential contests between Barack Obama and Mitt Romney in 2012 (20%), Barack Obama and John McCain in 2008 (18%), and George W. Bush and John Kerry in 2004 (15%). An open animosity between candidates and their campaigns was evident on the campaign trail, in political advertisements, during debates, and in television reporting. In turn, this behavior produced a super-charged political environment ripe with examples of elite polarization.

Such conditions provided an excellent opportunity to study what effects, if any, elite polarization has on mass polarization.

The phenomenon of rising elite polarization has also been accompanied by a concurrent rise in the use of social media as a vehicle for political communication. One particularly popular social media platform for this type of communication has been Twitter, which allows users to instantly share thoughts, opinions, and reactions via words, images, and HTML links. One of the most valuable aspects of Twitter is that it provides an immediate snapshot of a person's state of mind; reactions to external stimuli can be measured in near real-time. Just as the conditions of the 2016 presidential race provided an excellent opportunity to study possible links between elite polarization and mass polarization, the emergence of social media as a popular form of political discussion provides an extremely valuable tool for measuring such possible links.

This dissertation examines the question: "Does elite polarization contribute to mass polarization in social media?" Such a question is not only important for better understanding the dynamics of political polarization, but it is especially timely given the convergence of increases in political incivility and the rise of social media as a political platform. The bulk of prior literature on political polarization has studied the phenomenon through the lens of the traditional media environment, while using methods appropriate for such settings. However, this dissertation argues that the social media environment is vastly different than the traditional media environment, most notably with respect to the structural dynamics of social media that redefine what it means to be a political elite. As such, this dissertation utilizes a mixed methods approach including

social network analysis and visualization in order to better understand how members of social networks react to polarizing behavior on the part of elites.

One major benefit of studying social networks is that it allows researchers to examine how interpersonal relationships and social neighborhoods form in response to "real world" events. Prior research on causal links between elite and mass polarization has primarily relied upon evidence citing individuals' positions on public policy issues and party sorting (Fiorina and Abrams 2008; Hetherington 2009; Levendusky 2009; Abramowitz 2013). While these are definitely useful measures, any ostensible effects are often separated from their purported causes by a considerable amount of time. This time lag allows for a significant muddying of the waters, as individuals have increasingly more opportunities to be influenced by multiple intervening variables as the time horizon between cause and effect increases. Perhaps even more importantly, such measures of a causal relationship largely rely upon self-reporting and proxies, which are prone to subjectivity, reporting error, and imperfect comparisons.

I approach the main research question of this dissertation in three stages. First, I use daily measurements to test whether there is a relationship between rates of affective rhetoric and the temporal proximity to major political events over a one-year time span. Second, I use minute-by-minute measurements to test for similar relationships over the 100-minute span of a presidential debate. Third, I test a contemporary theory which claims a presidential candidate's patterns of speech sows the seeds of mass polarization in the form of resentment, fear, or incivility. In all three cases, I employ cutting-edge network analysis and visualization tools to measure the extent to which any such rises in affective rhetoric are successful in gaining influence in their respective networks.

The following overview summarizes how the following chapters pursue the question regarding the effects of elite polarization on mass polarization in social media:

Chapter Two: Theoretical Model and Literature Review

Communication technology made significant advancements during 1990's, as the Internet and the World Wide Web made it possible for people across the globe to access news and information instantly, while also gaining the ability to share information with others. The rapid advancement of social media technology during the first decade of the 21st Century made it possible for people to share information in ways that were previously unimaginable. Further, these shifts allowed the common citizen to communicate in ways that could potentially rival the influence and reach of traditional mainstream media outlets.

These seismic shifts in communication technology represent a paradigmatic shift which require a rethinking of traditional political communications theory – especially with respect to the extent elite cues play in establishing possible causal links between elite polarization and mass polarization. Further, more needs to be known regarding how the social media environment modifies the impact and reach of affective rhetoric, as well as how such language contributes to the formation of polarized communities. Chapter Two presents a theoretical model for examining such causal links while, presents a review of the relevant communications literature, and outlines the chronological development of social network analysis as a methodological discipline.

Chapter Three: Data and Methods

A primary goal of social network analysis is to better understand how members in a network share information, gain influence, and create communities of discussion. As Chapter Two argues, achieving this goal requires a rethinking of existing communications theory and the use of a new set of methodological tools. However, many of these tools employ concepts and analyses that are foreign to the traditional study of political science. As such, Chapter Three provides an overview and explanation of the specific methods used in this dissertation.

Chapter Four: Social Network Patterns of Affective Rhetoric during the 2016 Republican Primaries

Chapter Four presents the first of three sequential empirical studies designed to test the main research question of the dissertation. The first empirical chapter draws upon an original data set of 366 daily observations collected over a one-year period during the 2016 Republican Presidential Primary comprised of roughly 8.4 million tweets containing approximately 160,000,000 words. This chapter uses a broad time frame for examining whether political events cause an increase in mass affective rhetoric, while also seeking evidence regarding the impact extreme affective rhetoric has on the architectural structure of networks.

Chapter Five: The Instant Effects of Confrontation, Controversy, and Contempt: A live-stream analysis of mass polarization on Twitter during the 2016 presidential debates

Whereas Chapter Four examines the dissertation's main research question from a broad temporal perspective, Chapter Five examines the same question with a far higher level of temporal detail. Drawing upon a second unique data set of approximately 1,500,000 tweets captured live and categorized into one-minute intervals during the first 2016 U.S. presidential debate, this chapter seeks to measure the immediate reactions by individuals when they are exposed to potentially polarizing elite cues. Few studies have sought to analyze causal relationships between elite cues and mass affective rhetoric with the level of detail which is pursued in Chapter Five. The findings presented in this chapter suggest compelling evidence that such reactions are often extreme and have an immediate impact on the formation of polarized communities.

Chapter Six: Mountains or Molehills? Examining the "Trump Effect" on Twitter

The third and final empirical chapter applies the theory and methods developed in this dissertation to test a prevailing argument forwarded by media and political elites during the 2016 U.S. Presidential Election: the existence of a "Trump Effect". Specifically, this argument posited that the language and actions of Donald Trump during the presidential election caused feelings of fear and anxiety among several segments of U.S. society, while simultaneously encouraging anger and aggression in others. The possible existence of such an effect raised important questions regarding how presidential candidates' patterns of speech may influence feelings of resentment, fear, or incivility. However, such a purported effect has not received a level of rigorous scientific inquiry befitting of its ostensibly serious implications.

Chapter Six seeks to help fill this void by analyzing the effects of Donald Trump's most controversial remarks during the U.S. presidential election on social networks. This chapter draws upon an original data set of approximately 4,500,000 tweets consisting of nearly 86,000,000 words collected over the span of 548 consecutive days from 9/1/2015 through 3/1/2017. Findings in this chapter suggest that while there was evidence of a causal relationship between controversial remarks by Donald Trump and a resulting rise in anxiety, fear, and aggression, similar effects were found when examining remarks made by his opponent, Hillary Clinton.

Chapter Seven: Future Research, Implications, and Conclusions

The final chapter reviews the findings presented in this dissertation while discussing how these findings better inform our understanding of the relationship between elite and mass polarization, the implications for the study of political communication in the social media era, and the ways in which future research can expand and refine the findings presented in this dissertation. Further, I discuss the wider societal implications of this dissertation with respect to prevailing utopian and dystopian perspectives regarding social media and the Internet, while placing the research into context with emerging networked technologies.

Dissertation Chapter 2

Theoretical Model and Literature Review

Eric C. Vorst

Introduction

In this dissertation, I expand upon existing theories regarding the relationship between elite polarization and mass polarization in order to answer the question: Does elite polarization contribute to mass polarization in social media? I propose that a completely different approach is required when testing for potential causal relationships in the social media environment than the approach used when testing for potential causal relationships in the traditional media environment. Such a different approach is necessary because of fundamental differences between the two communications environments. These differences are discussed in greater detail within this chapter and following chapters.

The hypotheses presented in the empirical chapters of this dissertation are based upon the following theoretical model:
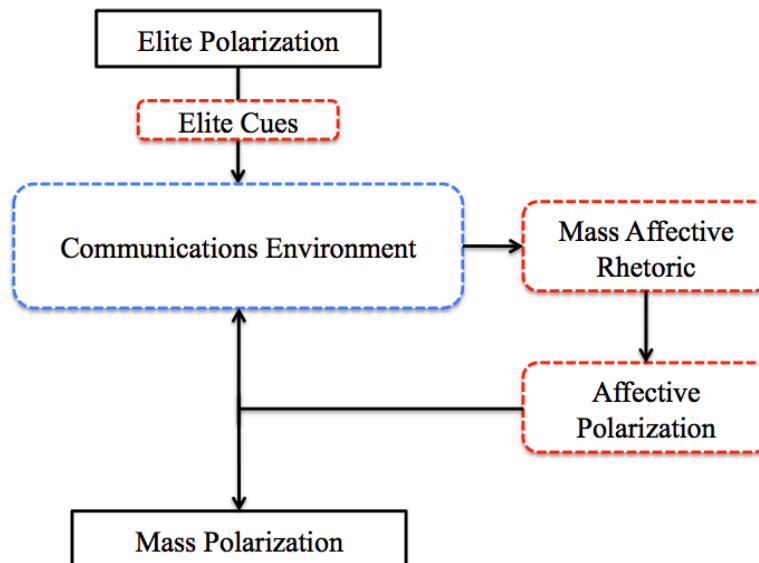
This theoretical model proposes that potentially polarizing cues originate from elites and enter the communications environment. When the mass public is exposed to these cues, there is a likelihood of increases in mass affective rhetoric which, in turn,

could contribute to increases in affective polarization. Due to the unique nature of the social media environment, the mass public is able to re-enter the communications environment to express polarized cues of their own – not unlike the cues originating from elites. In theory, this process reinforces an increasingly polarized communications environment that creates spaces where mass political polarization can develop. In other words, social media allows for affective rhetoric to not only spread efficiently among the mass public, but to be amplified by members of the mass public as well. Further, it allows for individual communities of ideological homogeneity to form with far greater ease than was previously possible in the traditional media environment.

The vast majority of prior research on elite cues, political polarization, and media effects have been conducted within the context of the traditional media environment. It is critical to acknowledge the unique nature of the social media communications environment, how it differs from the traditional communications environment, and why this matters when testing this theoretical model. Given the completely different structure of the social media communication environment, it is possible that the influence and reach of elite cues disseminated through social media sources will be different than the same elite cues would be in traditional media sources. As such, the existing literature is in need of revisiting and, in some cases, revisions.

One of the main purposes of this dissertation is to illustrate how and why traditional approaches to political communication analysis may not be the best fit for studying political communication in social media. To build this case, I first draw upon literature addressing three predominant forms of polarization: elite, mass, and affective, while also discussing the role played by incivility. Second, I discuss elite cues, self

selection, and how these variables react differently within the traditional and social media communication environments. Last, I provide support for network analysis as a robust means for testing these relationships by presenting a chronological overview of the development of social network analysis as a reliable and effective methodology within the social sciences.

Political Polarization

Elite Polarization and Mass Polarization

Political polarization is a defining feature of the contemporary American political landscape. By most measures, polarization amongst political elites has reached record levels (Heatherington 2009). A primary tool for measuring polarization among elites is DW-NOMINATE (Dynamic Weighted Nominal Three-Step Estimation), originally developed by Keith T. Poole and Howard Rosenthal in the early 1980s. This tool utilizes roll-call vote records by members of Congress as a means for estimating their position on the liberal/conservative ideological continuum. After multiple iterations over multiple congressional sessions, trends have emerged over time which demonstrate a clear ideological divergence in voting behavior among political elites. In short, Republicans are voting in a more exclusively conservative manner, Democrats are voting in a more exclusively liberal manner. More importantly, there has been progressively less overlap in the moderate areas of liberal Republicans and conservative Democrats.

Recent research suggests polarization in Congress has become so pronounced that congresspersons sharing district borders, yet representing different parties, consistently vote in opposition to each other – even when congresspersons share heavily

gerrymandered borders where one would expect some geographical common interests (Andris 2015). These phenomena are indicative of the widening levels of polarization amongst American leaders and are widely considered to influence our political system in a way that causes more harm than good. For example, an increasingly polarized U.S. Congress faces more scenarios where compromise is difficult to achieve, leading to gridlock and – in some cases – threats of a government shut down (Farina 2015).

The extent to which polarization manifests itself in the American electorate is still an open question. Fiorina has provided strong support for the argument that most voters have not been influenced by increased levels of polarization amongst elites (2011). At the same time, polarization can be observed through increased levels "sorting", wherein voters' party identification and ideological self-placement are increasingly aligned (Levendusky 2009). Polarization is also evidenced by a tendency of supporters of one party to follow to demonize supporters of the opposing party (Abramowitz 2013). Further, there is evidence to suggest mass polarization is fueled by deep-seated psychological impulses of "fear and loathing" of members in the opposing political party, especially amongst those who are in the "out party" (Kimball, et al. 2014).

Recent national polls support the conclusion that the American public is increasingly divided along party lines and, more importantly, separated by an increasing gap of partisan identification. The Pew Research Center (2014) found the percentage of Democrats who were consistently more liberal than the median Republican rose from 70% to 94% from 1994 to 2014. Similarly, the percentage of Republicans who were consistently more conservative than the median Democrat rose from 64% to 92%. During the same time span, the levels of antipathy towards members of the other political

party more than doubled, with the percentage of Democrats viewing Republicans very unfavorably rising from 16% to 38% and the percentage of Republicans viewing Democrats very unfavorably rising from 17% to 43%.

Just as levels of elite polarization can be measured by observing behavior on the part of political elites such as voting records or other elite cues, levels of mass affective polarization can be measured by observing variances in mass affective rhetoric. Questions remain as to whether or not high levels of affective polarization translate into high levels of mass political polarization. However, it is reasonable to believe that such a relationship could exist, as an atmosphere filled with strong psychological divisions could be primed for divisions along other lines, given the proper elite cues are delivered.

Such a possibility appears more likely when one considers the possibility that expressions of political polarization in the form of elite cues may have a kind of framing effect on the mass public, wherein expressions of political polarization by elites influences and shapes the mass public's understanding of political reality. Broadly defined, political framing occurs when a story or issue is portrayed using a specific perspective or through a particular lens. Despite being presented with the same set of facts, a person may reach different conclusions depending upon the way an issue is framed. Framing has the potential to be a powerful persuasive tool, as it occurs in a manner that is far less obvious than the traditional means of outlining an argument based upon clearly stated premises and conclusions.

If viewed from a framing theory perspective (Blumer 2015), the framing potential of elite cues would equate to elites affecting not only polarized behavior on the part of the mass public (or, "what to think about") but also potentially affecting polarized political

positions on the part of the mass public (or, "what to think about it"). Given the influence of political figures' ideological differences on affective mass polarization (Rogowski and Sutherland 2015), such a causal link is not out of the question. In attempting to answer questions regarding the extent and effects of mass polarization on political participation, the vast majority of research has been conducted through the lens of traditional forms of communication, such as mass media messages, candidates' campaigning tactics, or voting behavior of elected officials.

Affective Rhetoric, Incivility, and Affective Polarization

An increasing body of literature has defined mass polarization in terms of affect. While related to the concept of emotion, affect is best defined as emotion that persuades. When applied to political polarization, this school of thought argues that rather than being driven by political ideology, political divisions in the mass public are driven by hostility towards the opposing party. Instead of a person with one party identification opposing someone with a different party identification based upon ideological differences or policy disagreements, such hostility is the product of psychological mechanisms. Drawing upon a definition of affect as emotional persuasion, it can be viewed as a type of argument that is less cerebral and more base.

When such persuasion is married to party identification and infused within political debate, the results can be detrimental to reasoned discussion. Such partisan discrimination fuels levels of affective polarization that can, in some cases, be equally as strong as levels of polarization based on race (Iyengar and Westwood 2015). These

tendencies are troubling, especially given what social scientists know about the myriad divisions rooted in race related issues.

Another potentially troubling manifestation of affective rhetoric is embodied through incivility. In broad terms, incivility can be defined as "rude or impolite attitude or behavior" (Merriam-Webster 2017). Incivility is of particular interest to political scientists, given its potential for eroding democratic norms, trust in institutions, and a healthy political discourse. If the media, presidents, and congresspersons are seen as progressively mean and nasty, this could diminish citizens' perceptions of media and political elites' ability to reach reasoned, rational, and dispassionate conclusions about how government should operate. Further, citizens who are averse to (or simply tired of) constant disagreements, bickering, and nastiness may simply turn away from political discussion and choose less stressful pastimes.

It is no surprise that the political communication literature is ripe with empirical studies of incivility's potentially negative effects on the political process. For example, incivility depicted on television during political debates has been found to create a corresponding distrust of government (Mutz and Reeves 2005). Evidence exists of a negative relationship between discursive incivility and deliberative attitudes such as open-mindedness and favorable assessments of opposing viewpoints (Hwang 2008). Further evidence suggests that exposure to increased levels of incivility leads to decreased political trust and political efficacy (Borah 2013), while even more evidence points to a demobilizing effect on voters exposed to increased levels of incivility (Wolf, Strachan, and Shea 2012). Such research represents but a few examples; however, it is consistent with a wider literature detailing the potential negative effects of incivility.

At the same time, there is some evidence that the negative consequences of incivility are not as extreme as others may fear. Brooks and Geer (2007) found that despite uncivil attacks by political elites being viewed by the public as less fair or less informative, such behavior did not lead to detrimental effects among the public. Others have argued that incivility itself is too broadly defined, and that researchers should instead be measuring more pronounced forms of uncivil behavior, such as expressions of outrage (Sobieraj and Berry 2011). Regardless of the degree to which incivility has a negative impact, it must be stressed that the bulk of these studies have been conducted within the context of the traditional media environment – a context which is markedly different than the social media environment.

If it is true that affective rhetoric in the form of incivility has potentially deleterious effects within the traditional media communication environment, it is reasonable to conclude that similar effects may occur within the social media communication environment. The main goal of this dissertation is not necessarily to confirm or debunk the extent of incivility's deleterious effects on political discourse. Rather, the main goal is to find evidence of *how* forms of affective rhetoric and incivility propagate throughout social networks and, the ways in which such behavior contributes to political polarization, and – most importantly – how the nature of these relationships differs when viewed within the context of social media.

Regardless of whether a causal linkage exists that flows from elite polarization, through elite cues, affective rhetoric, and affective polarization, and results in mass polarization, the political communications literature can be strengthened by better understanding how different types of elite cues influence affective polarization in

different types of interpersonal environments. This understanding is especially important

with respect to how elite cues delivered in a live, confrontational, and politically charged

atmosphere contribute to affective polarization which, in turn, may be creating conditions

that may foster mass polarization in online spaces.

Traditional Media as One-Way Streets

Elite Cues

It is widely accepted that traditional media has the power to shape political

discussion. Much of this power is derived from its role as an agenda setter. Early

research in the study of media effects found the media not only influences what people

think about an issue, but it influences how important people think that issue is. It was

further argued that the power to influence issue salience was in fact the primary effect of

mass media, to such an extent that the power to influence opinion formation was

negligible (McCombs and Shaw 1972). However, such perspectives were challenged by

the demonstration of a causal link between opinion formation and issue salience, built

upon the premise that "the distinction between 'what to think' and 'what to think about'

is misleading" (Entman 1989). Further, opinion formation likely involves a more

complicated process than simply being told what is most important. Rather, opinions are

formed through a process that combines an individual's prior beliefs, an issue's

alignment with these beliefs, and perceived issue salience (Zaller 1992).

Given that it is nearly impossible to force someone to adopt a specific position

that they would have otherwise not taken, the more likely route for engaging in

successful agenda setting is through selective dissemination of information. Despite the

fact that mass media cannot necessarily make people think a certain way, it can influence what people are thinking about. In turn, opinion formation can be influenced by promoting the importance of one issue while suppressing or omitting other potentially important issues.

More recent research indicates that some areas of mass media play a role in opinion formation that is not necessarily dependent upon issue salience. Specifically, media bias in editorials has been shown to influence voters' evaluations of candidates as well as the choices voters' make at the ballot box (Druckman and Parkin 2005). While it is true that media outlets engage in selective dissemination of information (and indicating what people should think about) by choosing whether or not to publish an editorial on a particular issue, it is also true that the direct impact of an editorial's persuasive power on the merits of an issue can not be ignored. Rather put, while the power to influence "what to think" flows directly from the power to influence "what to think about", one cannot assume that salience alone is sufficient. Once the importance of an issue is communicated, the way in which the issue is presented or argued will have a direct impact on the types of opinions that are formed.

Overt bias in the form of editorials is but one way information consumers can be influenced "what to think". There is also significant evidence that the ways in which news is framed plays a central role in how people form their opinions. Such evidence was presented by Shanto Iyengar (1992), who classified framing as being either thematic or episodic. According to Iyengar, episodic framing presents and discusses issues within a narrow or "zoomed in" perspective. Episodic framing often focuses on the immediate impact on individuals in society, and can often incorporate a dramatic element. A

popular example of such a news story is of a struggling single mother who must survive on welfare. Because of the episodic frame, viewers or readers of the story are more likely to focus on how to help the single mother or other single mothers directly and personally. Conversely, thematic framing uses a "zoomed out" perspective, typically focusing on an issue's broader impact, wider themes, or general implications for society. An example of thematic framing in a news story might discuss the issue of welfare from an institutional standpoint, the challenges of funding, or the negative stigma society sometimes applies to those on welfare. In contrast to episodic framing, which tends to be more easily accessible, thematic framing is often more abstract and, as such, is often more difficult for the general population to digest intellectually.

Ultimately, the types of actions people feel must be taken in response to these news stories depends heavily upon whether they are presented in an episodic or thematic frame. For example, an episodic frame may lead a viewer to favor actions that occur at the individual level, such as finding ways to directly assist the single mother on welfare. Likewise, a thematic frame may lead a viewer to favor actions that occur at the institutional level, such as welfare reform or more effective application of need-based social resources. The distinction between episodic and thematic framing matters because the majority of televised news is presented in episodic frames. As a result, information consumers are more likely to form opinions that favor an individual-based need agenda versus an agenda favoring institutional or systemic modifications. Such an agenda-setting power by the media is more subtle than selective information dissemination and editorial bias, but it is no less as effective in its end results.

Self-selection

It is widely accepted that traditional media engages in agenda setting and that this behavior is a means by which bias can be expressed. However, the extent to which traditional media serves a function as gatekeeper of political information must also be acknowledged. The traditional media's power as gatekeepers provides it with the potential to exert a powerful influence on citizens' political discussions, including content, intensity, and tone. Traditional media sources decide what information to provide, how to frame it, and how frequently to provide it. Within this structure, information consumers have little direct power to influence the content or delivery of the information provided by information sources.

Despite information consumers having little direct power to influence the content or delivery of the information provided by information sources, consumers have the power to choose from a range of available traditional media information sources. While it is true that information providers may adjust content to meet the perceived expectations of their audiences, it is also true that an information provider's slant will influence the composition of its audience (Levendusky 2013). Such a tendency allows for information consumers to "self-select" the information they receive, thus allowing them a modicum of autonomy when choosing the manner in which they are exposed to political information (Mutz 2006). Opportunities for self-selection increased as the number of television news providers expanded during the shift from network broadcasts to 24/7 cable news providers in the 1980's. Still, the range of information consumers' choices within the traditional media marketplace has been and remains somewhat constrained,

while the providers of such information retain a significant amount of discretion

regarding the type of information provided.

The Changing Landscape of Political Communication: Social Media as a Platform

      Since its widespread adoption in 2009, Twitter has emerged as a major tool for

sharing political information and for engaging in political debate, all within a 140

character-per-message limit. By 2015, 23% of adult Internet users utilized Twitter, up

from 16% in 2012 (Pew 2015). The demographic characteristics of Twitter users are also

important to note. When measured as a percentage of all Internet users, Twitter users are

disproportionately young (32% between 18-29), urban (30%), and non-white (28% black,

non-Hispanic; 28% Hispanic) when compared to the overall demographics of the United

States (Pew 2015).

      The increased use of social media by politicians is a reflection of the increased

use of social media by the American public. In ten short years, the frequency with which

Americans use social media platforms such as Facebook and Twitter has increased

dramatically, rising from 7% in 2005 to 65% in 2015 (Pew 2015). As recently as 2012,

34% of Americans used social media to share their thoughts or comments on political and

social issues. Social media users on opposite ends of the political spectrum did so even

more frequently, with 42% of liberal Democrat social media users and 41% of

conservative Republican social media users utilizing such platforms to engage in political

discussion (Pew 2012). The level of engagement on social media did not end with

merely discussing politics; by the end of 2014, 18% of Republican social media users and

15% of Democratic social media users were actively following candidates for office, political parties, or elected officials (Pew 2015).

One particularly noteworthy aspect of Twitter that sets it apart from social networking sites like Facebook is that Twitter allows for a much higher level of anonymity in its users. While Twitter does prohibit users from impersonating someone else, it does not prohibit the use of pseudonyms, or "fake names". This policy is in stark contrast to Facebook's "real name" policy, which requires people to "provide the name they use in real life" so that other users "always know who [they're] connecting with" (Facebook 2016). These differences between anonymity policies matter, as recent research indicates there is a positive correlation between online anonymity and incivility when discussing "politically sensitive and potentially divisive issues" (Rowe 2015).

The type of social media platform also matters, whether it be Facebook, YouTube, Twitter, or any one of the myriad social network platforms, because different rules govern the types of interactions users can engage in. As such, these rules have a direct effect on the types of networked publics that can be shaped within these environments (Boyd 2010). This is important to acknowledge because it means one social media platform may foster incivility more readily than another social media platform, depending upon the extent to which architectural allowances are made for anonymity. Given the ability for Twitter users to remain anonymous if they choose, it is possible that the Twitter environment will be more uncivil than online environments that take extensive steps to verify users' identities. In turn, one would expect levels of observed mass polarization to be more pronounced on Twitter than on Facebook. If true,

this makes Twitter a prime testing ground for the effects of polarizing language on large scale social networks.

Twitter is also useful for studying mass polarization due to its user networks being far more flexible and open than user networks on Facebook, due largely to the fact that Twitter is a predominantly public network. With the rare exception of users with private accounts, every tweet is publicly visible and available to be viewed by anyone who searches for the correct hashtag, user, or key words. Conversely, Facebook networks are largely restricted to those who make an active choice to request – and to be accepted into – a user's circle of friends or into a discussion group. While it is true that content and network analyses can be conducted on Facebook networks, such analyses provide a perspective of networks that form *after* an individual has made the conscious decision to enter into a particular discussion community.

While these differences may seem minor, they have significant implications when attempting to observe the link between elite polarization, affective rhetoric, and mass polarization. Part of finding evidence of such a link requires the ability to link elite cues to the decision-making process in the mass public, and to do so within a measurable temporal perspective. For example, if a candidate were to make an incredibly offensive statement on a Monday, it would be difficult (if not impossible) to measure the effect this had on the decision of individuals to join specific Facebook groups on the following Tuesday. It would also be difficult (if not impossible) to analyze the extent to which individuals with differing political ideologies on Facebook either intermingled or became more polarized. Granted, one could perform focused content analysis within each Facebook group; however, the power of comparison across multiple groups would be

nullified. On the other hand, the fluid and open nature of Twitter networks allows

researchers to observe causal effects of elite polarization and affective rhetoric as they

unfold. More importantly, it allows for researchers to observe how communities form in

reaction to such elite behavior *as they form*, rather than studying these communities *after*

they have formed. Again, the differences in Twitter and Facebook networks seem minor

at first, but they have significant meaning when viewed within the context of this

dissertation's main research question.

The Changing Landscape of Political Communication: Social Media and Politics

Social Media and Networked Publics

A brief review of fundamental structural changes in the political communication

environment over the last few decades is helpful in understanding social media's special

place in this evolution. Newspapers and television news programs are primarily one-way

streets, where journalists have an agenda setting influence through the process of

gathering and disseminating information (McCombs & Shaw 1972). While it is true that

consumers can self-select their sources of information, their power to influence

information content is limited.

Traditional Media: one-way streets

Information Provider	Self-selection
Information Provider	Information Consumer
Information Provider	
Information Provider	Agenda Setting
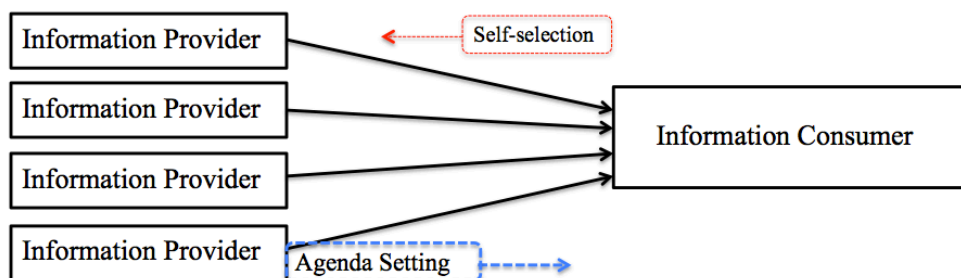
The Internet can be both a one-way or two-way street, as it not only allows

consumers to self-select online sources of information, but it also allows consumers to

share their opinions in the comment sections typically appearing at the end of an online

article (Marchionni 2013). After reading the original author's article, information

consumers are then able to read the opinions of other readers. While these opinions may

be reviewed by a forum moderator to censor obscene language or hate speech, the content

often lacks the same level of journalistic integrity that is typically expected of most

online news sources. Readers of these comments are thus exposed to sources of

information that are neither vetted, verified, or anticipated by the online publisher or

original author. More importantly, readers – not the author – have power to shape the

story in ways that were previously impossible prior to the introduction of the Internet as a

source of news and opinion.

Traditional Internet: two-way streets

| Information Provider |
| Information Provider |
| Information Provider |
| Information Provider |

Self-selection

Information Consumer

Comments Sections

Agenda Setting
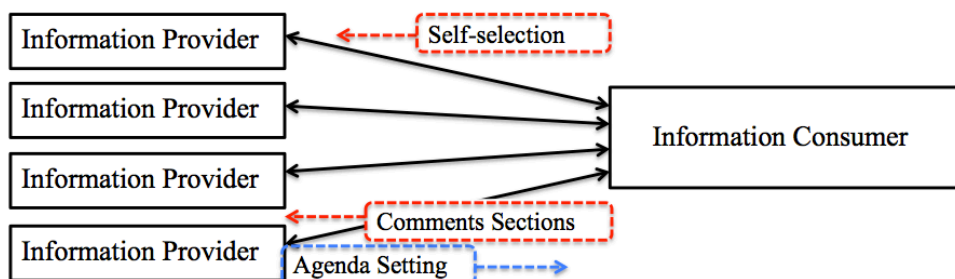
Social media changes the information sharing dynamic completely, as it fosters a

flexible media environment that is shaped almost entirely by each individual's choice of

connections. Not only can users shape the social media environment they experience, but

they can also shape the social media environment experienced by others. This

phenomenon has given rise to the notion of *networked publics*, which are defined as

spaces created by networked technologies where an "imagined collective … emerges as a

result of the intersection of people, technology, and practice" (Boyd 2010). In this

respect, information consumers have the power to become influential information

providers. This new paradigm shatters the traditional hierarchy of news dissemination

and commentary, allowing active social media participants to become "leaders of opinion

and creators of noise and buzz" (Sebastião 2014).

Networked
Publics
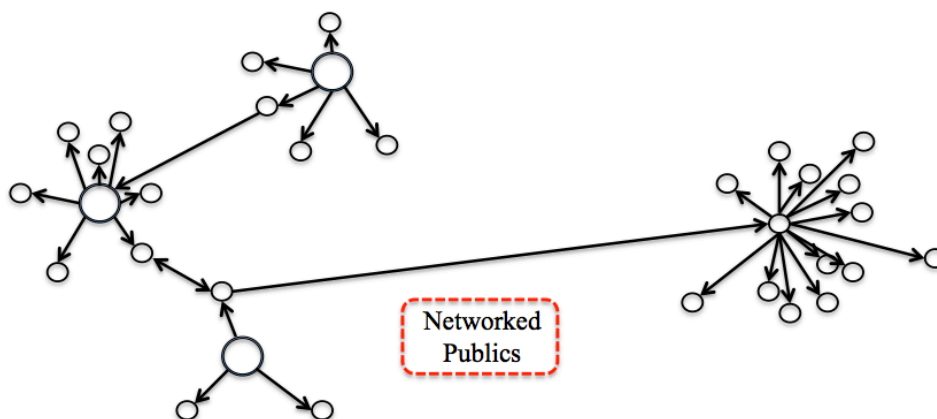
Social media's unique architecture requires analysts to rethink conventional

wisdom regarding information gatekeepers and agenda setters. In contrast to the top-

down hierarchical information dissemination process of traditional media sources and the

capability for bottom-up information dissemination in the form of user comments on

websites, members of social networks are parts of user generated neighborhoods of

discussion that are largely the result of an individual centric information dissemination

network. Social network users create and expand upon information sharing and

discussion opportunities through a constantly shifting lattice of the user's network of

connected "friends" and, more importantly, the secondary network of those "friends'

friends".

Considering the significant structural differences between the traditional and

social media information environments, an examination of mass polarization in social

media cannot begin and end with only a consideration of elite cues. Rather, equal consideration must be given to how members of a social network react when exposed to polarizing messages, with whom those messages are shared, and the extent to which those messages reach influential members in the network. Doing so provides insight into how deeply polarizing language is able to penetrate into discussion within a social network at a given point in time. As a result, it allows for a better glimpse into how social media networks facilitate mass polarization differently than traditional sources of information sharing.

Self-selection

The opportunities for members of online social networks to self-select the information they receive – as well as the people with whom they share information – are exponentially greater than they are in the traditional media environment. Such control over self-selection is especially salient when measuring for a relationship between elite polarization and mass polarization in social media environments. There is little doubt that social media has fast become a virtual town square for citizens to discuss politics (Kavanaugh, Perez-Quinones, Tedesco, & Sanders, 2010). As this town square has become more populated, it has also created increasing opportunities for citizens to be exposed to elite cues – especially from political candidates and elected officials. This is due in large part to the fact that the structure of the social media environment and the methods individuals use to share information are fundamentally different than other forms of media.

The ability for information consumers to actively choose the type of information

they receive and the providers who supply it, coupled with the ability to actively avoid types of information and providers they do not wish to consume, can lead to Internet users creating a virtual environment consisting mostly of information they agree with (Sunstein 2001). Also referred to as "filter bubbles" (Pariser 2011), such environments have potentially negative consequences, as information consumers are able to craft their own information realities which could favor a disproportionately high percentage of information with which they agree versus information with which they disagree. One potential downside of such an outcome is that it can lead to ideological homogeneity in networks. People in "filter bubbles" ostensibly experience more comfort in being surrounded by those with whom they agree, yet could be argued to be at a disadvantage due to being sheltered from attitudes and beliefs that are different than theirs. Clearly, there are potentially negative implications of such outcomes if one considers the benefits of ideological diversity.

Elite Cues: Social Media as a Strategic Resource for Political Entrepreneurs

Just as social media has evolved into an inexpensive and powerful tool for the mass public to share political information and engage in political discussion, the same is true for political elites. Politicians have modified their communications strategies to take advantage of the medium as a low cost means to reach potential voters and supporters. One recent example of how politicians have begun to utilize social media as a low-cost and high-impact medium occurred during the 2008 Democratic election. As Gainous and Wagner (2011) observed, Barack Obama's "Yes We Can" YouTube video was able to reach a far larger audience at a far lower cost than a competing televised town hall by

Hillary Clinton. The Obama Campaign seized upon other forms of social media, including direct text messaging, in order to reach potential voters with very little investment in campaign finances. Other politicians were quick to note the potential benefits and incorporated their own social media strategies. By 2012, the use of YouTube videos was a standard element of most campaign strategies (Gainous and Wagner 2014), as was the incorporation of other forms of social media like Facebook and Twitter. This relatively recent and pronounced influx of political elites into the social media environment cannot be underestimated, as it has most assuredly resulted in an atmosphere that not only contains more elite cues, but also likely contains increased expressions of elite polarization.

Political elites' use of social media as a strategic tool for gaining advantages over their opponents often extends beyond the act of communicating policy objectives, attempting to sway potential voters, or striving to turn out the vote. It can also be used to employ more negative strategies such as those employing fear appeal (Borah 2014) or personal attacks. In this respect, Hillary Clinton and Donald Trump have been especially active on Twitter during the 2016 presidential election. Using the Twitter API to collect the 3,200 most recent tweets sent by Clinton and Trump as of 10/21/2016, it was found that Trump's official Twitter account averaged 12.1 tweets per day since 2/1/2016 and Clinton's official Twitter account averaged 24.8 tweets per day since 6/14/2016. The accusations leveled within these official tweets ran the gamut, including – but certainly not limited to – verbatim candidate depictions of their opponents as being dangerous, risky, terrifying, clueless, crooked, corrupt, hateful, shameful, bullies, bigots, whiners,

and, of course, liars (Twitter 2016). Such an environment provides rich testing grounds for examining the effect elite polarization may have on mass polarization in social media.

Conclusions

There is a great deal of existing scholarship examining the extent to which elite and mass polarization is linked; however, the vast majority of this literature examines such a relationship within the scope of traditional communications environments. Within this chapter, I have presented an argument for why researchers must take into account the paradigmatic differences between traditional communication environments and the social media environment. Not only does the study of social networks require a different contextual perspective on the part of the researcher, but it also requires a unique set of tools. In Chapter Three, I outline the reasons why these differences matter when questioning whether or not elite polarization contributes to mass polarization in social media and why social network analysis is an effective method for answering this question. Chapter Three also presents a brief history of the methodological development of social network analysis, while describing the processes used to gather data and the specific social network analysis tests employed when testing various aspects of the main research question in Chapters Four, Five, and Six.

Dissertation Chapter 3

Data and Methods

Eric C. Vorst

Introduction

The main research question of this dissertation asks whether or not elite polarization contributes to mass polarization in social media. As noted in Chapter 1, this question is pursued through the investigation of three related and more specific questions, using three individual mixed-methods empirical studies and three unique data sets to investigate the main research question from three different perspectives.

It is important to note that most large *n* Twitter studies have focused upon single conglomerate networks comprised of an aggregation of messages collected over the span of multiple days, weeks, or months of observations. To be sure, such studies are breaking ground with respect to our understanding of how mass polarization manifests itself in social media. While these sorts of data projects have allowed for unique research questions to be tested rigorously, little has been done in the area of performing extended time-series sentiment analysis. Specifically, no known works have attempted to use large scale network analysis methods over an extended period of time by using individual daily observations as a means for examining how changes in the real-time political climate influence levels of political polarization in social media networks.

Measuring the frequency of affective rhetoric in social networks is an important first step in understanding the role such language may play in fostering political polarization. However, measurements of frequency alone do not tell the full story of how negative messages impact participants of social networks, where connections between individuals vary widely depending upon their immediate personal networks as well as the personal networks of those with whom they are connected. Answering questions such as these will shed more light on how social networks react and respond to affective rhetoric,

as well as how this informs our understanding of the relationship between elite and mass polarization.

With social media, it is less a question of how many times a message is sent; it is more a question of how deeply that message spreads throughout the network and how long that message is able to sustain itself. These are defining features of the social media environment which represent a paradigmatic shift compared to the traditional media environment. Understanding these differences is a critical part of answering this dissertation's question regarding the extent to which elite polarization influences mass polarization in social media. Social network analysis allows for the identification of trends in reactions by the mass public in response to elite cues. More importantly, it allows for an examination of how effectively the mass public's reactions are disseminated throughout the communications network. In lay terms, this dissertation probes the question, "If civility falls in the forest and no one is around to hear it, does it really fall?"

Chapter Three begins with a brief chronological overview of the development of network analysis methodology in the social sciences. After establishing this context, I provide a summary of the data sets used in this dissertation, along with how they were collected and how they were applied to test the hypotheses presented in Chapters Four, Five, and Six. Additionally, this chapter discusses the different methods used to test each hypothesis, including content analysis and network analysis.

Chronological Overview of Network Analysis Methodology

Methodological Foundations of Network Analysis: The Pioneers

The roots of social network analysis can be traced back to J. L. Moreno's development of the concept of "sociometry" (1934), which describes social systems as "attraction-repulsion-neutrality systems" that form as a product of human preferences. Moreno argues that such human preferential systems do not lend themselves well to traditional statistical and observational methods that rely heavily upon "objective fact finding". Instead, an approach to examining social systems requires a "process of subjectification" when observing interactions between members of these systems (Moreno 1934, p 56).

Derived from the field of linguistics, subjectification accounts for contextual meanings as well as inherent meanings. Such an approach not only accounts for the preferences of one person, but it accounts for how those preferences relate to the preferences of others within the same system. Simply put, preferences within an interactive communications system must be viewed subjectively because they are dependent on the preferences of others in the system while, at the same time, influencing the preferences of others. This is directly relevant to analyzing online social networks, given the nature of networked publics outlined previously in this chapter.

The process of subjectification is at the heart of network visualization and is primarily expressed through the use of sociogram charts, which provide a physical representation of individual members of a group during a discrete period of time. It should be stressed that this picture can change if the temporal period of observation grows shorter or longer, as the individuals participating in the group may decrease or increase during different time spans. For example, if a researcher employed the use of sociograms to analyze the social network structure of students in a high school cafeteria,

he or she would almost assuredly find different results in a sociogram constructed during breakfast and one constructed during lunch, just as he or she would find different results in sociograms constructed during lunch periods on a Monday versus a Friday.

Sociograms are especially useful when examining networks of social interactions due to their flexibility in observational scope. Researchers can examine social interactions on a wide range of scales and across a wide range of time frames. Further, sociograms allow researchers the ability to survey the structure of an entire social system, identify communities of interest, and examine these communities with greater detail. Placing this process into the context of modern social network analysis, researchers can identify tightly clustered groups of discussion, "zoom in" on these groups to identify individual users, then pull out relevant characteristics of these users to answer more deterministic questions such as prevailing ideological leanings, frequency in communication, or participation in other social systems – to name a few.

Moreno's sociograms provide the basis upon which modern social network visualization is built. It must be stressed that social network analysis is largely non-deterministic; social network visualizations cannot be read in the same way as one reads a traditional Cartesian graph. In other words, one cannot isolate an individual within a sociogram and reach immediate conclusions based upon the size or "X/Y" grid location of that individual. Instead, the size and location of the individual within a sociogram must be interpreted *in relation to* other individuals within the sociogram. In contrast to the deterministic nature of traditional graphs, social network analysis is based upon relationships and, more importantly, the power of strong, interconnected, and influential relationships within a defined system.

Social Network Analysis Renaissance

Few major developments occurred within the study of social networks for several decades following Moreno's initial work. Some of the more notable contributions included research by Lévi-Strauss (1951) on rules governing kinship systems and work by Haray (1959) in the area of sociometric matrices analysis. Despite these significant steps forward in the mathematical foundations of social network analysis, the field went largely ignored by the social sciences. Perhaps the most impactful scholarship to originate from the 1950's school of social network analysis came from communications studies pioneer Everett Rogers's *Diffusion of Innovation* (1962), in which Rogers expanded upon concepts initially presented in Moreno's original 1934 work.

The Modern Era of Social Network Analysis

Despite these early advancements, there is a relative scarcity in volume of academic literature on social network analysis prior to 1990. As Bernard observes, "20 articles about social network analysis [are] listed in Sociological Abstracts" between the years 1960 and 1975. Conversely, "from 1990 to 2005, the number [is] over 3,000" (2005). Clearly, social network analysis gained traction within numerous areas of social research. Given the geometrically multiplying increases in computing power during the 1990's and through the 21st Century, it is likely that this sudden surge in academic study was largely facilitated by the availability of inexpensive and powerful computers capable of putting mathematical theory into practical action.

More recent work has been successful in clearly defining social network analysis as an organized paradigm. A prime example of such literature is found in Linton Freeman's *Development of Social Network Analysis: A Study in the Sociology of Science* (2004), in which Freeman defines social network analysis as consisting of four features. First, social network analysis does not focus on the attributes of actors, but rather upon the connections between actors. Second, social network analysis is built upon systematic data collection that focuses upon these ties. Third, social network analysis is strongly reliant upon the use of graphical representations. Last, social network analysis is dependent upon mathematical and computational tools to "make sense of the welter of information" that describes these ties (Freeman 2004, p 3). Modern social network analysts employ these techniques as a means for identifying trends in structural patterning; the *structure* of a network tells researchers a great deal about the *patterns* of speech, discussion, and relationships.

The emergence of social media – and of Twitter in particular – has provided a wealth of new opportunities to utilize social network analysis tools as a means for studying human behavior. Social network analysis goes far beyond the ability to produce "eye candy" in the form of striking and often beautiful visualizations. Social network analysis draws upon empirical data to provide context for relationships between individuals and, in doing so, reveals insight into issue trends, influential participants, and a treasure map for learning more about their predominant characteristics. In this respect, social network is a powerful tool for organizing massive amounts of empirical data and allowing the analyst to identify and focus upon empirical data that is most germane to his or her research question. Today, social network analysis is an invaluable tool for making

sense of the millions of interactions that occur on an hourly basis across multiple social network platforms.

Increasing numbers of researchers have enjoyed improved access to powerful social network analysis tools in recent years, due largely to the convergence of social media's widespread popularity with researchers' access to progressively powerful computers at reasonable costs. Such a convergence allows social network analysts in the "Twitter Age" to design research frameworks capable of sufficiently accommodating Freeman's four features of social network analysis. Twitter data provides information on the connections between actors and an application program interface (API) that allows for systematic collection of this data, while modern personal computers have the ability to process complex algorithms and convert them into graphical representations of social networks containing tens of thousands of actors. Additionally, powerful software is readily available that allows these graphics to be presented in a manner that clearly illustrates where neighborhoods of discussion form in relation to each other.

Despite the availability of such tools, academic contributions in the field of social network analysis have been sparse until recently. As Williams, Terras, and Warwick (2013) observed, only three academic papers published in 2007 focused upon Twitter in some form or another. This number rose to eight in 2008 and 36 in 2009, with the volume of academic research increasing significantly in the 2010's. It is likely that these increases was due to a combination of several factors. First, the 2010's saw a boon in the availability of progressively powerful and increasingly inexpensive hand-held mobile smart devices which were ideal for using a lightweight and easy-to-use application like Twitter. Second, researchers enjoyed a concurrent rise in computing power alongside a

corresponding drop in cost. Third, third-party software developers began releasing numerous inexpensive open-source tools allowing researchers to access the Twitter API at little to no cost to the researcher.

As social network analysis has gained more acceptance within academia, researchers have begun to focus on the issue of political polarization in social media. Early examples include studies examining the extent to which Twitter users cross ideological lines (Himelboim, McCreery, and Smith 2013) and challenge conventional wisdom with respect to media echo chambers (Barberá et al. 2015). These studies have represented valuable efforts to examine and quantify the nature of mass polarization and filter bubbles in social media. There are also increasing numbers of studies drawing upon large *n* datasets spanning several months worth of Twitter messages, many of which are designed to better understand how political information is shared and discussed within social media networks as a whole (Gruzd 2014; Morales 2015).

Instant Effects during High Stakes Political Events

This dissertation examines the relationship between elite polarization and mass polarization in social media through different temporal lenses. Chapters Four and Six test this question using very broad time horizons, measuring shifts in daily sentiment over the span of hundreds of days. However, Chapter Five tests this question using a comparatively small time horizon, measuring shifts in minute-by-minute sentiment over the span of a 100-minute presidential debate. Without a doubt, there is value in identifying shifts in sentiment over broad time horizons, as this offers to reveal trends with staying power. If a particular sentiment predominates for multiple days, it could

indicate a high level of salience. At the same time, broad time horizons are limited when it comes to measuring the mass public's reaction to specific types of elite cues. When measuring daily shifts in sentiment, the researcher must make a best guess as to the primary independent causal variable, given that a great deal of information is disseminated throughout a single day within the 24/7 news cycle.

Conversely, when measuring minute-by-minute shifts in sentiment, the researcher is equipped with far higher resolution when it comes to identifying the causal independent variables. Recent research in the field of political communication has focused on the phenomenon of hybrid media events, wherein social media is used alongside and during televised political events. Referred to as "dual screening", the viewing public expands its role from being a mostly passive consumer of information to being an active member of the political event itself. These roles often involve members of the general public using social media to engage in "lay tutelage" behavior, including the acts of fact-checking, correcting, counter-claiming, or highlighting biased reporting (Vaccari, Chadwick, & O'Loughlin 2015). Within this environment, the power of event participants to shape the narrative increases in proportion to their relative influence when measured by their number of active followers. In such a scenario, these event participants may serve as "bridging elites" with an influential power that can rival that of media and political elites (Freelon & Karpf 2015).

Other recent research has examined ways in which collective patterns of behavior on social media during media events differ from behaviors observed in traditional media contexts. One such example with particular relevance to this dissertation found a multitude of notable behavioral characteristics among individuals who engage in "dual

screening" during media events (Lin et al. 2014). Specifically, individuals who engaged in social media communication during televised political events exhibited significantly lower levels of interpersonal communication. There was also a tendency for individuals to exhibit a higher level of concentrated attention to topics, while also engaging in far more "re-tweeting" and replying. Most significantly, Lin et al. found that elites tended to be the predominant beneficiary of these sorts of behaviors.

Findings such as these have been extremely valuable in better describing and explaining the roles individuals play in politically charged social media environments, as well as the types of individuals who play these roles. At the same time, more needs to be known regarding how elite messages originating in the traditional media environment shapes the network structures that facilitate social media participants' roles, whether they be participating in a bridging elite, lay tutelage, or passive observer function. Specifically, a better understanding of the nature of how these network structures shift can provide valuable context for interpreting the relative influence of participants regardless of their roles.

Data Collection

This dissertation draws upon an original data set created by accessing the Twitter API via the NodeXL Excel template (Social Media Research Foundation, 2017) on a daily basis from September 1st, 2015, through February 1, 2017. The full data set used in the three empirical chapters of this dissertation consists of approximately 13,000,000 tweets and 260,000,000 words. This data is part of a larger daily collection regimen which (at the time of this writing) conducts searches for nearly 40 unique names,

hashtags, or Twitter accounts, allowing for the curation of approximately 250,000 tweets per day. While other options exist for conducting Twitter API searches, NodeXL was chosen due to its ability to perform a wide range of search functions while keeping the financial costs to the researcher extremely low.

A primary limitation of using the Twitter API for data collection is that the results returned for high-frequency search terms represent a sample of approximately 1% all tweets during the specified search time frame. According to Twitter, these results are "a statistically relevant sample". Such a rather vague explanation is somewhat bedeviling to social scientists, as it limits the ability to establish the extent to which this ostensibly "random" data is representative of the larger population.

Recent research indicates that data acquired via the Twitter API may not be very random after all. For example, when comparing data sets compiled through multiple Twitter API searches, Joseph et al. (2014) found that on average, more than 96% of tweets found in one sample were also found in all other samples. Despite such similarities, the content found in the subset of non-matching samples did not differ significantly in terms of tweet structure or user popularity. It should be stressed that such limitations apply to any scientific study using high volumes of data acquired via the Twitter API.

Given that Twitter operates as a publicly-traded for-profit business, it is likely that Twitter has a financial motivation for not allowing potential business competitors to have insight into the nature of their randomization models, or any other type of proprietary algorithms or code. It is also worth noting that use of the Twitter API seems to be the preferred method for social media researchers, as the only option for avoiding Twitter's

black box of "statistically relevant samples" is to pay for access to the Twitter "firehose"

or to purchase data from companies specializing in storing hundreds of billions of

archived historical tweets. Such access allows researchers access to every single tweet

ever sent, typically acquired by purchasing a given volume of tweets (e.g. 100,000)

mentioning a give key term (e.g. "Donald Trump") over a given time frame (e.g.

11/1/2016 – 11/31/2016). While such an option provides researchers with the ability to

fine-tune the creation of their data sets without needing to perform daily searches, it is

also an option that is often prohibitively expensive.

In sum, while there are limitations with respect to how representative the Twitter

API's "random" data is of the larger population, these limitations are shared by most

researchers in the social sciences. Rather than being a condition that disqualifies the

validity or generalizability of results obtained through Twitter API data, it is more of a

caveat to be considered when analyzing the results of any study using such data.

Content Analysis

Content analysis is a valuable method for quantifying the frequency of words in

bodies of text. Typically, content analysis utilizes specialized dictionaries which

organize words within specific categories. Through this process, various meanings and

themes within the text begin to emerge. Content analysis is used in this dissertation in

two main ways. First, I use content analysis to measure positive and negative sentiment.

Second, content analysis is used to measure the frequency of words expressing specific types of reactions, including emotional and higher-level thought processes.

The primary tool for conducting content analysis was Lexicoder 3.0, a software application developed by Mark Daku, Stuart Soroka, and Lori Young at McGill University. This software was used in conjunction with the Lexicoder Semantic Dictionary (Daku, Soroka, and Young 2016) and the Regressive Imagery Dictionary (Martindale 1975, 1990). The Lexicoder Semantic Dictionary draws upon a dictionary of approximately 5,000 words and is designed to measure the positive and negative sentiment in political texts. The Regressive Imagery Dictionary is comprised of approximately 3,000 words divided into three primary categories with 44 sub-categories. These categories are divided into Primary Processes, which include drive, sensation, defensive symbolization, regressive cognition, and Icarian imagery; Secondary Processes, which include abstraction, social behavior, instrumental behavior, restraint, order, temporal references, and moral imperative; and Emotions, which include positive affect, anxiety, sadness, affection, aggression, expressive behavior, and glory. The aforementioned categories are used in this dissertation to provide a more accurate measurement of affective rhetoric. Measuring specific types of thoughts and emotions (such as abstraction and aggression) provides an element of specificity that is more precise than mere positive and negative sentiment.

Content analysis is performed by processing text obtained via daily NodeXL searches of the Twitter API through Lexicoder 3.0. The initial results provided by Lexicoder 3.0 are in the form of raw frequency counts. However, given fluctuations in total volume of total tweets observed on individual days, the use of basic frequency

counts lacks descriptive accuracy. Rather put, a measurement of 100 occurrences of positive words on a day with 10,000 total words is substantively different than a measurement of 100 occurrences of positive words on a day with 20,000 total words. Hence, an essential step in content analysis involved converting raw frequency counts into rates by dividing them into the total number of words tweeted during each day of observation. In order to improve the comparative utility of these rates, they are reported as "rate of anxiety per 1,000 words", "rate of aggression per 1,000 words", and "rate of negativity per 1,000 words".

Time-series Analysis

Measuring temporal relationships between actions by political elites and reactions by the mass public is a central part of testing the main research question of this dissertation. Social media provides an excellent environment for testing whether or not a relationship exists between elite and mass polarization because of its ability for the mass public to express their instant reactions to comments made by political elites. Time-series analyses were a natural choice for such tests, especially given the rich and extensive data sets that were collected during the composition of this dissertation. Time series analysis is especially useful for examining trends, as well as for identifying potential cause and effect relationships. If it is true that certain types of elite cues elicit polarizing types of behavior in the mass public, evidence of such relationship will become evident if repetitive patterns emerge through time-series analysis.

Time series analysis is used in this dissertation to measure for shifts in sentiment and affective rhetoric in several ways. In Chapter 4, I use time-series analysis at the

macro level, by drawing upon 366 daily observations spanning 7 different candidates over a full calendar to test for cause-and-effect relationships between proximity to a presidential primary election or debate and a resulting rise in mass affective rhetoric. In Chapter 5, I use time-series analysis at the micro level by analyzing each individual minute of tweets captured live during the first presidential debate.

In Chapter 6, I use time-series analysis to test for the existence of the proposed "Trump Effect", which argues that controversial remarks by Donald Trump during the primary and general elections were encouraging aggressive behavior in some Americans while creating fear and anxiety among others. Time-series analysis is performed by measuring specific reactions in the mass public to especially controversial remarks made by a high profile political elite: Donald Trump. In sum, time-series analysis can be a valuable tool for testing temporal relationships between multiple variables. This dissertation takes advantage of the methodological flexibility of time-series analyses to measure such relationships within the scope of a variety of unique political scenarios.

Network Analysis

The empirical studies presented in subsequent chapters make extensive use of network metrics and network visualizations created with Gephi 0.8.2 for Macintosh, which an open source and multiplatform application especially designed for creating visual graphs of any type of network. It must be stressed that social network analysis is primarily about relationships between people and groups. These relationships are observable through the use of both descriptive network metrics and visualizations. Network analysis is especially useful for answering questions about mass political

behavior in social network environments because it allows researchers to observe relationships in a way that is not feasible through traditional means such as survey data or regression models.

Network Visualization

Gephi is an especially useful tool for testing the hypotheses posed in this dissertation because it allows for a comparison in networks between days with high rates of polarizing language and days with low rates of polarizing language. Such differences are manifested either by graphs demonstrating clusters of mostly unconnected discussion on the fringes of the network map or graphs demonstrating a high concentration of interconnected nodes near the center of the network. In this respect, daily variances in network architecture serve as dependent variables, while the rates of polarizing language serve as the independent variables. If it is true that polarizing behavior on the part of elites causes polarizing behavior in the mass public, such tendencies can be observed in the shape and structure of groups within these networks, as well as through an examination of shifts in network metrics.

Social network analysis allows researchers to identify links between the types of discussion occurring within social networks and the resulting neighborhoods they form. This is an especially important aspect to keep in mind when interpreting network visualization graphs. A cursory understanding of how Gephi accomplishes this is important for making the most sense of the network visualization portion of the dissertation.

First, a specially formatted data file is imported into Gephi. The very least the software requires is a "node" (or the sender, receiver, or re-tweeter of a message) and an "edge" (or the connection between two nodes), to which Gephi assigns what can best be described as magnetic values to the nodes and edges. In simplest terms, nodes repel each other while edges draw nodes together. This repulsion is especially strong where individual nodes with no edges are concerned. The net result (after many hundreds of algorithmic iterations) is that as the number of nodes and edges in a network increase, the forces of attraction and repulsion create a picture of areas in the network where the web of communication is most dense. As a result, nodes that are more interconnected with other nodes form "neighborhoods" of discussion and, in doing so, create a stronger gravitational force that repels smaller "neighborhoods" of nodes.

These neighborhoods are of special interest because they define areas where people are gathering in virtual spaces. Since the data sets for this dissertation are built around specific search terms (e.g. "Donald Trump" or "Hillary Clinton"), these neighborhoods illustrate where people are gathering to discuss a particular candidate. More importantly, it can bee seen whether communication is predominantly occurring between influential members of the network, or whether it is more widely dispersed amongst less influential members of the network. An explanation of the network metrics used to estimate influence is outlined in the following subsection.

Network analysis – when combined with content analysis – provides a picture of both the nature of political discussion and the efficiency with which this discussion spread throughout members in the network. For example, if content analysis on a specific date demonstrates a relatively high rate of aggressive affective rhetoric, but

network analysis suggests a weakly connected network, one could infer that the impact of such rhetoric has been mitigated. Conversely, if content analysis alone was used in this scenario, the more likely inference would have been an overestimation of the aggressive rhetoric's overall impact on members in the network as a whole. Simply put, content analysis provides valuable aggregate measures – but network analysis puts these aggregate measures into context by taking into account the critical variable of network structure.

Network Metrics

While there are a wide range of methods available for testing various characteristics of networks, this dissertation focuses primarily upon metrics describing the frequency, centrality, and influence of members in a given network. Of these three measures, centrality and influence are especially critical to this dissertation's research question, as they provide additional information that goes beyond the aggregate measure of frequency. Frequency is most easily defined as the volume of messages originating from or being directed to an individual member of the network. Centrality is measured by using network visualization to observe an individual's relative location within the wider network, using the methods described in the previous subsection. Influence is measured by calculating the influence of an individual within a network based upon the influence of the people he or she is connected with.

These measures matter in the social media communication environment because frequency is not the same as influence – nor is it the same as centrality. These differences are important because Individual A may have a large number of immediate

connections, but if those connections are not influential, Individual A's relative influence in the network is limited. Conversely, Individual B may have a small number of immediate connections who are very influential and, in such a scenario, Individual B would likely have far more influence in the network than would Individual A. Acknowledging and testing for these differences is a critical aspect of this dissertation because it accounts for network dynamics that enhance or diminish the frequency of a message. A message may appear relatively frequently, but if it is not shared with influential or central members in a network, the impact of the message will be diminished.

The primary network metrics used in this dissertation to test influence and centrality are *Average Community Size* and *Average Path Length*. Average Community Size provides an indication of how densely clustered neighborhoods of discussion are in the network. This metric is calculated by first running a modularity algorithm. Generally speaking, modularity is a measure of a network's tendency to gravitate towards clusters of communities. High modularity values suggest the existence of more sophisticated internal structures and, in turn, help describe how a network is compartmentalized into sub-networks (Blondel et al. 2008). Accounting for modularity matters because it provides concrete values when looking for evidence of mass polarization. Networks with high average community sizes indicate a tendency of individuals in a particular network to cluster together around certain topics, themes, or pieces of information and, in turn, are consistent with polarized behavior.

Average path length provides an indication of a network's efficiency by providing a measure of how easily messages can travel throughout a given network. It is best

understood as a measurement of the number of steps it takes for a message to travel

across a network (Brandes 2001; Albert and Barabási 2002; Newman 2003).

Specifically, a message originating from one member (node) in a network and

terminating with another member (node) in a network represents a path length of one. A

message originating from one member (node) in a network, being received by a second

member (node) in a network, and being sent to a third member (node) in a network

represents a path length of two. It should be noted that in this second example, the path

length is still two even if the intermediary member (node) sends this message to multiple

other members before the message reaches its terminus.

In order to establish context for average path length values, I report these values

as a percentage of network diameter. In brief, network diameter is defined as the longest

path distance between any two members (nodes) in a network. As such, it provides one

estimation of the network's overall size. Since network diameter varies from graph to

graph, so too does the relative value of the average path length. For example, members

of a network with an average path length of 3 and a diameter of 3 are far more

interconnected than members of a network with the same average path length of 3 but

with a diameter of 6. Such a measure is particularly useful to investigating this

dissertation's research questions because it helps identify networks which are conducive

to homophily. Specifically, networks with low average path length to network diameter

ratios indicate the likelihood of smaller and potentially polarized clustering of network

members. As is the case with each of the network metrics used in this dissertation, this

measure is best used in conjunction with other network metrics. It is unlikely that a

single network metric can best describe network structure on its own. Rather, the most

accurate assessment of a network's dynamics is best attained by accounting for multiple characteristics.

Conclusion

This dissertation relies heavily upon a mixed-methods approach combining large scale data collection, extensive content analysis of hundreds of millions of words across nearly two years of daily observations, and creative application of cutting edge network visualization tools. While time-series analyses are extremely useful in measuring for trends or cause-and-effect relationships, network analysis helps to put any such findings into context. Analyzing social networks requires special attention to be given towards relationships and, more importantly, how the strength of relationships modify the influence and centrality of messages being sent and received. This purpose of this brief chapter has been to familiarize the reader with the methods used for analyzing these relationships in order to maximize their descriptive power in addressing the main research question of this dissertation. Each of the following empirical chapters employs network analysis methods in a slightly different way to provide three unique perspectives on the relationship between elite polarization and mass polarization in social media.

Dissertation Chapter 4

Trolling Twitter:

Social Network Patterns of Affective Rhetoric during the 2016 GOP Presidential Primary

Eric C. Vorst

Introduction

In this first empirical chapter, I test the theory that as a political event draws nearer, there will be an increase in the general combativeness among those who discuss politics. This is similar to a "big game" effect, where people argue about sports near the Super Bowl or World Series. The 2016 Republican primary provided an excellent testing ground for this chapter, as the primary featured an initial field of 19 viable candidates. The unpredictable and often volatile nature of the primary field was magnified by the candidacy of businessman Donald Trump.

This chapter first tests the proposition that the use of affective rhetoric in social networks is positively correlated with its temporal proximity to a political event. Additional tests are performed to determine if a political figure's ranking in national polls increases the quantity of affective rhetoric. This first set of tests are designed to answer the first half of a two-part question that recurs in subsequent chapters: "Does it happen?"

Second, this chapter tests how social networks react to increases in affective rhetoric. As has been discussed in Chapter 2, there is ample evidence demonstrating the nature and structure of the social media communication environment is significantly different than the nature and structure of traditional communication environments. While prior studies have examined the effects of affective rhetoric in traditional environments, very few have done so within the unique communications environment of social media. In brief, the second set of tests within this chapter were designed to measure the extent to which increases in affective rhetoric in the aggregate impacted the reach and influence of such rhetoric. This second set of tests were designed to answer the second half of a two-part question that recurs in subsequent chapters: "Does it matter?"

Research Questions

In order to test for a link between the proximity to a political event and a subsequent rise in social network polarization, this chapter investigates two sequentially dependent research questions. Specifically, the first research question looks for a causal link between politically-charged events in the "real world" and rises in affective rhetoric on social media, while the second research question measures how the unique nature of social networks impact the extent to which any such effects matter.

Research Question #1:

Does the proximity to a political event result in an increase in affective rhetoric in social media?

Research Question #2:

Do participants in open social networks tend to self-police especially controversial forms of affective rhetoric?

Hypotheses

Hypothesis #1: The use of polarizing language in social networks is positively correlated with its temporal proximity to a political event.

Hypothesis #2: Online social networks tend to sequester unproductive messages in favor of constructive debate.

Data and Methodology

Data Collection: NodeXL

Data for this chapter was gathered by using NodeXL to perform a search for each of the top three GOP presidential candidates over the span of 366 consecutive days during the primary election season. The full data collection consisted of 1,047 individual data sets, with a total number of observations tallying roughly 8.4 million tweets consisting of approximately 160,000,000 words.

The data returned from these searches was comprised only of messages sent roughly within the preceding 24 hours, so it was essential to engage in these searches every single day over the seven month span of the data gathering process. These searches were often hampered by idiosyncrasies of the NodeXL data collection process, including unpredictable search interruptions due to shifting Twitter volumes. For example, in some cases searches would be cut short due to the NodeXL API shifting into a "summary" mode to account for extremely high volumes of Twitter data at that given time, while in other cases simple fluctuations in Internet connectivity could trigger the software to end collection. Further, Twitter sets a limit to how many free requests for data can be made during a set period of time, which was a cause for frequent concern. However, due to an unyielding tenacity in data collection, the vast majority of daily observations ranged from between 5,000 and 25,000 tweets.

Candidate Search Terms

The selected search terms focused on the candidates' full names (e.g. "Donald Trump" or "Jeb Bush") rather than their Twitter tags (e.g. "@realDonaldTrump" or "@JebBush") in order to filter out tweets that were coming directly from the candidate themselves. Actual candidate tweets have been collected separately using a search term that only gathered messages being sent from their official Twitter account. This approach allowed me to approach the data from an alternate perspective. First, it allowed me to focus only on people who were talking about the candidates, rather than including messages from the candidates themselves. Second, it placed the focus on the semantic content of the tweet rather than the "tagging" process. Rather put, I wanted to look at how people talked about these candidates when using their names in a sentence that was not necessarily "tagged" with the candidate's Twitter name (e.g. @realDonaldTrump). Another way to view the difference is to describe it as looking at how people talk *about* someone, rather than how they talk *to* someone.

Admittedly, the process of examining how people talk *about* someone (e.g. "Donald Trump") versus how people talk *at* someone (e.g. "@realDonaldTrump") is not entirely in line with conventional approaches to studying Twitter communication. After all, the foundation of personal identity on Twitter lies in the ampersand tag; Donald Trump's personal identity on Twitter is "@realDonaldTrump", not "Donald Trump". Further, the vast majority of studies on Twitter communication focus on users' ampersand tags as the point of reference. However, the field of social network analysis is one that invites researchers to approach questions from alternative perspectives. There are few known formal studies on how people communicate differently on Twitter when

talking *about* someone versus talking *at* someone. This dissertation will offer a first step in this line of inquiry, thus opening up opportunities for future research.

Content Analysis

Content analysis was performed to determine the total number of positive and negative words within the individual data set for each candidate on each day of observation. Positive and negative words were chosen as a proxy for affective language, which can be an indicator of polarization (Iyengar 2012). The primary tool for conducting the content analysis was Lexicoder 2.0, a Java-based software application developed by Mark Daku, Stuart Soroka, and Lori Young at McGill University. This software was used in conjunction with the Lexicoder Sentiment Dictionary, which is designed to capture the sentiment of political texts. The Lexicoder Sentiment Dictionary does so by assigning a positive or neutral value to a defined set of over 3,500 words. After text is processed through Lexicoder 2.0 and the Lexicoder Sentiment Dictionary, a raw count of total positive and negative words per tweet was produced. These counts were collected, compiled, and recorded for each of the candidates' daily data sets. In order to make these numbers more usable, they have been reported as "positive words per 1,000 words" or "negative words per 1,000 words".[1]

While positive and negative words have the potential to highlight affective language, mere positive and negative language alone cannot account for some of the most polarizing types of language. As such, I selected for analysis one of the most profane words in the English language: the "F Word" (or "F*ck"). It should be noted that the

---

[1] See Chapter 2

different incarnations of this word are not always necessarily negative in tone. Indeed, this word is often used to express joy, happiness, confusion, excitement, or any number of emotions that are not expressly negative. However, this word represents the epitome of vulgarity and profanity and, as such, serves as a valuable proxy for analyzing affective rhetoric, regardless of the tone with which it is used.

A basic Excel "find and replace" function was used to gather a raw count of how often any variation of the "F Word" appeared in each of the 105 individual days of observation across all three candidates. A wildcard symbol was used at the end of the "F Word" in order to capture as many variations of the word as possible, such as "F*cks", "F*cker", "F*cking", "What the f*ck", and so on.

Methods

Time Series Analysis

Hypothesis #1 and the first half of Hypothesis #2 were tested using a battery of time-series analyses measuring the rise and fall of affective and uncivil language in proximity to a political event. The dependent variables measured include the rate of affective language and the rate of extremely uncivil language. The independent variables are represented by the days during which a primary debate or primary election occurred. I record the estimated viewership of each debate and the delegates at stake on each primary election date in order to place its political importance in context. The logic behind this decision is that debates with higher viewership and primary elections with more delegates at stake are likewise more important political events. If there is a positive

relationship between political events and mass polarization, we should see a spike in the use of uncivil or affective language on these dates.

Given the wide time frame of daily observations gathered over a full year, it was necessary to identify time frames within the primaries. Simply put, it is difficult to create a timeline that is easily interpreted when observing 366 data points in time. In order to best present the time-series data, I divided the primary season into three individual stages. One unintended benefit of this approach is that it allowed me to isolate and compare three different phases of the primary season, each of which featured distinctly different levels of competitiveness.

The first phase was called the "Cattle Call Stage" (9/1/2015 – 1/31/2016) and was a period during which there were six different candidates who occupied a top three spot in the Real Clear Politics polling average at any given time. The "Cattle Call Stage" featured six televised debates, but there were zero primary election dates. The second phase was called the "Competitive Stage" (2/1/2016 – 3/31/2016) and was a period which tracked four candidates who reached a top three spot in the Real Clear Politics polling average. The "Competitive Stage" featured five televised debates and eleven primary election dates. The third and final phase was called the "Confirmation Stage" (4/1/2016 – 6/30/2016) and was a period during which the same three candidates consistently rested in the top three spots of the Real Clear Politics polling average. There were zero televised debates during the "Confirmation Stage"; however, there were nine primary election dates.

Network Analysis

The second half of Hypothesis #2 was tested using Gephi for Mac, which is an open source and multiplatform application especially designed for creating visual graphs of any type of network.[2] It is worth stressing again that social network analysis is primarily about relationships between people and groups. Social network analysis allows for the identification of links between the types of discussion occurring within social networks and the resulting neighborhoods they form. This is an especially important aspect to keep in mind when interpreting network visualization graphs, as their interpretive power lies in comparative relationships rather than in absolute Cartesian values.

Gephi is an especially useful tool for testing Hypothesis #2 because it allows for a comparison in networks between days with high rates of affective rhetoric and days with low rates of affective rhetoric. Such differences are manifested either by graphs demonstrating clusters of mostly unconnected discussion on the fringes of the network map or graphs demonstrating a high concentration of interconnected nodes near the center of the network. In this respect, daily variances in network architecture serve as dependent variables, while the rates of polarizing language serve as the independent variables. Additionally, Gephi is an extremely useful tool for isolating specific types of language and identifying the areas of networks where such language occurs. Network visualization graphs are also used to test Hypothesis #2 by analyzing the extent to which members of Twitter networks using extreme forms of affective rhetoric are successful in gaining visibility and influence for their messages.

---

[2] See Chapter 3

Results and Analysis

Hypothesis #1: The use of polarizing language in social networks is positively correlated

with its temporal proximity to a political event.

When testing Hypothesis #1, results suggested that there is a moderate and

positive relationship between rates of affective language usage on a given day and its

proximity to a political event. During the "Cattle Call Stage" (9/1/2015 – 1/31/2016),

positive and negative word usage rate increased following 5 of 6 debates (83%), while "F

Word" usage rate increased slightly during 3 of 6 debates (50%). During the

"Competitive Stage" (2/1/2016 – 3/31/2016), positive and negative word usage rate

increased either on or immediately following 7 of 11 primary election dates (64%) and 3

of 5 debates (60%), while "F Word" usage rate increased either on or immediately

following 8 of 11 primary election dates (73%) and 5 of 5 debates (100%). During the

"Confirmation Stage" (4/1/2016 – 6/30/2016), positive & negative word usage rate

increased either on or immediately following 6 of 9 primary election dates (67%), while

"F Word" usage rate increased either on or immediately following 7 of 9 primary election

dates (78%).

Initial results also suggest there is a moderate and positive relationship between

rates of affective language usage and the *importance* of a political event. When looking

at all primary election dates (2/1/2016 – 6/7/2016), the rate of "F Word" usage in tweets

mentioning Donald Trump increased on election dates with 50 or more delegates at stake

in 11 of 12 cases (92%). Conversely, rates of "F Word" usage in tweets mentioning

Donald Trump increased on election dates with fewer than 50 delegates at stake in 4 of 8 cases (50%). During the "Competitive Stage" (2/1/2016 – 3/31/2016), the rate of "F Word" usage in tweets mentioning Donald Trump increased surrounding the debates preceding Super Tuesday (3/1/2016). The rate of "F Word" usage in tweets mentioning Donald Trump spiked significantly on Super Tuesday (3/1/2016) and continued to do so for each subsequent primary election date during March. During the "Confirmation Stage" (4/1/2016 – 6/30/2016), the rate of "F Word" usage in tweets mentioning Donald Trump decreased and remained stable after he became the sole remaining primary candidate (5/10/2016). A notable exception was found on 6/12/2016, which was the date of the Pulse Nightclub shooting in Orlando, Florida. On this date, Donald Trump registered the 2nd highest rate of "F Word" usage out of almost 900 daily observations in the data set.

In sum, time series analyses of the Hypothesis #1 provided consistently compelling support for a temporal relationship between affective rhetoric and a political event. Affective rhetoric in the form of positive and negative sentiment rose significantly during or immediately following days where a primary election or primary debate occurred. Increases in the rates of "F Word" usage aligned even more frequently with the same political events. Even stronger support for Hypothesis #1 was found when accounting for the relative importance of a political event. This was most evident when comparing primary election dates where fewer than 50 delegates were at stake with dates where 50 or more delegates were at stake.

Hypothesis #2: Online social networks tend to sequester unproductive messages in favor of constructive debate.

The first set of tests of the second hypothesis employed the use of network visualization graphs across a span of individual days of observation. This battery of tests was designed to identify differences in the overall shape and organization of networks with high and low incidences of extreme affective rhetoric: the "F Word". These tests were designed to provide a broad contextual picture of these networks' structures and, in doing so, provide visual evidence of their overall connectivity and density. Such contextual evidence is important because varying levels of density and connectivity are signs of varying levels of network polarization. First, I used Gephi for Macintosh[3] to create network visualizations for days with the five highest rates of "F Word" usage and days with the five lowest rates of "F Word" usage. Such an analysis is not intended to provide conclusive predictive powers, but rather is intended to demonstrate how unmoderated social networks behave differently when comparing periods of extremely high affective rhetoric to periods of low affective rhetoric.

[Insert Visualization 1.1]

The first set of visualizations compare days with the top five most frequent use of the "F Word" alongside days with the top five least frequent use of the "F Word". When comparing the visualizations side by side, it is clear that the networks are indeed different in structure, density, and neighborhood dispersion. This is evident in the visualizations in the first row (representing the days with the highest rates of the "F Word"), which

---

[3] See Chapter 3

demonstrate relatively small yet dense clusters of discussion dispersed throughout the peripheral regions of the networks. Additionally, the surrounding spaces in these networks is littered with large collections of "one-step" nodes, where individuals are either sending or receiving a message one time and no more. Last, there is little space between the communications in the center of the network and the largely unconnected ring of nodes encompassing the perimeter of the network. This suggests that there are fewer nodes in the center of the network and fewer connections between communications in the center of the network and surrounding neighborhoods of discussion. In turn, there is less of a repulsive force on surrounding nodes with few to no connections (indicative of non-influential members in the network).

Conversely, visualizations in the second row (representing the days with the lowest rates of the "F Word") demonstrate relatively large clusters of shared conversation. Further, the surrounding space consists largely of smaller neighborhoods of conversation connected directly to the larger central conversation. Last, the largely unconnected ring of nodes encompassing the perimeter of the network is, in most cases pushed, further out from communications in the center of the network. This suggests that there are more connected nodes in the center of the network and more connections between communications in the center of the network and surrounding neighborhoods. As a result, there is more of a repulsive force on surrounding nodes with few to no connections. These differences in network structure are more clearly evident when viewing the visualizations in higher resolution:

[Insert Visualization 2.1]

[Insert Visualization 2.2]

[Insert Visualization 2.3]

[Insert Visualization 2.4]

[Insert Visualization 2.5]

A final set of visualizations was created in order to locate these highly polarizing messages, examine their content, and identify their position with relation to the larger network of conversation and its influential participants. The following visualizations isolate specific tweets containing the "F Word" by changing the color and size of nodes containing these words. Doing so makes it easier to identify their location with respect to the center of discussion.

[Insert Visualization 3.1]

[Insert Visualization 3.2]

[Insert Visualization 3.3]

[Insert Visualization 3.4]

[Insert Visualization 3.5]

Isolating messages with the "F Word" and illustrating their position with respect to the center of discussion produced some fascinating results. When examining the days with the top five rates of "F Word" usage, it was found that messages containing the "F Word" were clustered outside of the center of discussion in all cases. While the number of tweets with the "F Word" may have increased in the aggregate, they were not shared with influential members of the network. As such, the impact and reach of these expressions of extremely affective rhetoric failed to gain influence in discussion about Donald Trump on these days.

These findings are important because they demonstrate two key points. First, they demonstrate that frequency does not equal influence in the social media environment. After analyzing the dates with the five highest rates of "F Word" usage (out of 366 days of observation), in no case did tweets containing the "F Word" successfully enter the core of shared discussion. Instead, in every instance these dense bubbles of extremely affective tweets were sequestered by members of the network. Second, these findings provide valuable insight into the behavior of unmoderated and unfiltered networks. Rather put, these findings provide evidence of "self policing" on the part of Twitter users when it comes to extremely vulgar language. As such, there is compelling evidence of an aversion to such types of language among those who wish to discuss, share, and even debate content related to a major political figure.

Conclusions and Discussion

This chapter has presented evidence that social networks experience both increases in positive and negative language as well as extremely polarizing and unproductive words as a political event nears. A candidate's movement in the national polls has some influence, but this appears to be moderate in comparison. Most importantly, there is strong evidence that despite an increase in such language, members of social networks tend to sequester such language in favor of constructive debate. Strikingly, such behavior is exhibited in the absence of significant controls over the content of the message or the anonymity of the messenger, which suggests there may be an element of "self policing" inherent in large, public, and open online social networks. The findings in this chapter could be considered as cause for a modicum of optimism in a
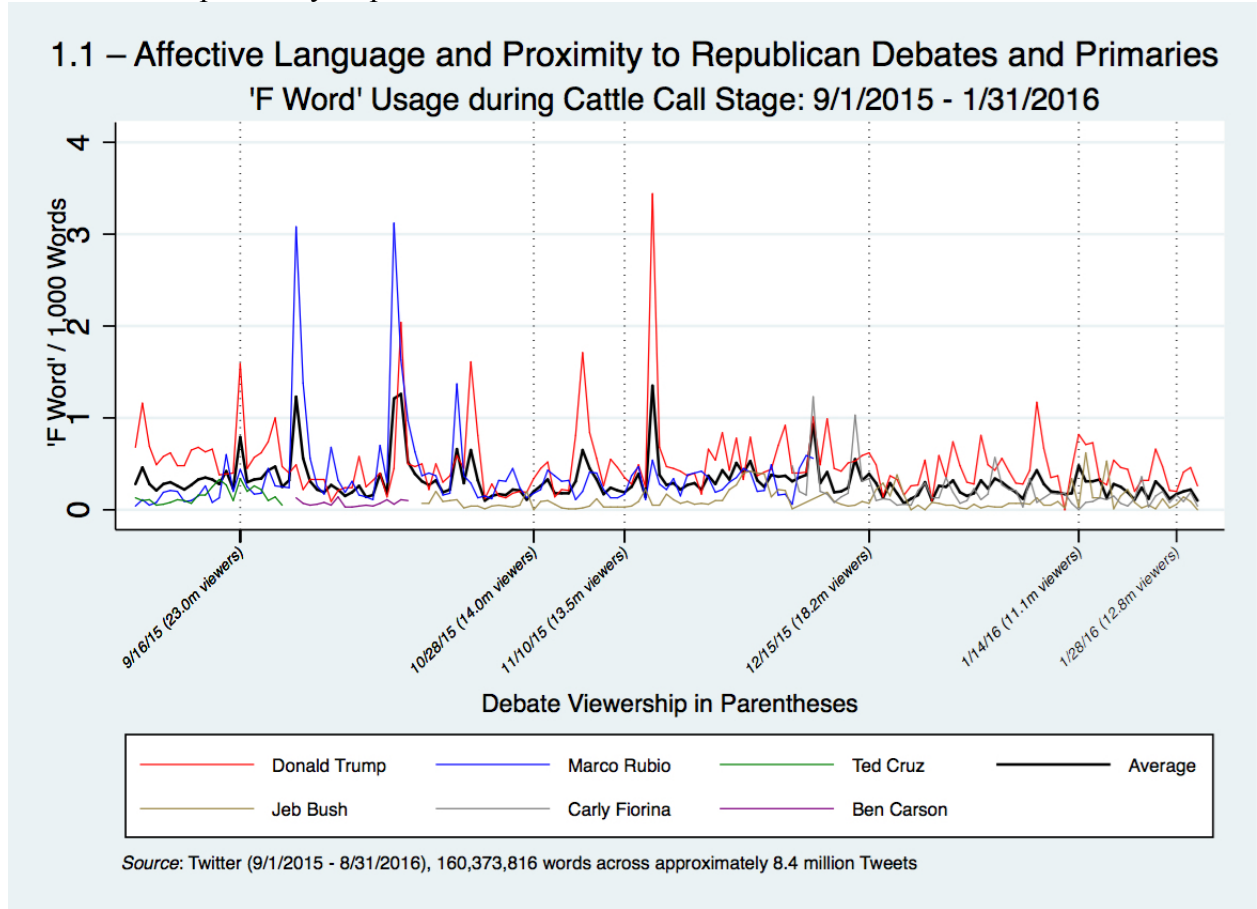
highly polarized climate where an inordinate amount of attention is often given to the loudest voices, and where perhaps a bit too much worry is wasted on the Twitter Trolls.

At the same time, it is important to acknowledge that this chapter examined these relationships within the context of a primary election. This context matters, because the nature of U.S. presidential primary elections is vastly different than the nature of U.S. presidential general elections. These differences are both fundamental and significant, as they include (but are not limited to) different campaign strategies, different election schedules, different media coverage, and different levels of citizen involvement. The nature, effects, and impact of these differences could – and likely do – influence the nature, effects, and impact of political discussion on social media. While subsequent chapters in this dissertation examine two aspects of the U.S. presidential general election, they do so while testing different hypotheses than were presented in this chapter. This provides several opportunities for future research examining whether the "big game effect" varies when comparing the unique battlegrounds of the U.S. presidential primary and U.S. presidential general election political environments.

As social media continues to grow as a platform for individuals to gather, process, share, and debate political issues, its potential for affecting mass polarization also grows. Measuring how the frequency of affective rhetoric on social media changes over time with relation to the occurrence of a political event is an important step in better understanding how elite cues influence mass polarization in different information environments. More importantly, measuring how social networks treat affective rhetoric provides valuable insight into whether increases in such language will have a net negative

effect on constructive debate, or whether it will be pushed to the fringes where its

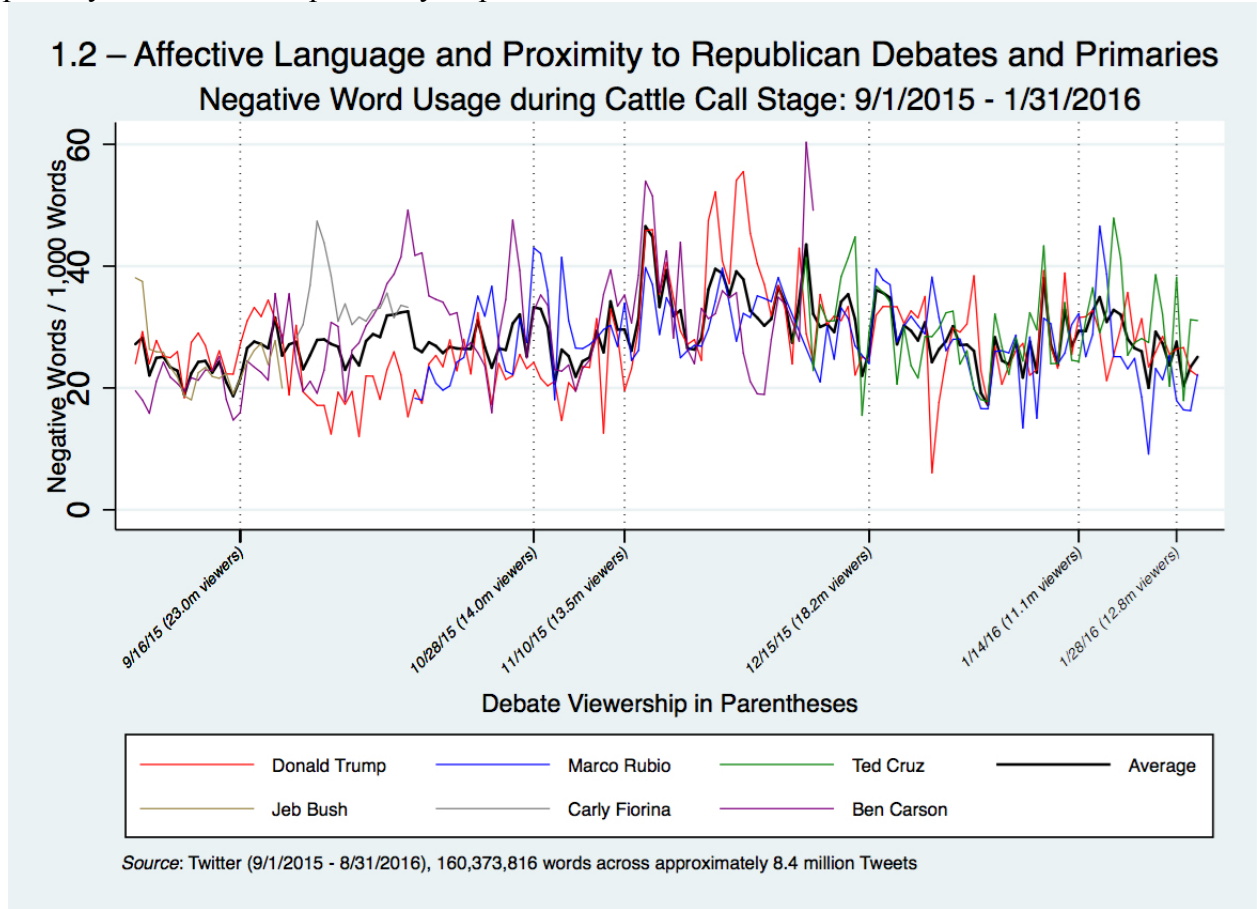audience will be smaller and less influential.

Graph 1.1: Comparing frequency of "F words" in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries

1.1 – Affective Language and Proximity to Republican Debates and Primaries
'F Word' Usage during Cattle Call Stage: 9/1/2015 - 1/31/2016

Source: Twitter (9/1/2015 - 8/31/2016), 160,373,816 words across approximately 8.4 million Tweets

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Frequency of "F words" in tweets identified using Lexicon 2.0 content analysis software. Frequency is measured in number of occurrences per 1,000 words.

Graph 1.2: Comparing frequency of negative affect in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries

1.2 – Affective Language and Proximity to Republican Debates and Primaries
Negative Word Usage during Cattle Call Stage: 9/1/2015 - 1/31/2016

Debate Viewership in Parentheses

Donald Trump   Marco Rubio   Ted Cruz   Average
Jeb Bush   Carly Fiorina   Ben Carson

Source: Twitter (9/1/2015 - 8/31/2016), 160,373,816 words across approximately 8.4 million Tweets

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Negative words in tweets identified using Lexicon 2.0 content analysis software. Frequency is measured in number of occurrences per 1,000 words.

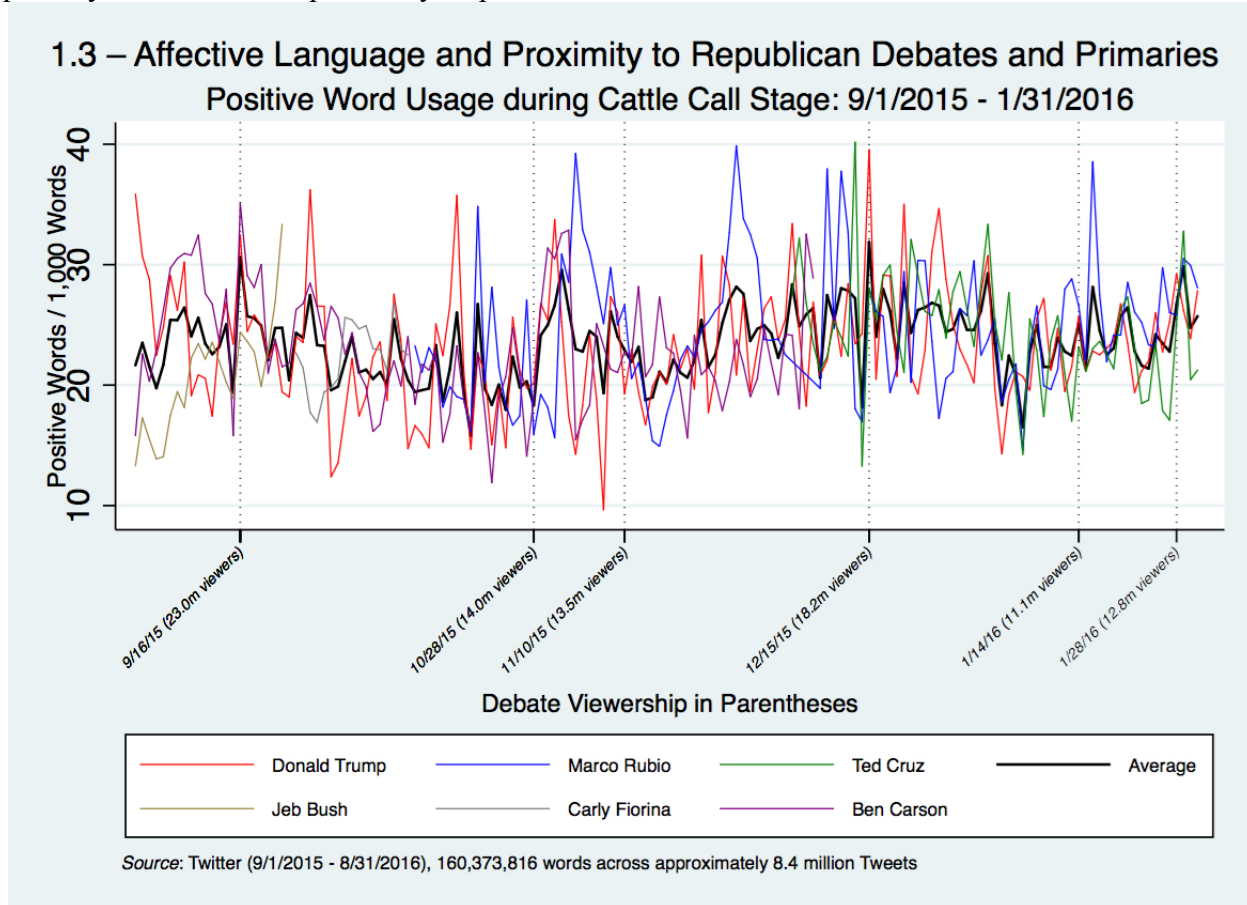Graph 1.3: Comparing frequency of positive affect in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries
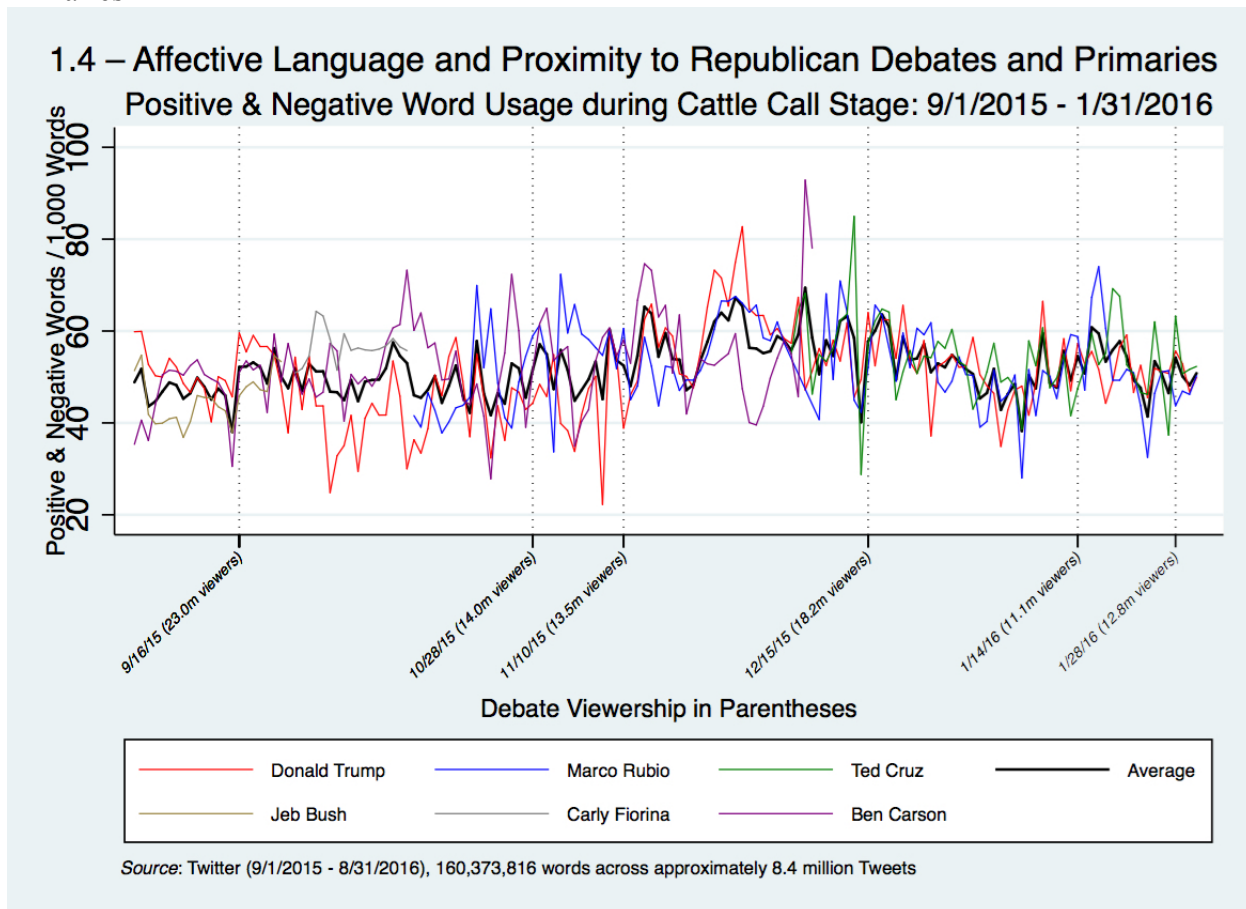

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Positive words in tweets identified using Lexicon 2.0 content analysis software.
Frequency is measured in number of occurrences per 1,000 words.

Graph 1.4: Comparing frequency of combined positive and negative affect in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries
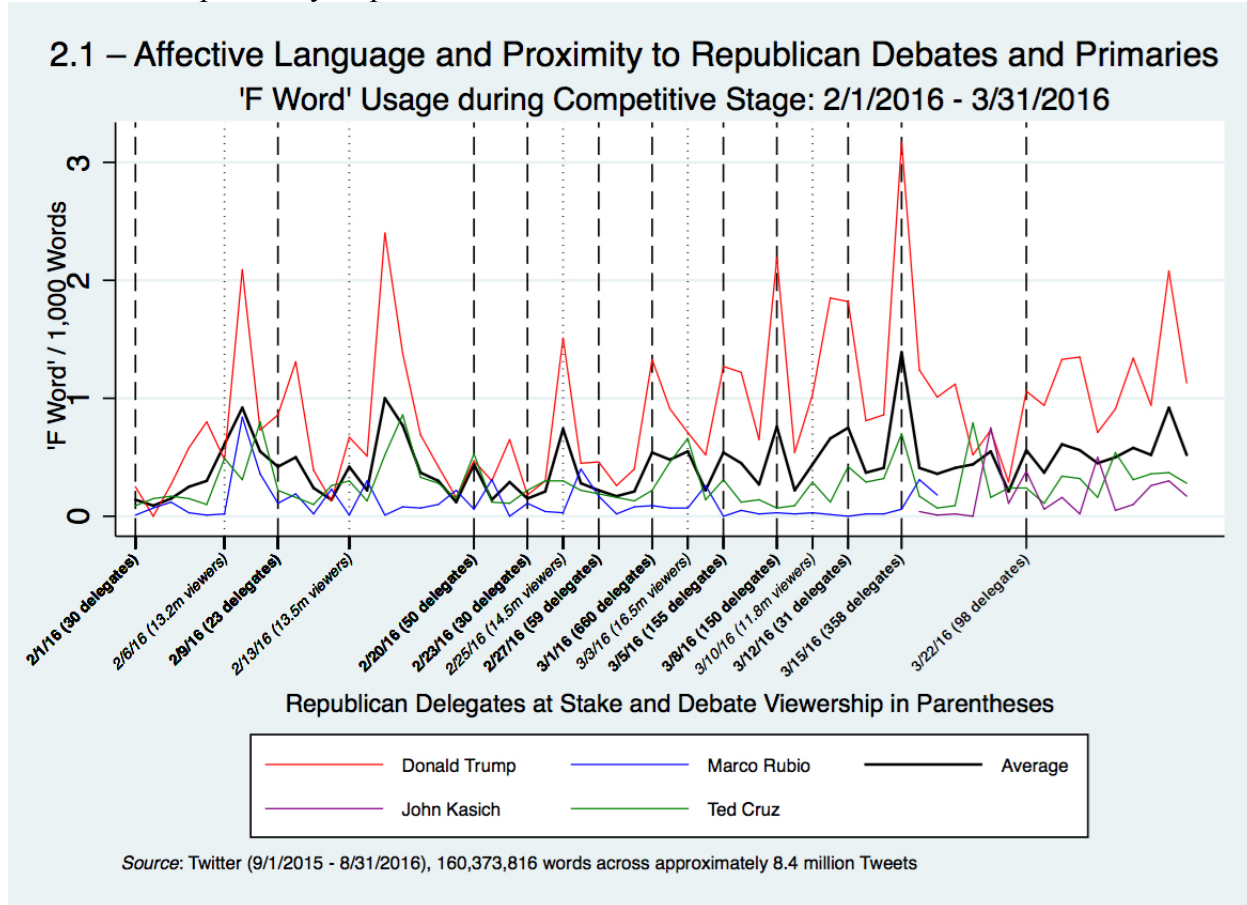

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Positive and negative words in tweets identified using Lexicon 2.0 content analysis software. Frequency is measured in number of occurrences per 1,000 words.

Graph 2.1: Comparing frequency of "F words" in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries

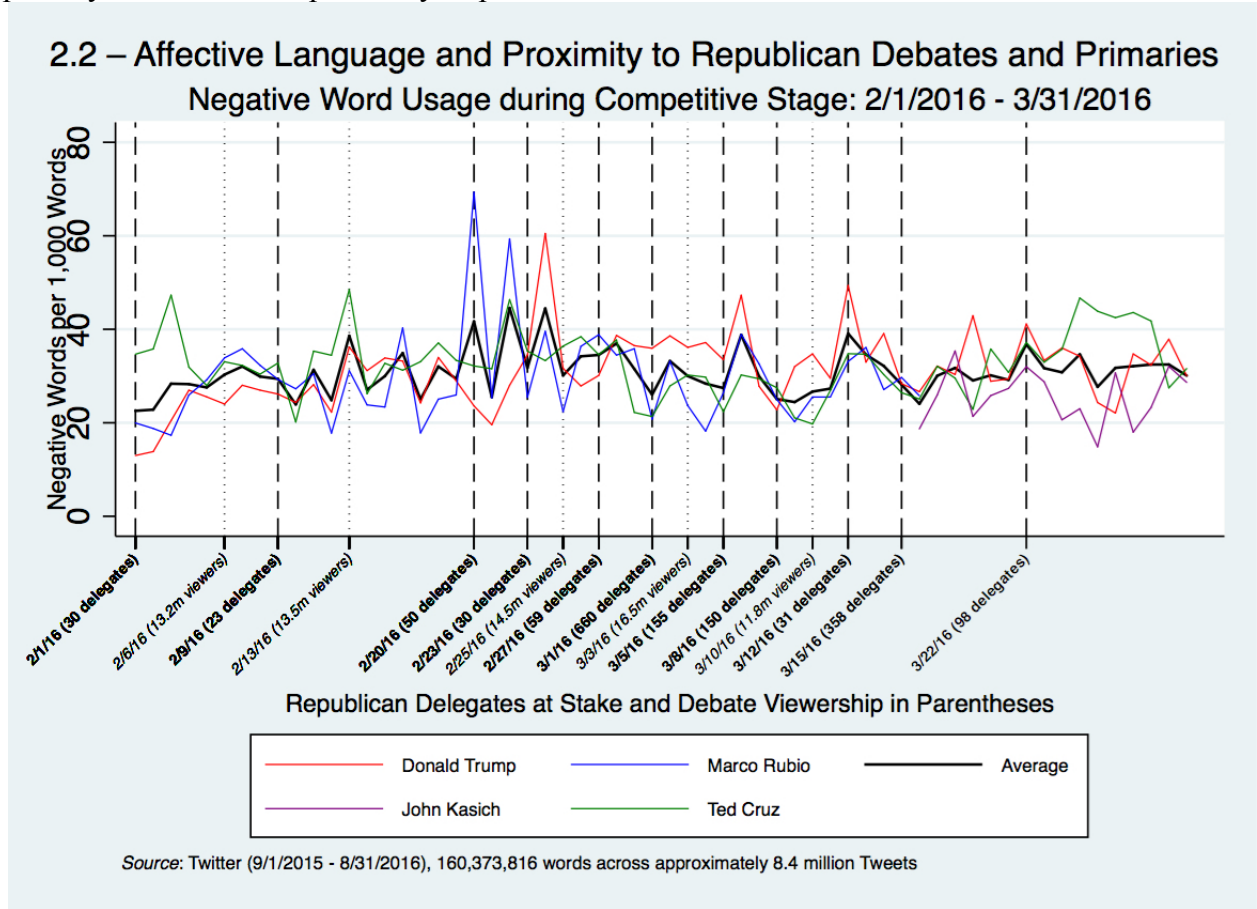

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Frequency of "F words" in tweets identified using Lexicon 2.0 content analysis software.
Frequency is measured in number of occurrences per 1,000 words.

Graph 2.2: Comparing frequency of negative affect in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries

2.2 – Affective Language and Proximity to Republican Debates and Primaries
Negative Word Usage during Competitive Stage: 2/1/2016 - 3/31/2016

Source: Twitter (9/1/2015 - 8/31/2016), 160,373,816 words across approximately 8.4 million Tweets

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Negative words in tweets identified using Lexicon 2.0 content analysis software.
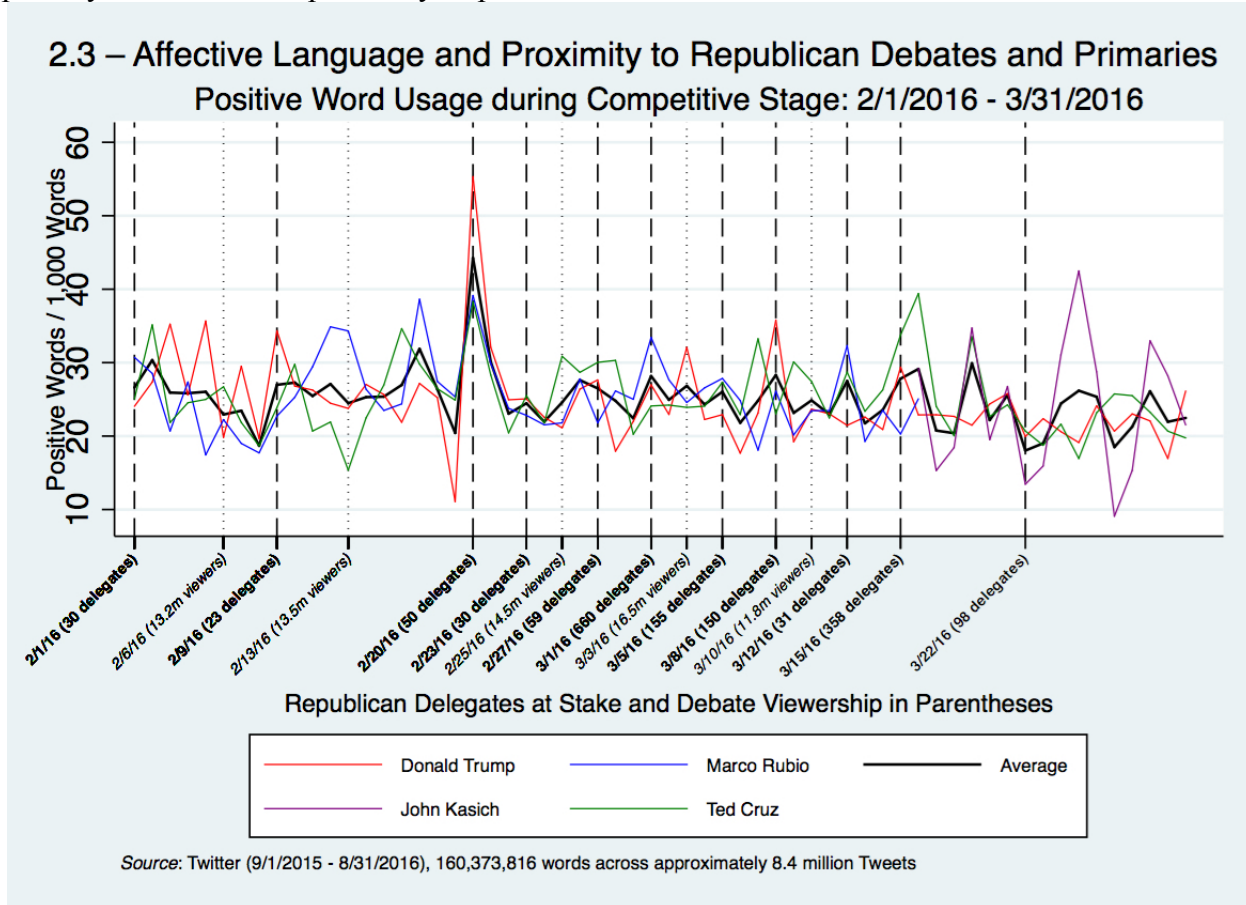Frequency is measured in number of occurrences per 1,000 words.

Graph 2.3: Comparing frequency of positive affect in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries
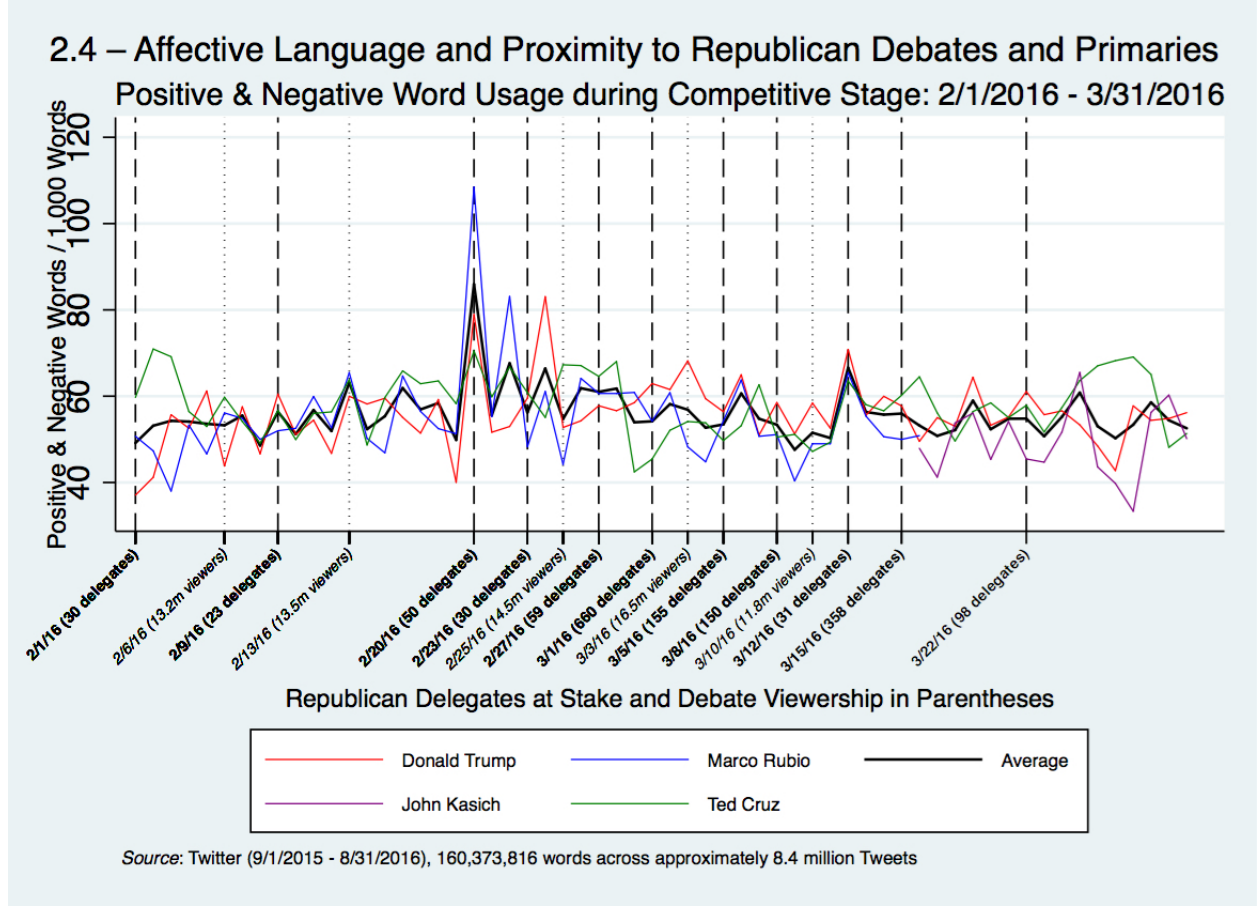

2.3 – Affective Language and Proximity to Republican Debates and Primaries
Positive Word Usage during Competitive Stage: 2/1/2016 - 3/31/2016

Positive Words / 1,000 Words

Republican Delegates at Stake and Debate Viewership in Parentheses

- Donald Trump
- Marco Rubio
- Average
- John Kasich
- Ted Cruz

Source: Twitter (9/1/2015 - 8/31/2016), 160,373,816 words across approximately 8.4 million Tweets

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Frequency of positive words in tweets identified using Lexicon 2.0 content analysis software. Frequency is measured in number of occurrences per 1,000 words.

Graph 2.4: Comparing frequency of combined positive and negative affect in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries
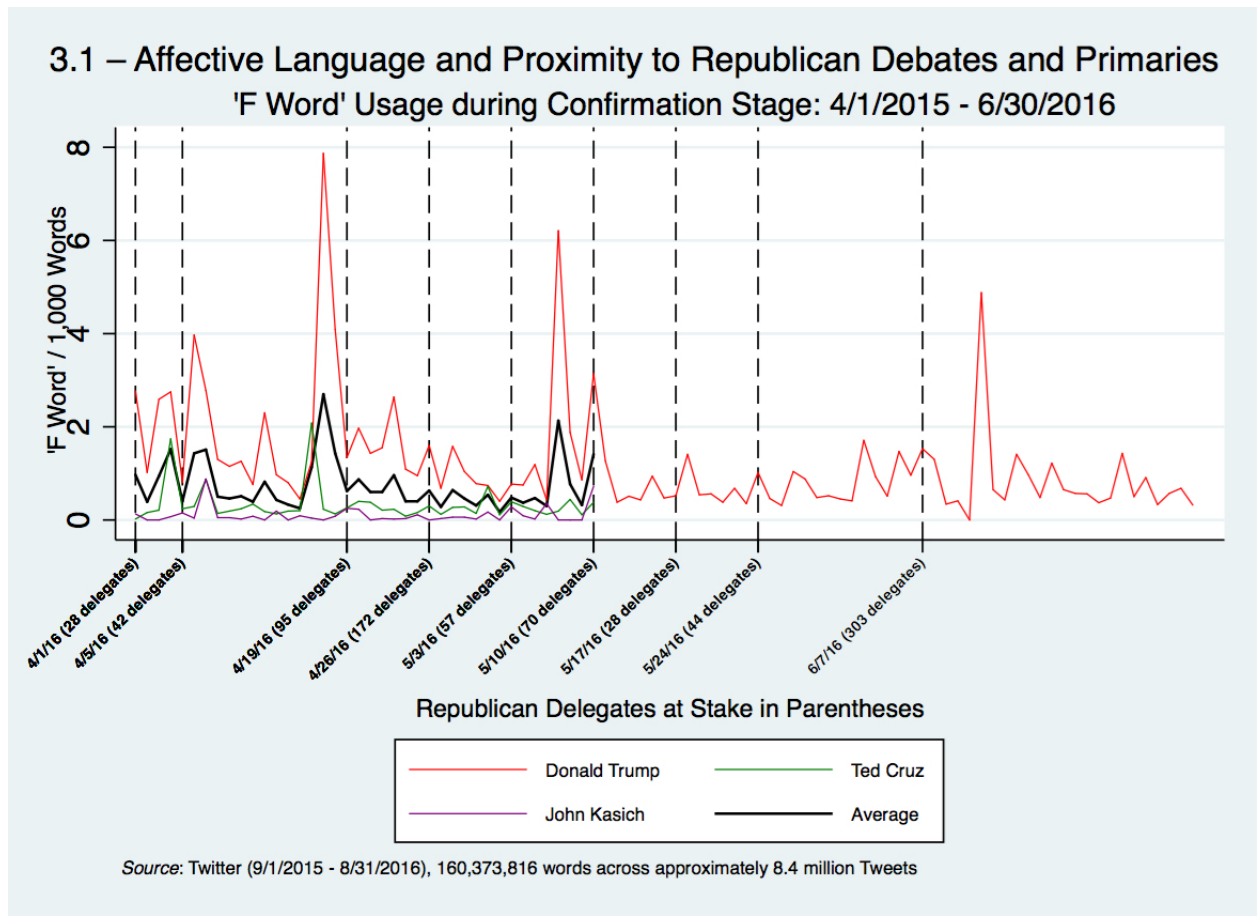

2.4 – Affective Language and Proximity to Republican Debates and Primaries
Positive & Negative Word Usage during Competitive Stage: 2/1/2016 - 3/31/2016

Source: Twitter (9/1/2015 - 8/31/2016), 160,373,816 words across approximately 8.4 million Tweets

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Frequency of positive and negative words in tweets identified using Lexicon 2.0 content analysis software. Frequency is measured in number of occurrences per 1,000 words.
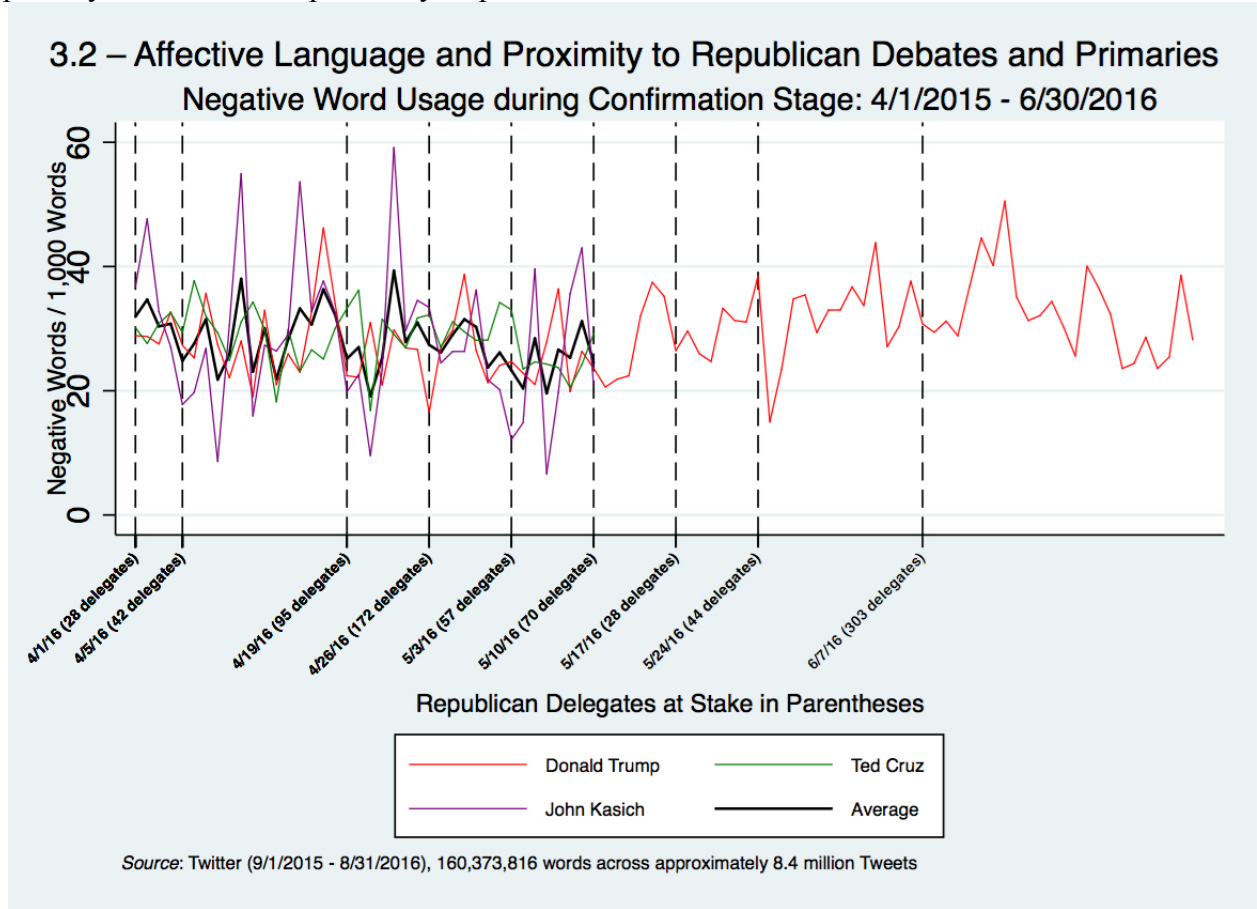
Graph 3.1: Comparing frequency of "F words" in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries

3.1 – Affective Language and Proximity to Republican Debates and Primaries
'F Word' Usage during Confirmation Stage: 4/1/2015 - 6/30/2016

'F Word' / 1,000 Words

Republican Delegates at Stake in Parentheses

Donald Trump   Ted Cruz
John Kasich   Average

Source: Twitter (9/1/2015 - 8/31/2016), 160,373,816 words across approximately 8.4 million Tweets

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Frequency of "F words" in tweets identified using Lexicon 2.0 content analysis software. Frequency is measured in number of occurrences per 1,000 words.
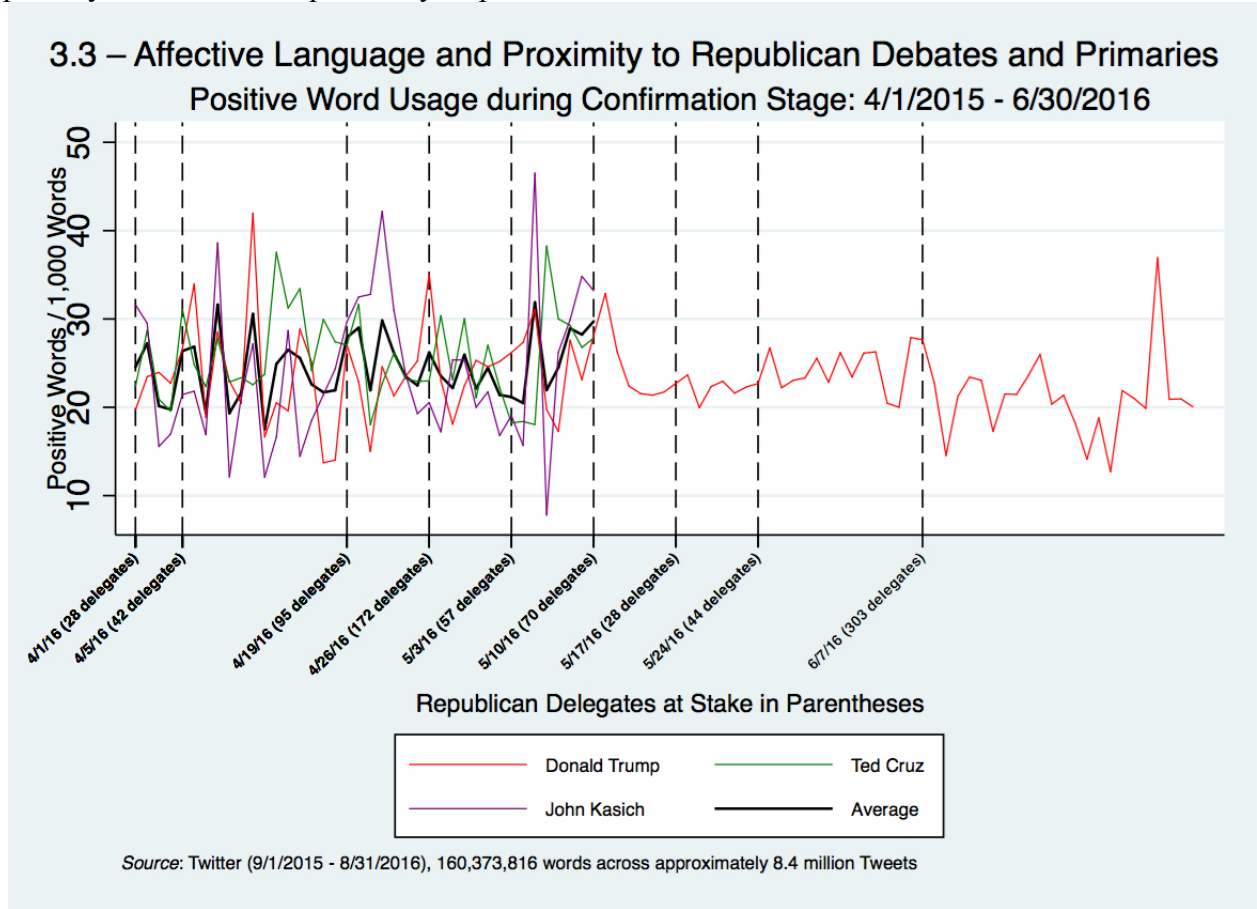
Graph 3.2: Comparing frequency of negative affect in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries

3.2 – Affective Language and Proximity to Republican Debates and Primaries
Negative Word Usage during Confirmation Stage: 4/1/2015 - 6/30/2016

Source: Twitter (9/1/2015 - 8/31/2016), 160,373,816 words across approximately 8.4 million Tweets

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Frequency of negative words in tweets identified using Lexicon 2.0 content analysis software. Frequency is measured in number of occurrences per 1,000 words.

Graph 3.3: Comparing frequency of positive affect in tweets[1] mentioning Republican
primary candidates and proximity Republican Debates and Primaries

3.3 – Affective Language and Proximity to Republican Debates and Primaries
Positive Word Usage during Confirmation Stage: 4/1/2015 - 6/30/2016

Source: Twitter (9/1/2015 - 8/31/2016), 160,373,816 words across approximately 8.4 million Tweets

Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Frequency of positive words in tweets identified using Lexicon 2.0 content analysis
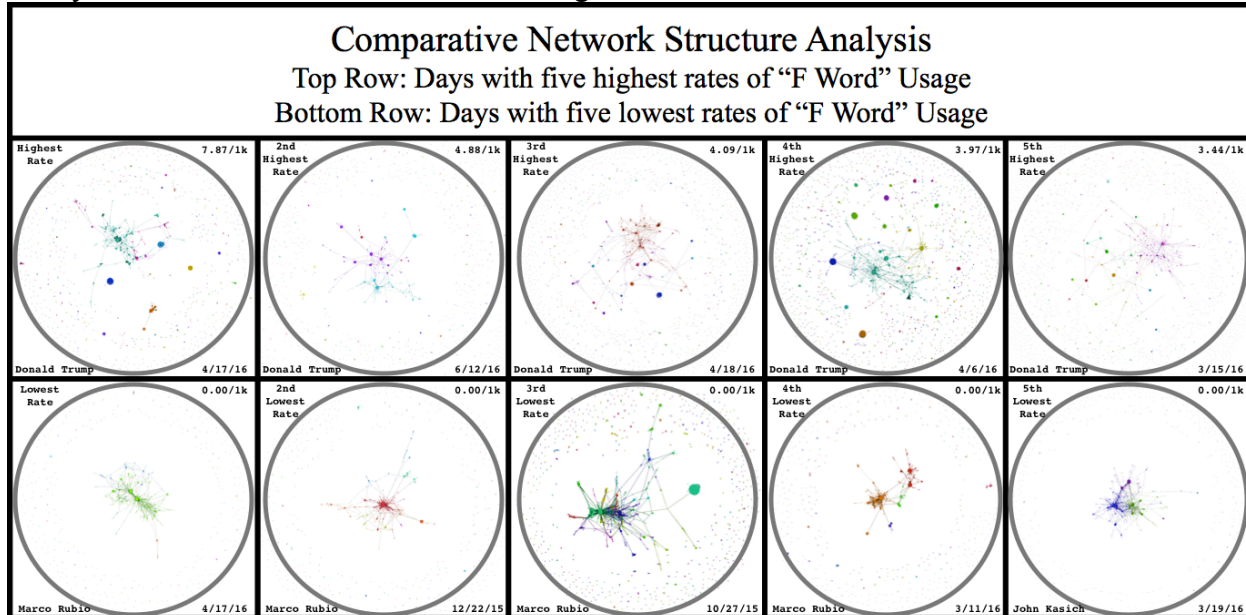software. Frequency is measured in number of occurrences per 1,000 words.

Graph 3.4: Comparing frequency of combined positive and negative affect in tweets[1] mentioning Republican primary candidates and proximity Republican Debates and Primaries
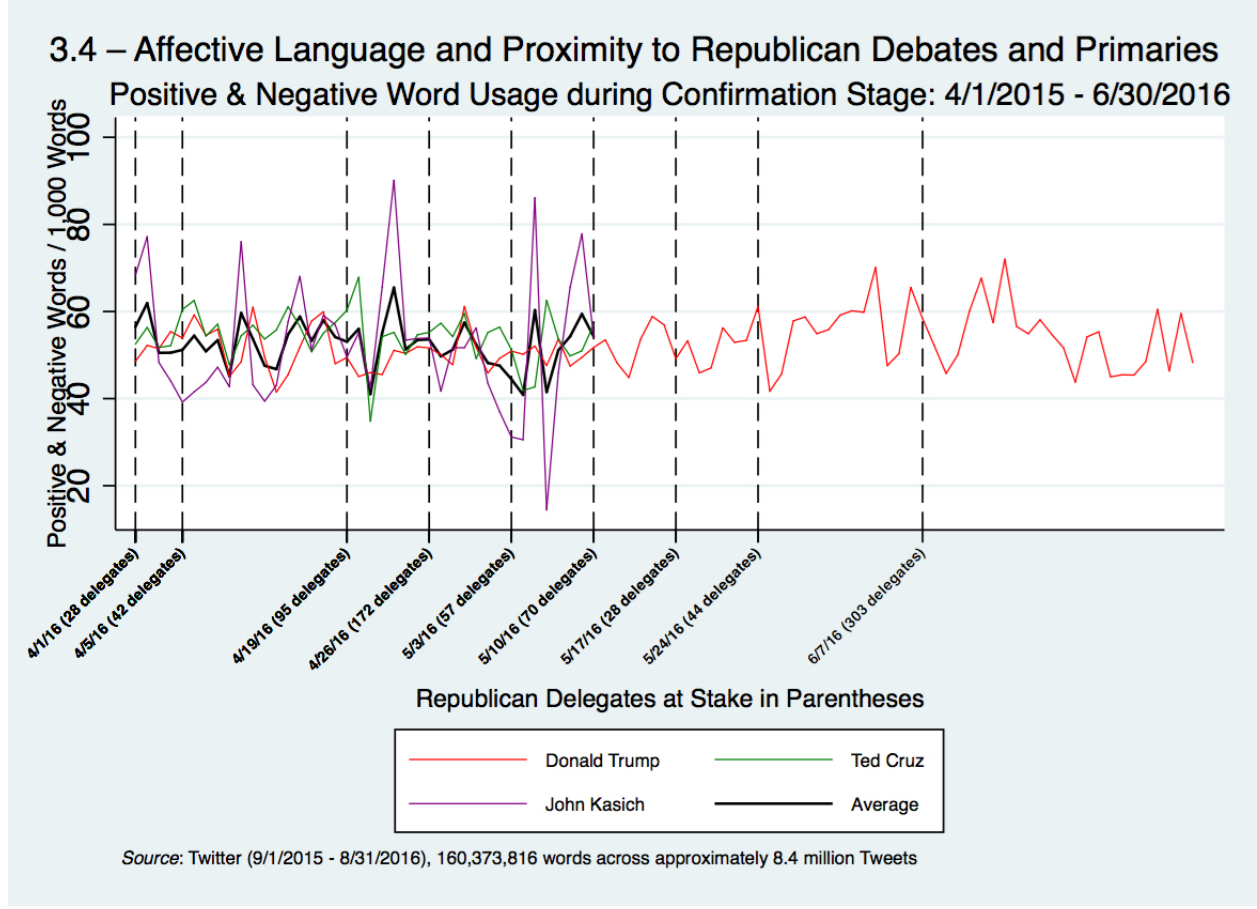
Source: Twitter (9/1/2015 – 8/31/2016)
160,373,816 words and 8.4 million tweets

Notes:
[1]Frequency of positive and negative in tweets identified using Lexicon 2.0 content analysis software. Frequency is measured in number of occurrences per 1,000 words.

Visualization 1.1: Comparing Network Structure[1] in Days with Highest Rate of "F Word" to Days with Lowest Rates of "F Word" Usage

Comparative Network Structure Analysis
Top Row: Days with five highest rates of "F Word" Usage
Bottom Row: Days with five lowest rates of "F Word" Usage

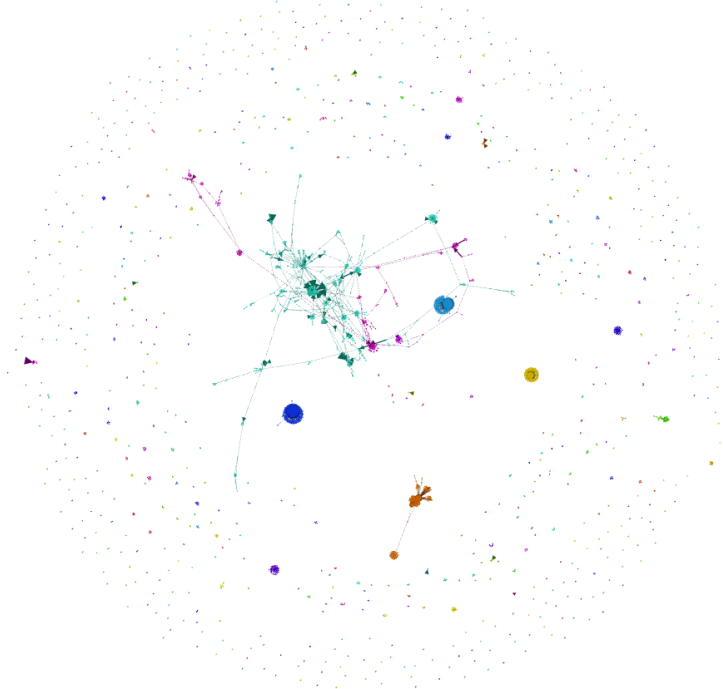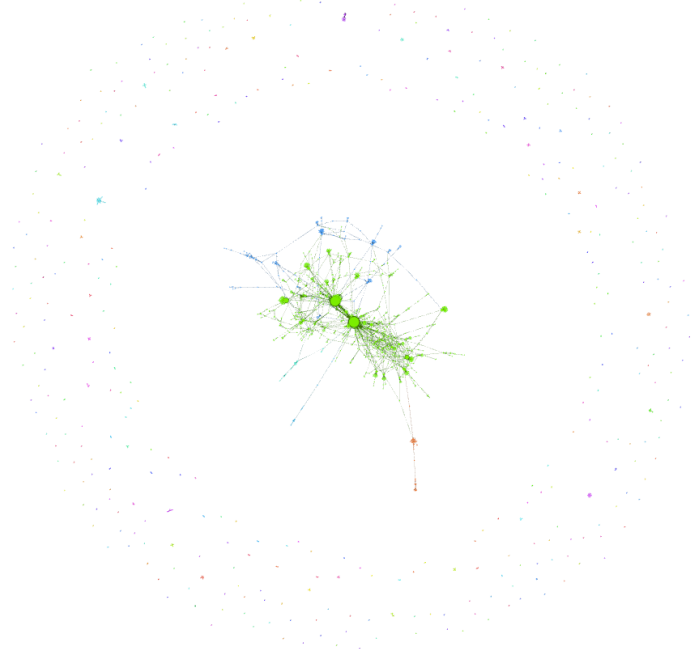*Source:* Twitter (9/1/2015 – 8/31/2016)
[1]Visualizations created using Gephi 0.8.1 for Macintosh

Visualization 2.1: Comparing Network Structure[1] in Day with Highest Rate of "F Word" to Day with Lowest Rate of "F Word" Usage

Highest Rate: Donald Trump (4/17/2016) – 7.87 per 1,000 words

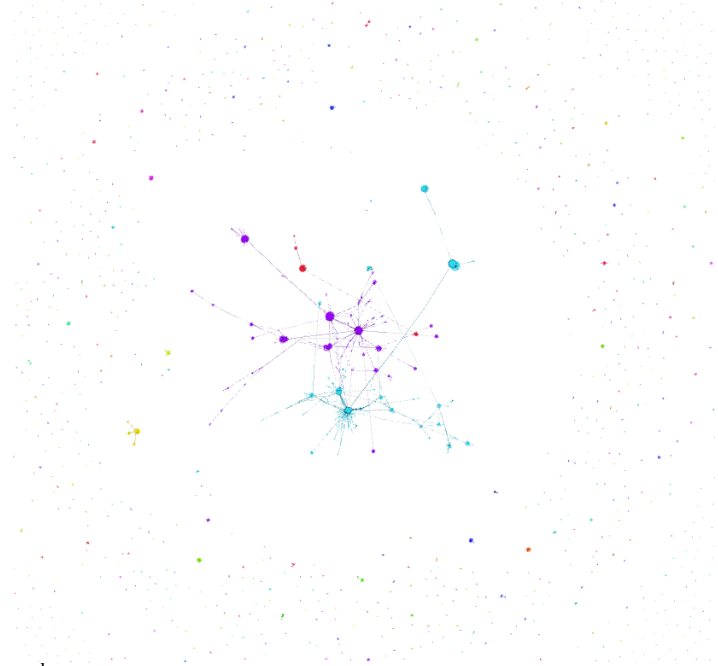Lowest Rate: Marco Rubio (4/17/2016) – 0.00 per 1,000 words

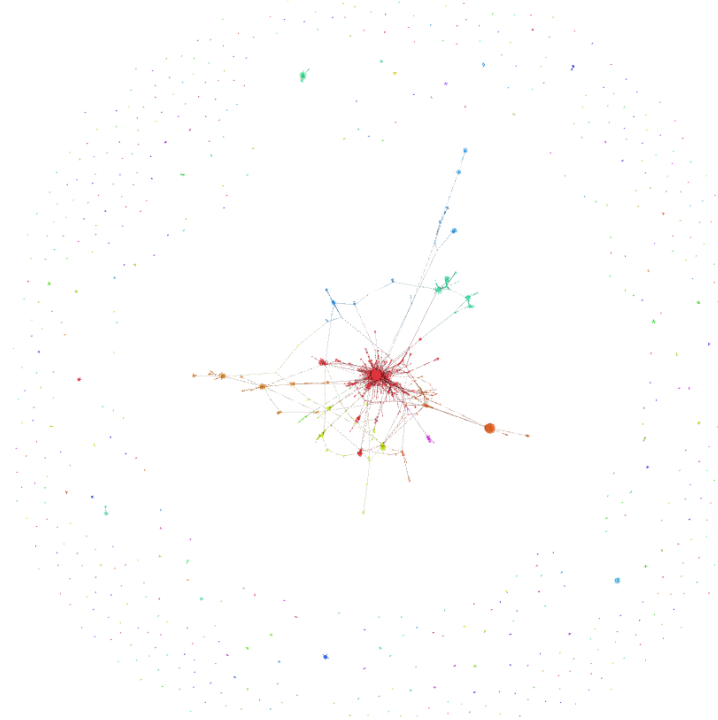*Source:* Twitter (9/1/2015 – 8/31/2016)
[1]Visualizations created using Gephi 0.8.1 for Macintosh

Visualization 2.2: Comparing Network Structure[1] in Day with 2nd Highest Rate of "F Word" to Day with 2nd Lowest Rate of "F Word" Usage

2nd Highest Rate: Donald Trump (6/12/2016) – 4.88 per 1,000 words

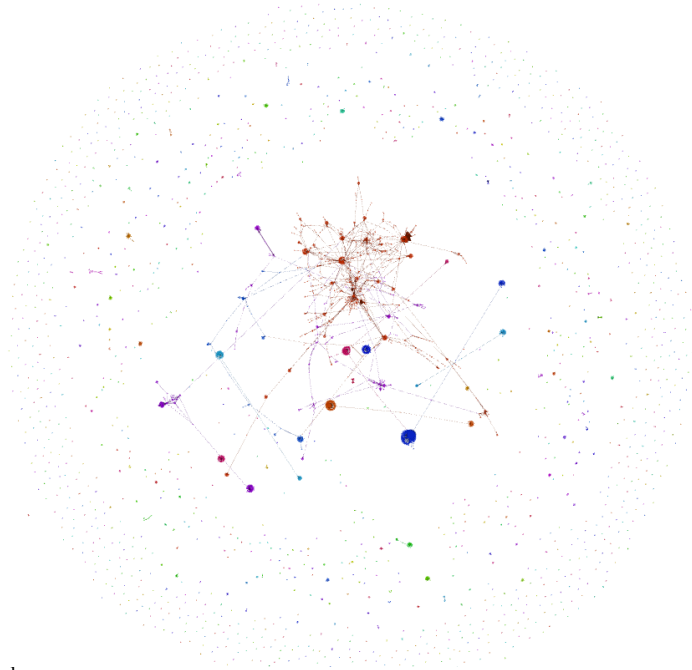2nd Lowest Rate: Marco Rubio (12/22/2015) – 0.00 per 1,000 words

Source: Twitter (9/1/2015 – 8/31/2016)
[1]Visualizations created using Gephi 0.8.1 for Macintosh

Visualization 2.3: Comparing Network Structure[1] in Day with 3rd Highest Rate of "F Word" to Day with 3rd Lowest Rate of "F Word" Usage

3rd Highest Rate: Donald Trump (4/18/2016) – 4.09 per 1,000 words

3rd Lowest Rate: Marco Rubio (10/27/2015) – 0.00 per 1,000 words

Source: Twitter (9/1/2015 – 8/31/2016)
[1]Visualizations created using Gephi 0.8.1 for Macintosh

Visualization 2.4: Comparing Network Structure[1] in Day with 4th Highest Rate of "F Word" to Day with 4th Lowest Rate of "F Word" Usage

4th Highest Rate: Donald Trump (4/6/2016) – 3.97 per 1,000 words

4th Lowest Rate: Marco Rubio (3/11/2016) – 0.00 per 1,000 words
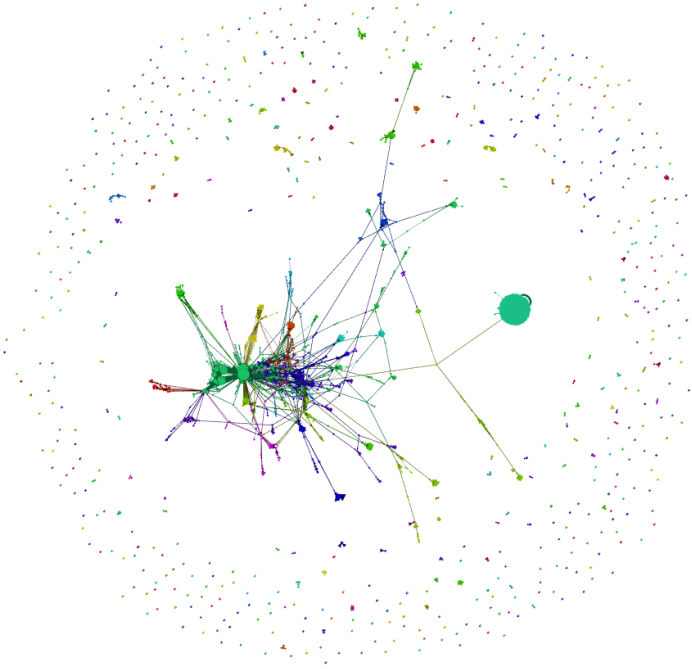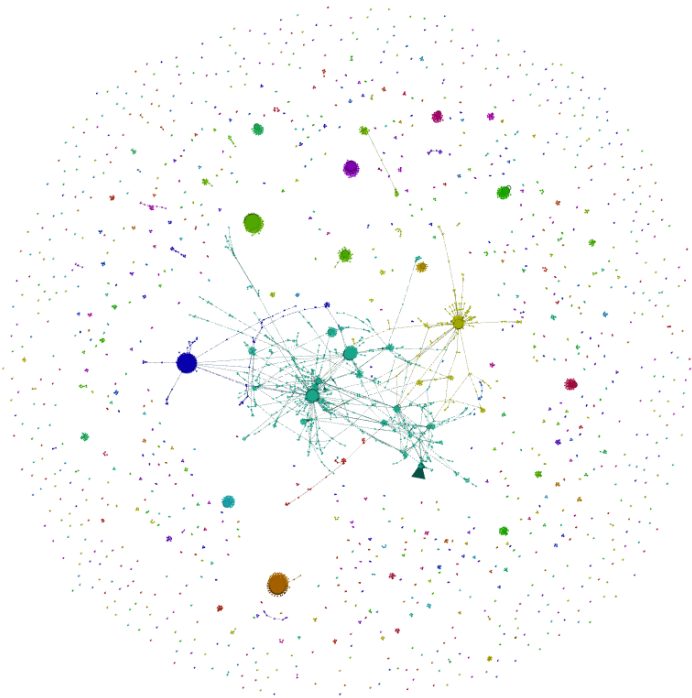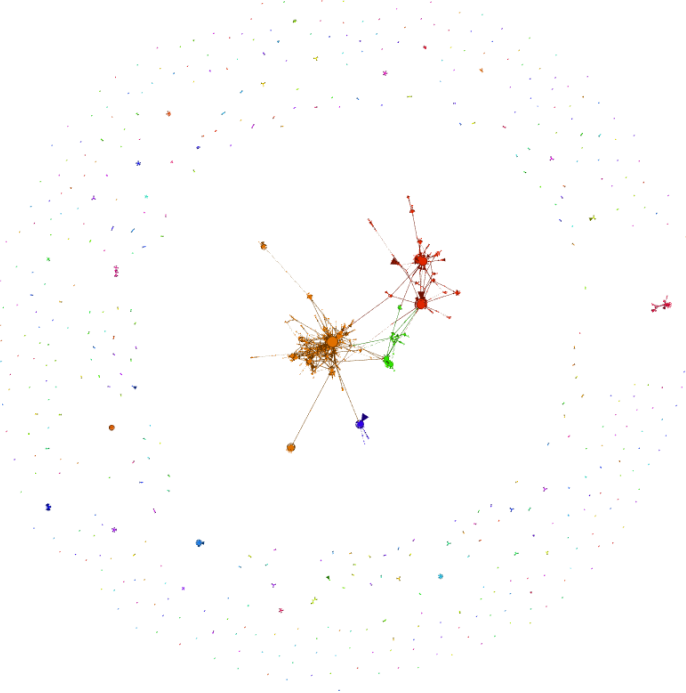

Source: Twitter (9/1/2015 – 8/31/2016)
[1]Visualizations created using Gephi 0.8.1 for Macintosh

Visualization 2.5: Comparing Network Structure[1] in Day with 5th Highest Rate of "F Word" to Day with 5th Lowest Rate of "F Word" Usage

5th Highest Rate: Donald Trump (3/15/2016) – 3.44 per 1,000 words

5th Lowest Rate: John Kasich (3/19/2016) – 0.00 per 1,000 words
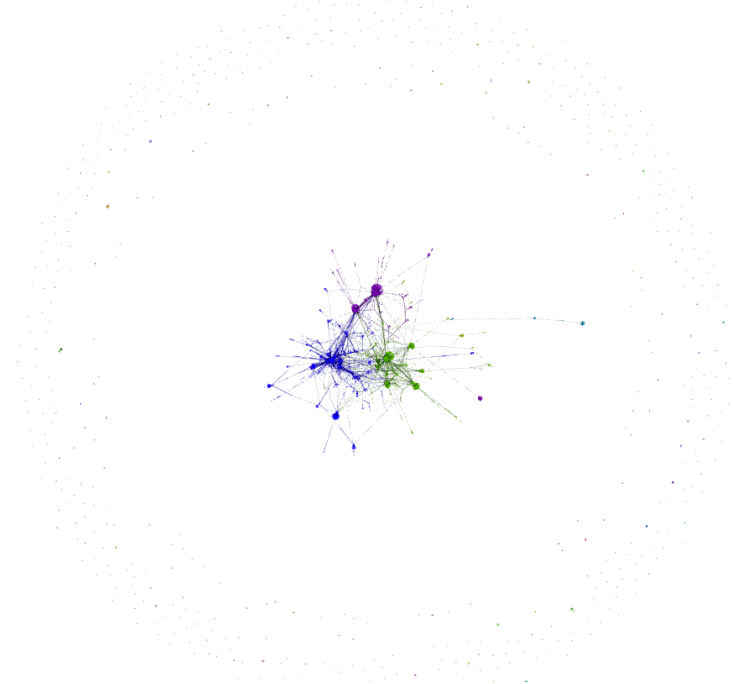
Source: Twitter (9/1/2015 – 8/31/2016)
[1]Visualizations created using Gephi 0.8.1 for Macintosh

Visualization 3.1: Analyzing Network Centrality and Influence[1] of Extreme Affective
Rhetoric on Day with Highest Rate of "F Word" Usage

Donald Trump (4/17/2016) – 7.87 per 1,000 words

Source: Twitter (9/1/2015 – 8/31/2016)
[1]Visualization created using Gephi 0.8.1 for Macintosh

Visualization 3.2: Analyzing Network Centrality and Influence[1] of Extreme Affective
Rhetoric on Day with 2nd Highest Rate of "F Word" Usage

Donald Trump (6/12/2016) – 4.88 per 1,000 words
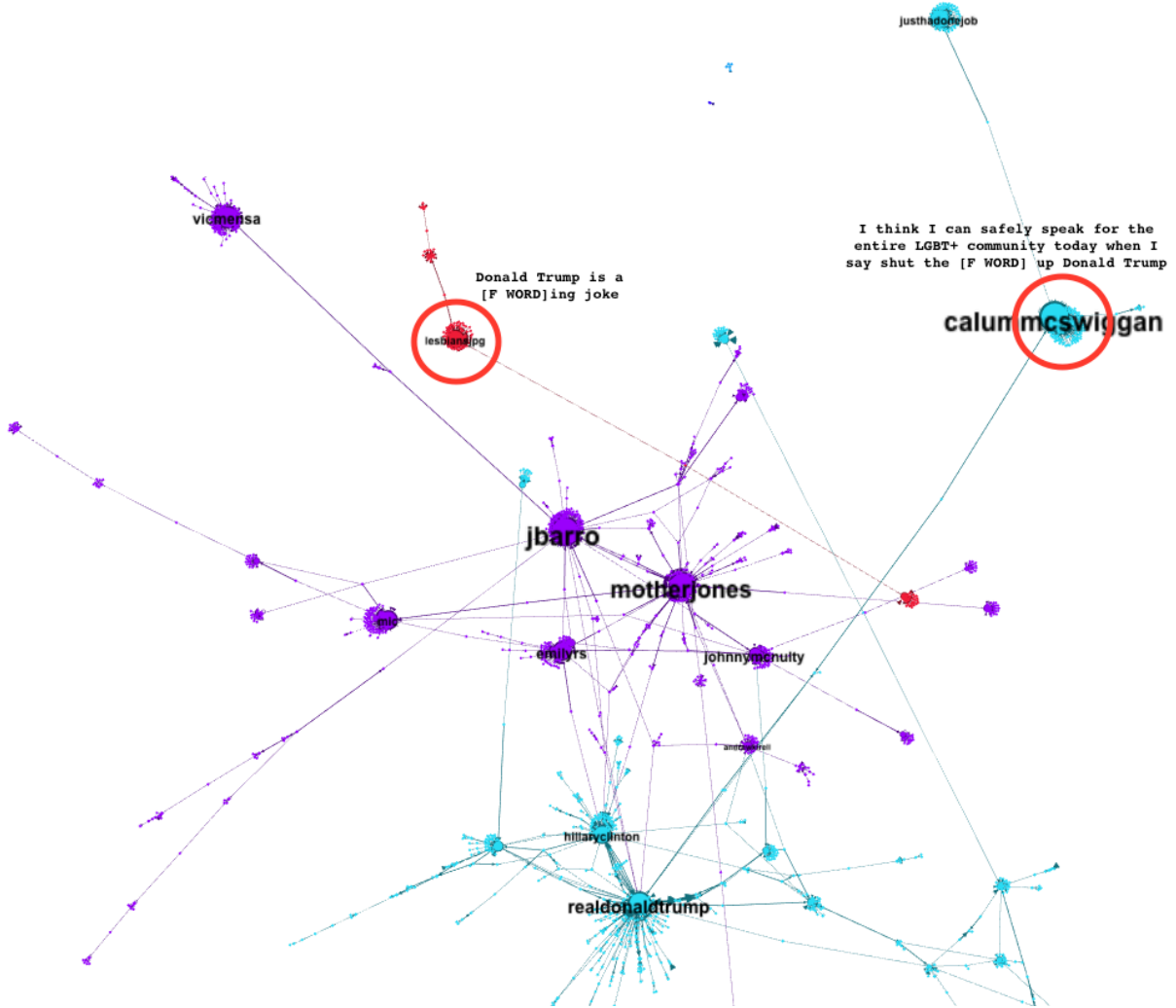
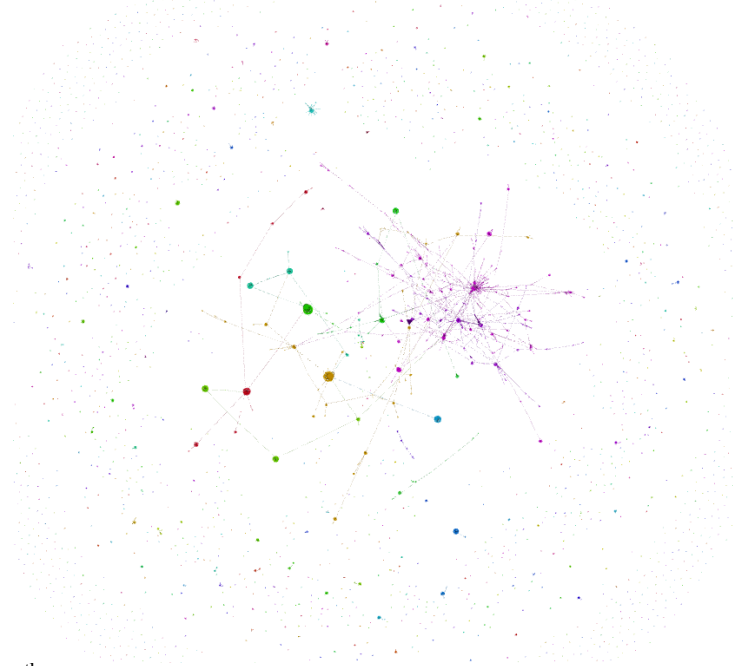*Source:* Twitter (9/1/2015 – 8/31/2016)
[1]Visualization created using Gephi 0.8.1 for Macintosh

Visualization 3.3: Analyzing Network Centrality and Influence[1] of Extreme Affective
Rhetoric on Day with 3rd Highest Rate of "F Word" Usage

Donald Trump (4/18/2016) – 4.09 per 1,000 words
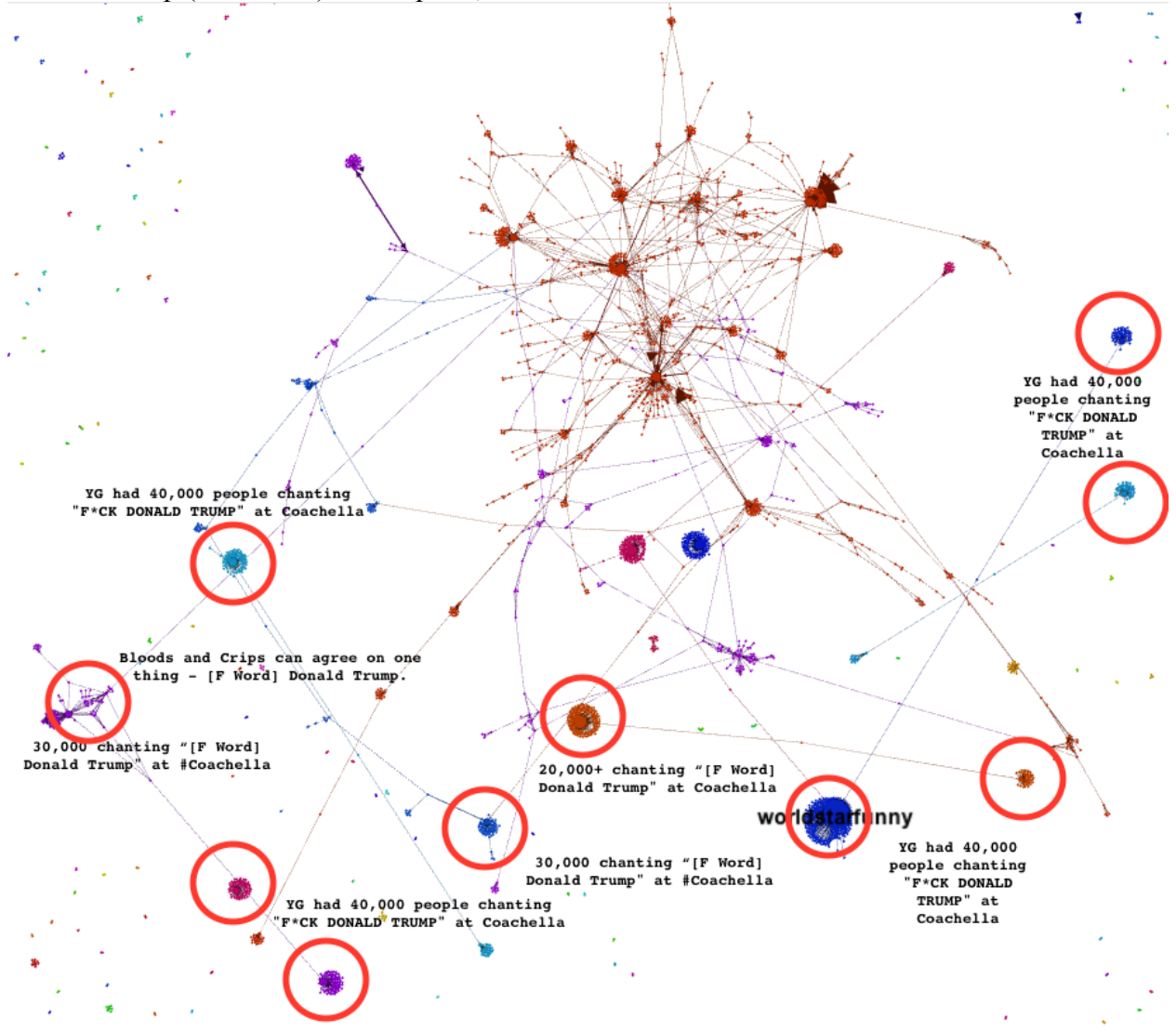
Source: Twitter (9/1/2015 – 8/31/2016)
[1]Visualization created using Gephi 0.8.1 for Macintosh

Visualization 3.4: Analyzing Network Centrality and Influence[1] of Extreme Affective Rhetoric on Day with 4th Highest Rate of "F Word" Usage

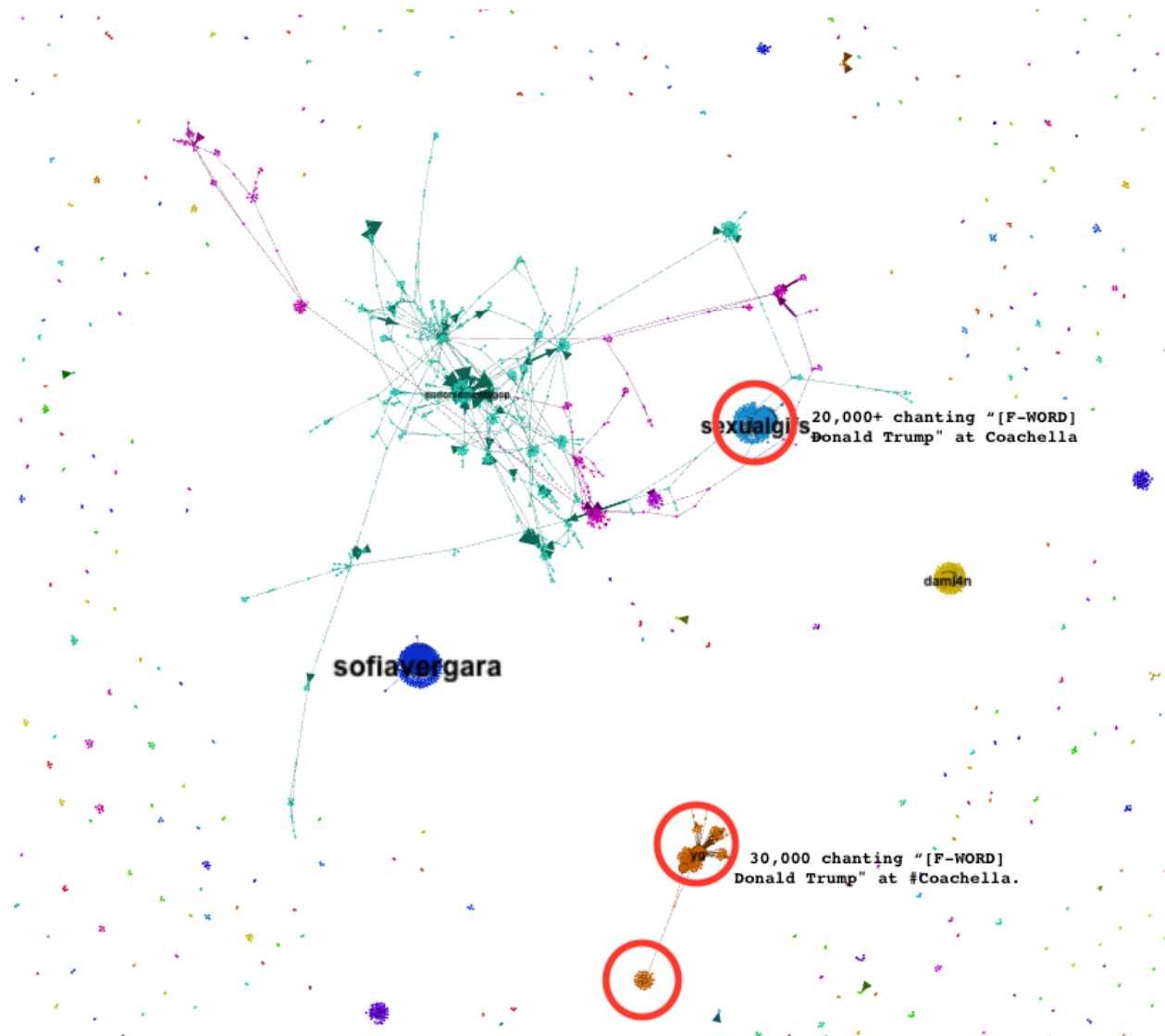Donald Trump (4/6/2016) – 3.97 per 1,000 words
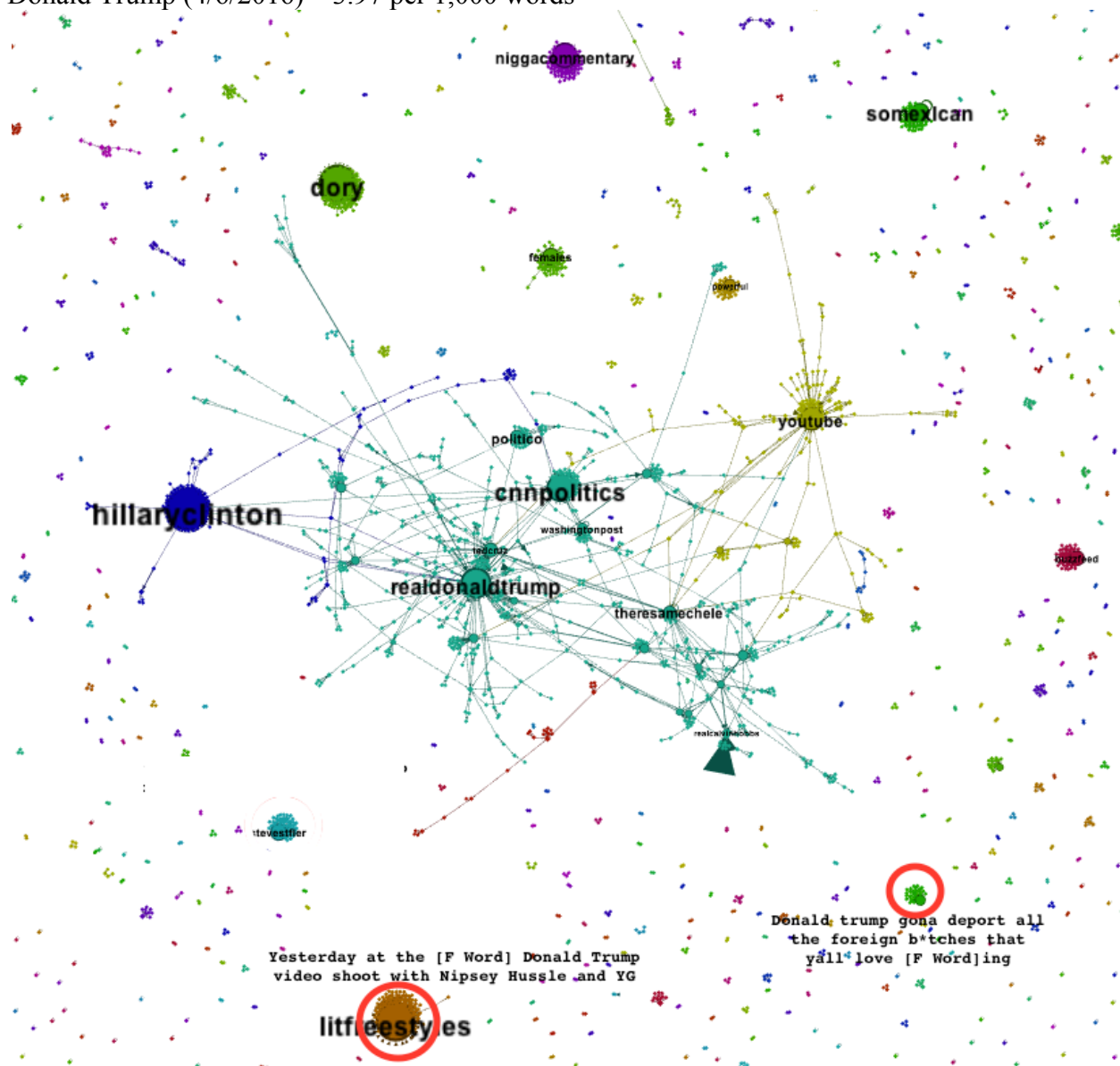
Source: Twitter (9/1/2015 – 8/31/2016)
[1]Visualization created using Gephi 0.8.1 for Macintosh

Visualization 3.5: Analyzing Network Centrality and Influence[1] of Extreme Affective Rhetoric on Day with 5th Highest Rate of "F Word" Usage

Donald Trump (3/15/2016) – 3.44 per 1,000 words
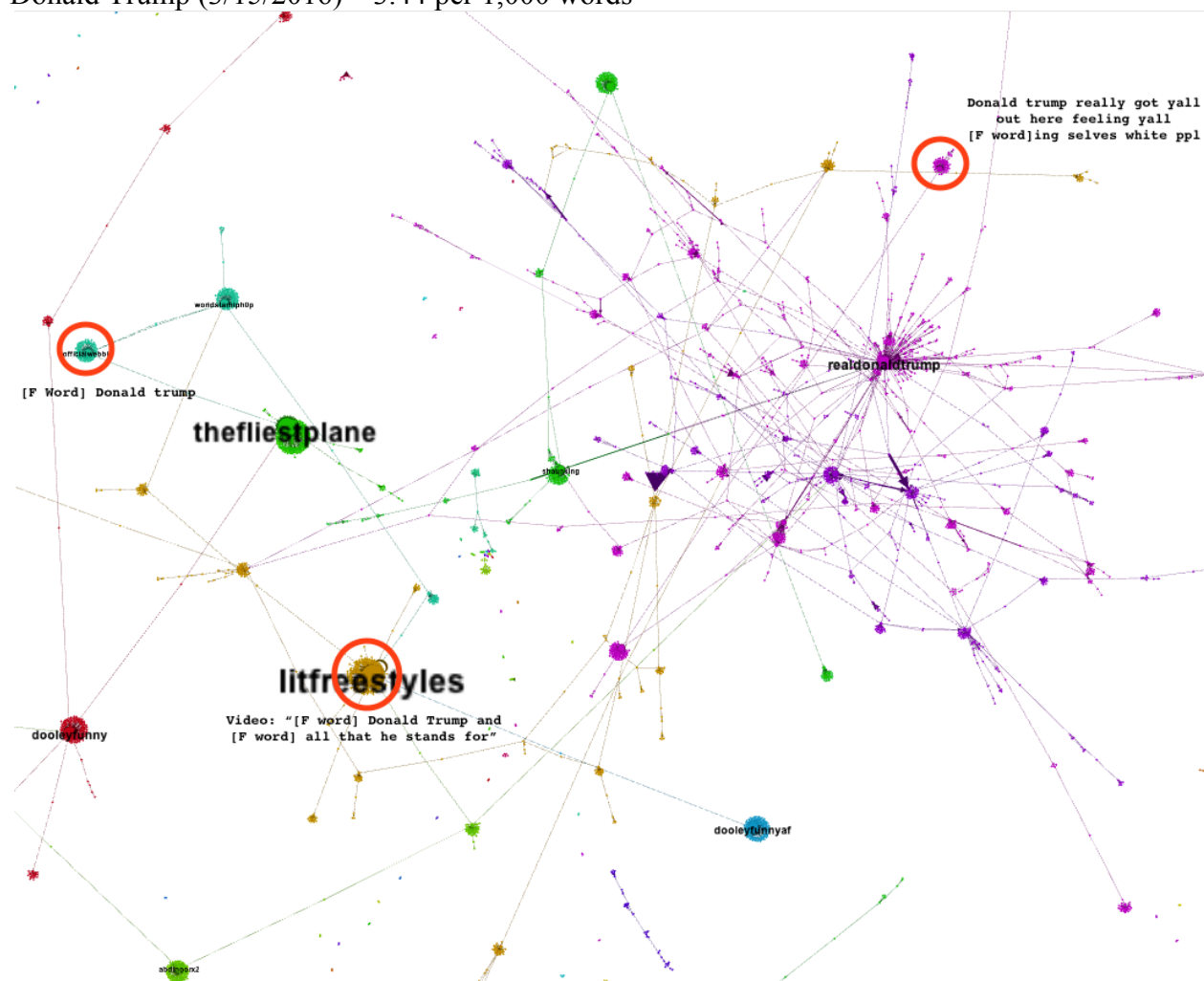
Source: Twitter (9/1/2015 – 8/31/2016)
[1]Visualization created using Gephi 0.8.1 for Macintosh

Dissertation Chapter 5

The Instant Effects of Confrontation, Controversy, and Contempt:

A live-stream analysis of mass polarization on Twitter during the 2016 presidential

debates

Eric C. Vorst

Introduction

The 2016 U.S. presidential campaign was marked by a daily onslaught of disputes between Hillary Clinton, Donald Trump, and their surrogates. Whether these confrontations involved e-mails, tax returns, the Russian government, reality television, temperament, white supremacist groups, health issues, a history of misogynistic behavior, or one of a number of other seemingly endless controversies, they all seemed to share a common tone in their discourse: contempt. This atmosphere of relentless confrontation, controversy, and contempt took center stage during each of the three presidential debates and in front of record-breaking television audiences, leading many to ask whether such visible divisiveness on the part of political elites may have contributed to mass polarization.

This chapter measures the real-time effects of elite polarization on social networks by drawing upon an original data set of approximately 1,500,000 tweets captured live during the first 2016 U.S. presidential debate. I measure shifts in sentiment, network dynamics, and neighborhood structure during the one-minute time frame immediately following especially controversial or confrontational candidate statements and exchanges. The nature of these causal relationships offers unique insight into the influence polarizing messages by elites may have on mass polarization, especially when observed in a live and extremely politically charged atmosphere.

Instant Effects during High Stakes Political Events

Recent research in the field of political communication has focused on the phenomenon of hybrid media events, wherein social media is used alongside and during

televised political events. Referred to as "dual screening", the viewing public expands its

role from being a mostly passive consumer of information to being an active member of

the political event itself. These roles often involve members of the general public using

social media to engage in "lay tutelage" behavior, including the acts of fact-checking,

correcting, counter-claiming, or highlighting biased reporting (Vaccari, Chadwick, &

O'Loughlin 2015). Within this environment, the power of event participants to shape the

narrative increases in proportion to their relative influence when measured by their

number of active followers. In such a scenario, these event participants may serve as

"bridging elites" with an influential power that can rival that of media and political elites

(Freelon & Karpf 2015).

Findings such as these have been quite valuable in better describing and

explaining the roles individuals play in politically charged social media environments, as

well as the types of individuals who play these roles. At the same time, more needs to be

known regarding how elite messages originating in the traditional media environment

shapes the network structures that facilitate social media participants' roles, whether they

be participating in a bridging elite, lay tutelage, or passive observer function. Further, a

better understanding of the nature of how these network structures shift can provide

valuable context for interpreting the relative influence of participants regardless of their

roles.

Research Design

This chapter examines the extent to which polarizing behavior on the part of elites

within a live and politically charged environment influences online spaces in a way that

creates conditions where mass polarization may develop. The analysis in this chapter is unique, as it offers a perspective that is both immediate and unfiltered. Further, it takes place in an environment where the variable of participant self-selection has less influence than would be the case in a traditional media setting. First, individuals' choice in their media source for the debate has virtually no impact on the way in which the political information is being received. Rather put, the content of elite cues is delivered in an unbiased and direct manner, regardless of whether an individual watches the debate unfold on Fox News, MSNBC, CNN, or some other channel. An individual cannot choose to receive these cues from a source that aligns with their ideology, as competing views are part and parcel of a televised debate. Second, individuals' reactions to elite cues are being observed in an unfiltered and ideological neutral environment. Specifically, reactions are only observed for individuals who have shared their messages using the official Twitter hashtag for the 2016 presidential debates.

Research Questions

In order to test for a potential link between elite polarization and mass polarization in these unique dual environments, this chapter investigates two sequentially dependent research questions. Specifically, the first research question looks for the effects of elite cues in online spaces, while the second research question measures how the unique nature of social networks impact the extent to which any such effects matter.

Research Question #1: How do polarizing messages by elites influence the nature of political discourse in online spaces?

The first section of this chapter seeks to determine if there are any relationships between expressions of polarizing behavior on the part of political elites and a resulting reaction among individuals that could indicate the seeds for mass polarization are being sown. Using the first presidential debate as my testing grounds, I identify moments during the debate where either of the two candidates made statements that could be perceived as confrontational, controversial, or contemptuous. These moments are used as proxies for the independent variable of elite polarization. Levels of affective rhetoric and abstract thought in individuals' tweets during each minute of the debate are used to measure the extent to which the preconditions for mass polarization are developing or diminishing at different moments during the debate. Finally, a time series analysis is utilized to determine if moments of confrontation, controversy, or contempt on the part of the two candidates result in a shift in frequency of affective rhetoric or abstract thought on the part of individuals watching the debate live. In essence, the first section seeks to answer the question: "Does it happen?"

Research Question #2:

How does the nature of political discourse in online spaces influence the creation of polarized communities?

While the first section of this chapter looks for causal links between polarizing messages by elites and shifts in mass affective rhetoric, the second section of this chapter seeks to determine whether there is a relationship between shifts in mass affective

rhetoric and the formation of polarized communities of discussion. Drawing upon periods with high rates of affective rhetoric in tweets, I conduct network metrics analyses to measure the size of communities of discussion and the relative efficiency of the networks within which these communities form. These metrics provide evidence as to the relative levels of polarization within a given network at a given time. If evidence of polarized communities exist, this would suggest mechanisms exist which would facilitate a causal link between elite polarization in the form of elite cues and mass polarization in social media. Further, this approach may provide even further clarity regarding the extent to which different *types* of elite cues lead to different *types* of mass polarization in social media. In essence, the second section seeks to answer the question: "Does it matter?"

Hypotheses

Hypothesis #1: Moments of confrontation, controversy, and contempt cause an instant increase in the rates of affective rhetoric in viewers' responses.

Hypothesis #2: Environments with elevated levels of affective rhetoric are conducive to the creation of polarized communities.

Methods and Data

Data Collection

Data for this chapter was gathered using the Twitter Live Stream function in Gephi version 0.9.1 for Macintosh. All tweets containing #Debates2016 were collected roughly five minutes before the start of the debate and continuing through approximately

one hour of post-debate discussion. The decision to track #Debates2016 was based upon Twitter's announcement that this would be their official hashtag for the 2016 presidential debates. For this analysis, I truncated the full data set of all #Debates2016 tweets to include only those occurring from the start of the debate through approximately six minutes following the debate's conclusion. The reasoning for this was two-fold: First, I assumed that discussion regarding the final minute (94) of the debate would continue for several minutes. Second, examining a full 100 minutes of debate reactions provided a balanced scope for data analysis allowing for an economical time frame divisibility. During the debate, a time log was kept to document the tweet count at each one minute mark, beginning with Lester Holt's first comment welcoming the audience to the debate.

Following the debate, the truncated data set of approximately 1,500,000 tweets was divided into 100 individual one-minute data sets of approximately 15,000 tweets each. Official transcripts for the debate were studied carefully to identify moments of confrontation, controversy, and contempt, then these moments were time stamped by synchronizing them with the official video of the debate. The first presidential debate had no shortage of such moments, so effort was made to choose the most obvious and egregious cases.

[see Appendix A]

Content Analysis

The primary tool for conducting content analysis was Lexicoder 3.0, a software application developed by Mark Daku, Stuart Soroka, and Lori Young at McGill University. This software was used in conjunction with the Regressive Imagery

Dictionary (Martindale 1975, 1990), which is comprised of approximately 3,000 words

divided into three primary categories with 44 sub-categories. These categories are

divided into Primary Processes, which include drive, sensation, defensive symbolization,

regressive cognition, and Icarian imagery; Secondary Processes, which include

abstraction, social behavior, instrumental behavior, restraint, order, temporal references,

and moral imperative; and Emotions, which include positive affect, anxiety, sadness,

affection, aggression, expressive behavior, and glory.[4] For this project, I selected the

Emotions category as a proxy for affective rhetoric. I also selected the Secondary

Processes category as a proxy for the *opposite* of affective rhetoric. Including this second

category provided a valuable comparative measure to provide context for the analysis of

affective rhetoric in the form of emotional responses.

<div align="center">[Insert Table 1.1]</div>

The full text for each minute of observation during the debate was processed

through Lexicoder 3.0 and the Regressive Imagery Dictionary, producing 100 individual

data sets containing raw counts for each of the 14 Emotion and Secondary Process

subcategories. These raw counts were then converted into rates by dividing them into the

total number of words tweeted during each minute of observation. In order to make these

numbers more usable, they have been reported as "positive words per 1,000 words" or

"negative words per 1,000 words".

Network Analysis

---

[4] See Chapter 2

Several network metrics were utilized to measure for signs of relative polarization between networks. First, I tested for modularity and average community size, as these measures provide an indication of how densely clustered neighborhoods of discussion are in the network. Generally speaking, networks with high levels of modularity feature communities of discussion containing dense internal connections and weak external connections to other communities of discussion. Measures of average community size are especially useful when analyzing two or more networks with each other, as this provides a tool for comparing the tendency of individuals in a particular network to cluster together around a certain piece of information.

Second, I tested for closeness centrality distribution and average path length, as these measures provide evidence of a network's efficiency. Generally speaking, networks featuring clusters of nodes with high closeness centrality and with low relative average path lengths tend to be more efficient networks. For example, a network with a closeness centrality distribution of 1 and an average path length of 1 would be a system of perfect information, where each participant would receive the same amount of information in only one step. Conversely, a network with a closeness centrality distribution of close to 0 and an average path length of 6 would be a system where information required a number of steps to travel across the network and where very few participants would be exposed to the same information.

Last, I used network visualization techniques to provide a big picture view of each network's architecture. While network visualizations are less precise than traditional network metrics, they can help to illustrate shifts in the networks as a whole. For example, a dense and interconnected network with a great deal of shared information will

be evident by tight clusters of nodes near the center of the visualizations, while polarized

and disconnected networks will be evident by relatively large clusters of nodes appearing

sporadically near the periphery of the larger network.

Findings

"Does it happen?"

      Hypothesis #1: Moments of confrontation, controversy, and contempt cause an

      instant increase in the rates of affective rhetoric in viewers' responses.

      The first test of this hypothesis involved a contextual view of how rates of

emotion and secondary process words rose and fell over the span of the debate. Rates of

words expressing emotion were used as broad proxies for affective rhetoric, while rates

of words expressing secondary processes were used as broad proxies for objective

discussion. The main goal of this first test was to get a "lay of the land" to see if there

were any identifiable patterns in tweet content and, more importantly, to see if these

patterns aligned with any of the preselected moments of confrontation, controversy, or

contempt during the debate.

[Insert Graph 1.1]

      The first time series analysis showed a clear flow of sentiment shifts throughout

the debate, both for the emotion and secondary process categories. Interestingly enough,

rates of secondary process language were generally higher than rates of emotional

language. This was surprising considering the heated nature of the first presidential

debate. It was also interesting to note that rates of secondary process language peaked

during the middle of the debate, while rates of emotion peaked near the end.

Such trends could have been due to at least two causes. First, it is possible that

the content of the candidates' speech was more policy oriented, less charged with

rhetoric, and more thought-provoking, due to the types of questions asked by the

moderator. Given that television is still largely an entertainment medium, it would not be

surprising if moderators (who are typically also network news anchors) intentionally

order the type and content of their questions in order to create a better "show" for the

audience. Second, it is possible that candidates strategically inserted a higher frequency

of emotional cues during the beginning and end of the debate in order to maximize

rhetorical impact during periods of the debate where viewers' attention was more

focused. Both possibilities are intriguing and could be tested by comparing the content of

Lester Holt's questions with an analysis of the content of the candidates' responses. Such

a content analysis could be performed using the excellent Laver and Garry *Dictionary of

Policy Position* (1992). These interesting observations aside, only rates of emotion were

examined closely in the time series analysis, as this provided the best means for testing

the first hypothesis most directly.

[Insert Graph 1.2]

[Insert Graph 1.3]

The time series analysis of emotion rates in tweets was conducted using two

graphs. The first reported the combined rate of all seven sub-categories of words

expressing emotion from the Regressive Imagery Dictionary, while the second the rates

for each individual sub-category. The purpose of this approach was to provide the opportunity to identify trends in overall emotion while also having the ability to focus on specific types of emotions that may have been driving overall shifts in the wider category.

Sentiment was relatively flat until the first moment where Clinton accused Trump of rooting for the housing crisis, to which Trump responded "That's called business, by the way." There were only slight shifts in emotion rates, driven primarily by expressions of anxiety and glory. A more noticeable upward trend in rates of emotion occurred following Minute 18, following a period of crosstalk between the candidates which ended with Clinton stating, "Well, Donald, I know you live in your own reality." This was driven primarily by increases in expressive behavior and positive affect.

Even more noticeable increases in rates of emotion occurred immediately following Minute 28, where the issues of Trump's tax returns and Clinton's e-mails took center stage. These increases were influenced by significant increases in anxiety and sadness. The shift in rates of sadness was interesting, as prior to this point in the debate rates of sadness were at or slightly above 0. Rates of emotion also trended upwards significantly following Trump's claim that "Secretary Clinton doesn't want to use a couple of words, and that's law and order" and the subsequent debate topic shift to the issue of "stop and frisk" starting at Minute 43. These increases were driven by upward trends in the rates of glory, affection, expressive behavior, and sadness. Interestingly, rates of anxiety remained at 0 during this topic.

Rates of emotion remained fairly steady with very few fluctuations in rates following this interchange and continuing through disputes over which candidate was

responsible for promoting the "birther" controversy, accusations of racist behavior, disputes regarding the role of Russia in the hacking of the Democratic National Committee e-mails, and confrontational moments involving Trump's position that he did not support the 2003 invasion of Iraq. These patterns in tweet content remained consistent through Minute 78.

To be clear, rates of emotion and secondary process did not flat line during this 30 minute period, as rate fluctuations varied by as much as 50% from any given minute to the next. Further, there were noticeable shifts in the rates of individual emotions. For example, rates of anxiety increased at to their highest rates at Minute 47 following Clinton's comment, "I've heard Donald say this at his rallies, and it's really unfortunate that he paints such a dire negative picture of black communities in our country", to which Trump responded, "Ugh." This interchange was accompanied by an increase in rates of anxiety from 0 to 0.273 per 1,000 words (674% above the average rate of anxiety for the entire debate). There was also a corresponding spike in words related to expressive behavior following the dispute between Clinton and Trump regarding whether or not murders in New York City increased following the end of "stop and frisk" policies. These instances aside, fluctuations on the whole did not align clearly and consistently enough with the identified moments of confrontation, controversy, and contempt to draw any conclusions. This could be due to a number of factors, not the least of which would involve the fairly high frequency with which these moments occurred during this 30 minute stretch. Rather put, the almost constant disagreements on a wide range of issues during this period may have created an atmosphere of "background noise".

If the 30 minute segment of the debate featuring a high frequency of back and forth exchanges on multiple issues created a form of background noise, the final 15 minutes were punctuated by two of the most memorable moments of the debate – both of which involved arguably the most controversial statements and confrontational atmosphere, as well as the strongest expressions of contempt. Specifically, these closing minutes of the debate began with the following statements made by Trump at Minute 78 (and 18 seconds):

> "Well, I have much better judgment than she does. There's no question about that. I also have a much better temperament than she has, you know? … I think my strongest asset, maybe by far, is my temperament. I have a winning temperament. I know how to win."

Clinton responded to Trump's one-minute long discussion about temperament with a few playful wiggles of her shoulders, followed by an amused and somewhat exasperated "Whew!" at Minute 79. This interchange elicited by far the most impressive increases in rates of emotion in tweets, rising from a rate of 0.696 per 1,000 words during Minute 77 to rates of 1.221 and 1.270 per 1,000 words during Minutes 78 and 79, and peaking at a rate of 6.555 per 1,000 words during Minute 80. This represented a 942% increase in the rates of emotion in tweets from the moments before Trump began discussing his temperament to the moments after Clinton's response. These increases in the rate of emotion were primarily driven by extreme increases in the rates of aggression and positive affect in tweets. Specifically, between Minute 77 and Minute 80, rates of positive affect increased from 0.0502 per 1,000 words to 0.6741 per 1,000 words (1,343% increase) while rates of aggression increased from 0.424 per 1,000 words to

5.443 per 1,000 words (1,284% increase). Rates of aggression and positive affect at Minute 80 were both at their highest recorded levels for the entire debate, indicating that this interchange had by far the biggest emotional impact on the #Debates2016 Twitter participants.

Rates of emotion in tweets remained high for several minutes after this interchange, staying well about the average rate of emotion over the entire span of the debate. However, the second of the two most memorable moments of the debate aligned with Clinton's scathing criticism of Trump's attempts to "switch (the discussion) from looks to stamina" at Minute 91, and the ensuing confrontation:

Clinton: "this is a man who has called women pigs, slobs and dogs, and someone who has said pregnancy is an inconvenience to employers, who has said …"

Trump: "I never said that."

Clinton: "women don't deserve equal pay unless they do as good a job as men."

Trump: "I didn't say that."

Clinton: "And one of the worst things he said was about a woman in a beauty contest. He loves beauty contests, supporting them and hanging around them. And he called this woman 'Miss Piggy.' Then he called her 'Miss Housekeeping,' because she was Latina. Donald, she has a name."

Trump: "Where did you find this? Where did you find this?"

Clinton: "Her name is Alicia Machado."

Trump: "Where did you find this?"

Clinton: "And she has become a U.S. citizen, and you can bet…"

Trump: "Oh, really?"

Clinton: "… she's going to vote this November."

This interchange between the two candidates, starting at Minute 91 and continuing through Minute 92 (and 30 seconds) aligned with a spike in emotion in tweets from 1.185 per 1,000 words to 2.597 per 1,000 words, representing an increase of 219% in one minute. Rates of emotion in tweets remained above average for the final minute of the debate following this confrontation. These increases in emotion rates were likely driven by increases in the individual measures of aggression, anxiety, glory, and positive affect.

The initial results of the time series analysis provided some support for the first hypothesis, as there were clearly some interchanges that elicited instant emotional responses in individuals discussing the debate on Twitter as it unfolded live on television. However, these results were mixed, as there were moments during the middle third of the debate where no definitive relationship was evident.

These mixed initial results could have been due to the one-minute time frame of observation being too fine of a unit of measurement; it is possible that some debate moments do not lend themselves well to precise measurements in minutes and seconds. Instead, such moments may be better defined as periods of exchange, rather than specific statements. Further, the one-minute time frame of measurement could exclude the responses from individuals who do not respond within 60 seconds. Rather put, while some debate viewers might have responded instantly on Twitter, others might have taken a minute or two to access their phone or computer and might have taken even more time to type out their responses. In order to account for the potential for observation error

when using a one-minute time frame as the unit of measurement, an alternate approach was used when analyzing the time series data.

[Insert Graph 1.4]

[Insert Graph 1.5]

For this brief follow-up analysis, I isolated four distinct segments within the debate where there were especially focused confrontational, controversial, or contemptuous interchanges between the candidates. These segments were six minutes in length, and consisted of interchanges between the candidates involving Trump's taxes and Clinton's e-mails (Minutes 28 – 33), the issue of law and order (Minutes 42 – 47), the issue of temperament (Minutes 78 – 83), and the "words matter" segment centering on Donald Trump's statements about women (Minutes 90 – 95). I also isolated two additional segments to be used as control sets: the opening moments of the debate (Minutes 1 – 6) and the segment of the debate when rates of emotional language in #Debates2016 tweets were at their lowest and steadiest (Minutes 13 – 18).

When examining the time series graph using six-minute segments as the dependent variable, a clear pattern emerged with respect to the moments of the debate which elicited the strongest emotional responses from the #Debates2016 audience. The four largest and most consistent spikes in rates of emotional language were centered around the "temperament", "words matter", "law and order", and "taxes and e-mails" candidate interchanges. These findings, when combined with the first set of findings, lend strong supporting evidence for Hypothesis #1: Moments of confrontation, controversy, and contempt cause an instant increase in the rates of affective rhetoric in viewers' responses.

"Does it matter?"

> **Hypothesis #2**: Environments with elevated levels of affective rhetoric are
> conducive to the creation of polarized communities.

On their own, the findings that there were increases in emotional language in Tweets following moments of confrontation, controversy, and contempt were not all that surprising. However, establishing evidence of such a relationship is an important precondition for testing the second hypothesis. For the second section of this analysis, I used network analysis metrics and visualizations to gain a better understanding regarding the nature and structure of the discrete networks that form in response to moments of confrontation, controversy, and contempt and, more importantly, facilitate increased rates of affective rhetoric in individuals' responses.

For this analysis, I chose to test the same five segments from the second half of the time series analysis. This decision was based upon the reasoning that too fine a level of detail would potentially create similar observational inconsistencies as were experienced in the first time series analysis. Confining the network analysis observations to a one-minute time frame reduced measurement of the independent variable to a single candidate statement, rather than taking into account the full content of an exchange between candidates or the development of a candidate's argument or defense. Further, given that the resulting networks are the product of individuals' comments, it was likely

that confining the analysis to one-minute time frames would hamper the value of the
dependent variable by omitting contributions by individuals who took more than 60
seconds to access their phone or computer and to post their responses to Twitter.

Visualizations

Given that the second hypothesis seeks to find a relationship between elevated
levels of affective rhetoric and the formation of polarized communities, the visualizations
were analyzed by comparing the network with the lowest rate of emotional language
(Minutes 13 – 18) with the networks containing the highest rates of emotional language
("Taxes & E-mails", "Law & Order", "Temperament", and "Words Matter"). Using rates
of emotional language as a proxy for affective rhetoric, there should be noticeable
differences in the general structure of these four networks compared to the control sample
network.

[Insert Visualization 1.1]

[Insert Visualization 1.2]

[Insert Visualization 1.3]

[Insert Visualization 1.4]

When analyzing each of these "high emotion" debate segments with the "low
emotion" control segment, there were some notable differences. For example, while
there were several clustered communities appearing outside of the center of discussion,
these were not located relative far from the center. Additionally, these clusters had
several clear connections with each other, suggesting that a measure of communication
was occurring between them. Conversely, visualizations for the four "high emotion"

debate segments featured a higher number of clustered communities appearing outside of the center of discussion. Further, these clusters were closer to the periphery of the network while demonstrating fewer connections with other clusters. This suggests these networks were somewhat more polarized than the network for the control segment.

While network visualizations are often useful analytical tools for examining interactions between individuals in a networked environment, there are times with such an analysis is limited in its descriptive power for testing hypotheses. Ultimately, while the comparison of network visualizations immediately following each of the five debate themes did illustrate variations in network architecture as a whole, these comparisons did not provide robust descriptive power for indentifying the formation of polarized communities. This could have been due to a combination of the high volume of tweets and the small time frame within which these tweets were being made.

Comparative Network Metrics

A final battery of tests was applied in order to look for evidence of network polarization during different segments during the debate. Such evidence was sought by measuring changes in network efficiency and community density. Rather put, if the network during a particular six-minute segment was less efficient (indicated by high average path lengths) and more dense (indicated by high average community sizes) than another six-minute segment, this would suggest a stronger likelihood of concentrated polarization in the former segment than in the latter.

[Insert Graph 1.6]

[Insert Graph 1.7]

[Insert Graph 1.8]

Where the network visualizations did not provide the most robust evidence of shifts in network polarization during varying periods of emotional language by network participants, the comparative network metrics analysis provided some compelling evidence. This evidence is particularly evident in Graph 1.7, which measures relative efficiency and intensity of the networks during different segments of the debate.

Generally speaking, networks with high average path lengths are less efficient for sharing information, as an increase in path length means information must make a higher number of "jumps" between one node to the next. Networks with an average path length of 1 represent an environment of perfect information, as all messages in that network reach all other nodes in that network in one step. Average path length can also be viewed as one indicator of polarization, since the network as a whole becomes more unified as this value decreases and becomes more dispersed as this value increases.

Average community size can be viewed as one indicator of network concentration, as it measures the average number of nodes connected within a shared neighborhood. In essence, these two values are best interpreted as a pair: Networks with high average path lengths and high average community sizes suggest high levels of concentrated polarization, while networks with low average path lengths and low average community sizes suggest low levels of diffuse polarization. Similarly, networks with high average path lengths and low community sizes suggest high levels of diffuse polarization, while networks with low average path lengths and high average community sizes suggest low levels of concentrated polarization.

This analysis provides strong evidence to support Hypothesis #2, as the two debate segments which elicited the highest rates of emotional language also produced the two highest combined measures of average path and average community size. Specifically, the networks that developed in the wake of the "temperament" and "words matter" interchanges were marked by especially concentrated polarized communities of discussion. Such findings are especially noteworthy considering these two segments also contained arguably the most emotionally charged and personalized interchanges of the debate. Conversely, the "lowest emotion" segment of the debate produced the lowest combined measures of average path and average community size, while the "debate start" segment (representing the segment with the second lowest rates of emotion) produced the second lowest combined measures of average path and average community size.

Conclusions and Discussion

The first goal of this chapter was to determine whether a relationship exists between polarizing behavior in the form of elite cues and increases in mass affective rhetoric. There were clear relationships between moments of confrontation, controversy, and contempt between Hillary Clinton and Donald Trump during the first presidential debate and subsequent rises in affective rhetoric in the form of emotional language among #Debates2016 participants on Twitter. This was especially evident in the emotional subcategories of positive affect and aggression. By offering evidence to the question of "Does it happen?", this finding provides a valuable tool for testing whether or not there exists a causal chain between elite polarization and mass polarization in social media.

The second goal of this chapter was to determine whether or not elevated levels of affective rhetoric in the form of emotional language shape the architecture of the networks in a way that is conducive to mass polarization. Significant evidence was found indicating that moments of intense confrontation, controversy, and contempt correlated with indicators of polarized networks. This was especially true for moments where interchanges between the candidates included especially personal character attacks. It should be stressed that while this evidence does not demonstrate the existence of mass polarization in response to exposure to elite polarization, it does demonstrate that elite polarization can influence the creation of polarized communities which could, in turn, facilitate affective polarization. In doing so, such evidence warrants a closer look into the nature of communication that is occurring within these polarized communities. Such an approach is pursued in Chapter 5.

Appendix A: Moments of Confrontation, Controversy, and Contempt during the first 2016 U.S. Presidential Debate

12:00   Trump: "That's called business, by the way"
(12)   (Clinton said Trump "rooted for the housing crisis")

13:45   Clinton: "Donald thinks that climate change is a hoax perpetrated by the Chinese"
(14)   Trump: "I did not. I did not. I do not say that."
            …
        Trump: "She talks about solar panels. We invested in a solar company, our country.
                That was a disaster."

15:40   Trump: "You've been doing this for 30 years."
(16-18)   [CROSSTALK]
                NAFTA "is the single worst trade deal ever approved in this country."

17:20   Trump: "You haven't done it in 30 years or 26 years" etc
                [CROSSTALK]
        Clinton: "Well, that's your opinion. That is your opinion."

18:20   Trump: "You called [the TPP] the gold standard."
(18)   [CROSSTALK]
        Clinton: "Well, Donald, I know you live in your own reality,"
                [CROSSTALK]
        Holt: "We're going to move to …"

25:30   Clinton: "Trump loophole"
(26)   [CROSSTALK]
                "Trumped up trickle down"
                Debt free college

28:00   [Holt begins Q&A about Trump releasing his tax returns]
(28)
30:00   Trump: "I will release my tax returns, against my lawyer's wishes, when she releases her
                33,000 e-mails that have been deleted."

31:00   Clinton: "Maybe he's not as rich as he says he is."
                "Maybe he doesn't want the American people … to know that he's paid nothing
                in federal taxes."
        Trump: "That makes me smart."
        Clinton: "it must be something really important, even terrible, that he's trying to hide."

33:00   Clinton: "I made a mistake using a private e-mail."
(33)   Trump: "That's for sure."
        [more talk about tax returns]

36:00   Clinton: "And maybe because you haven't paid any federal income tax for a lot of years."
(36)
37:00   Clinton: "Do the thousands of people that you have stiffed over the course of your
                business not deserve some kind of apology from someone who has taken
        their
                labor, taken the goods that they produced, and then refused to pay them?"

43:00   Trump: "Secretary Clinton doesn't want to use a couple of words, and that's law and
(43)   order. And we need law and order. If we don't have it, we're not going to have
                a country.

45:00   [Stop and Frisk]

47:00   Clinton: "I've heard Donald say this at his rallies, and it's really unfortunate that he paints
(47)   such a dire negative picture of black communities in our country."
        Trump: "Ugh."

51:30   Trump: "I agree … I think we have to look very strongly at no-fly lists and watch lists."

52:00   Trump: "you were the one that brought up the words super-predator about young
                black youth"

53:00   Clinton: "under the current mayor, crime has continued to drop, including murders.
(53)   So there is..."
        Trump: "No, you're wrong. You're wrong."
        Clinton: "No, I'm not."
        Trump: "Murders are up. All right. You check it.

54:30   Trump: "Look, the African-American community has been let down by our politicians.
                They talk good around election time, like right now, and after the election, they
                said, see ya later, I'll see you in four years."

55:20   Clinton: "think Donald just criticized me for preparing for this debate. And, yes, I did.

And you know what else I prepared for? I prepared to be president. And I think

that's a good thing."

56:00   [Obama birth certificate]
(56)
58:40   Clinton: "Well, just listen to what you heard. [Trump] tried to put the whole racist
            birther lie to bed … started his political activity based on this racist lie."

60:00   Clinton: "[Trump] has a long record of engaging in racist behavior."
            … "When they go low, we go high."

61:00   Trump: "you even sent out or your campaign sent out pictures of him in a certain garb,

            very famous pictures. I don't think you can deny that."

63:30   Clinton: "I know Donald's very praiseworthy of Vladimir Putin"
(64)
65:00   Clinton: "I was so shocked when Donald publicly invited Putin to hack into
Americans."

66:30   [Trump "400 pound hacker"]
        "You don't know who broke in to DNC. But what did we learn with DNC? We learned

            that Bernie Sanders was taken advantage of by your people, by
            Debbie Wasserman Schultz."

71:00   Clinton: "Donald supported the invasion of Iraq."
(71)   Trump: "Wrong."
        Clinton: "That is absolutely proved over and over again."
        Trump: "Wrong. Wrong."

76:30   [More on Trump supporting Iraq war]
        [CROSSTALK] with Holt

78:18   Trump: "Well, I have much better judgment than she does. There's no question about
(78)   that. I also have a much better temperament than she has, you know?"
        Trump: "I think my strongest asset, maybe by far, is my temperament. I have a winning

            temperament. I know how to win."
        Trump: "I don't know who you were talking to, Secretary Clinton, but you were totally

            out of control. I said, there's a person with a temperament that's got a
        problem."

79:00　Clinton: "Whew, OK."
　　　　[Shoulder shimmy]

80:40　[Dispute over Iranian sailors taunting American sailors]

81:10　[Dispute over use of nuclear proliferation]
(81)
82:00　Clinton: "a man who can be provoked by a tweet should not have his fingers anywhere
　　　　　　　　near the nuclear codes"
　　　　Trump: "That line's getting a little bit old, I must say."

84:45　Trump: "another one powerful is the worst deal I think I've ever seen negotiated that you
　　　　　　　　started is the Iran deal."

86:00　Clinton: "Well, let me -- let me start by saying, words matter. Words matter when you
(86)　　　　　run for president. And they really matter when you are president."

89:30　["I don't think she has the presidential look."]

90:45　Trump: "Hillary has experience, but it's bad experience. We have made so many bad
　　　　　　　　deals during the last -- so she's got experience, that I agree."

91:30　Clinton: "You know, he tried to switch from looks to stamina. But this is a man who has
(91)　　　　　called women pigs, slobs and dogs, and someone who has said pregnancy is an
　　　　　　　　inconvenience to employers, who has said..."
　　　　Trump: "I never said that."
　　　　Clinton: "… women don't deserve equal pay unless they do as good a job as men."
　　　　Trump: "I didn't say that."

91:50　Clinton: "And one of the worst things he said was about a woman in a beauty contest. He
　　　　　　　　loves beauty contests, supporting them and hanging around them. And he called
　　　　　　　　this woman "Miss Piggy." Then he called her "Miss Housekeeping," because she
　　　　　　　　was Latina. Donald, she has a name."
　　　　Trump: "Where did you find this? Where did you find this?"

92:15　Clinton: "Her name is Alicia Machado."

Trump: "Where did you find this?"
Clinton: "And she has become a U.S. citizen, and you can bet..."
Trump: "Oh, really?"
Clinton: "... she's going to vote this November."

93:00   Trump: "I was going to say something extremely rough to Hillary, to her family, and I

said to myself, "I can't do it. I just can't do it. It's inappropriate. It's not nice."

93:30   [Are you willing to accept the outcome as the will of the voters?"
Clinton: "Well, I support our democracy. And sometimes you win, sometimes you lose.

But I certainly will support the outcome of this election."

94:50   Trump: "The answer is, if she wins, I will absolutely support her."

Table 1.1: Emotion and Secondary Process Categories, Regressive Imagery Dictionary

SECONDARY PROCESS	
Abstraction	Know, may, thought
Social Behavior	Say, tell, call
Instrumental Behavior	Make, find, work
Restraint	Must, stop, bind
Order	Simple, measure, array
Temporal References	When, now, then
Moral Imperative	Should, right, virtue
EMOTIONS	
Positive Affect	Cheerful, enjoy, fun
Anxiety	Afraid, fear, phobic
Sadness	Depression, dissatisfied, lonely
Affection	Affectionate, marriage, sweetheart
Aggression	Angry, harsh, sarcasm
Expressive Behavior	Art, dance, sing
Glory	Admirable, hero, royal

Source: Provalis Research, https://provalisresearch.com/products/content-analysis-software/wordstat-dictionary/regressive-imagery-dictionary/

Graph 1.1: #Debates2016 Tweets during First Presidential Debate, Rates of Secondary Process vs. Rates of Emotion

#Debates2016 Tweets during First Presidential Debate:
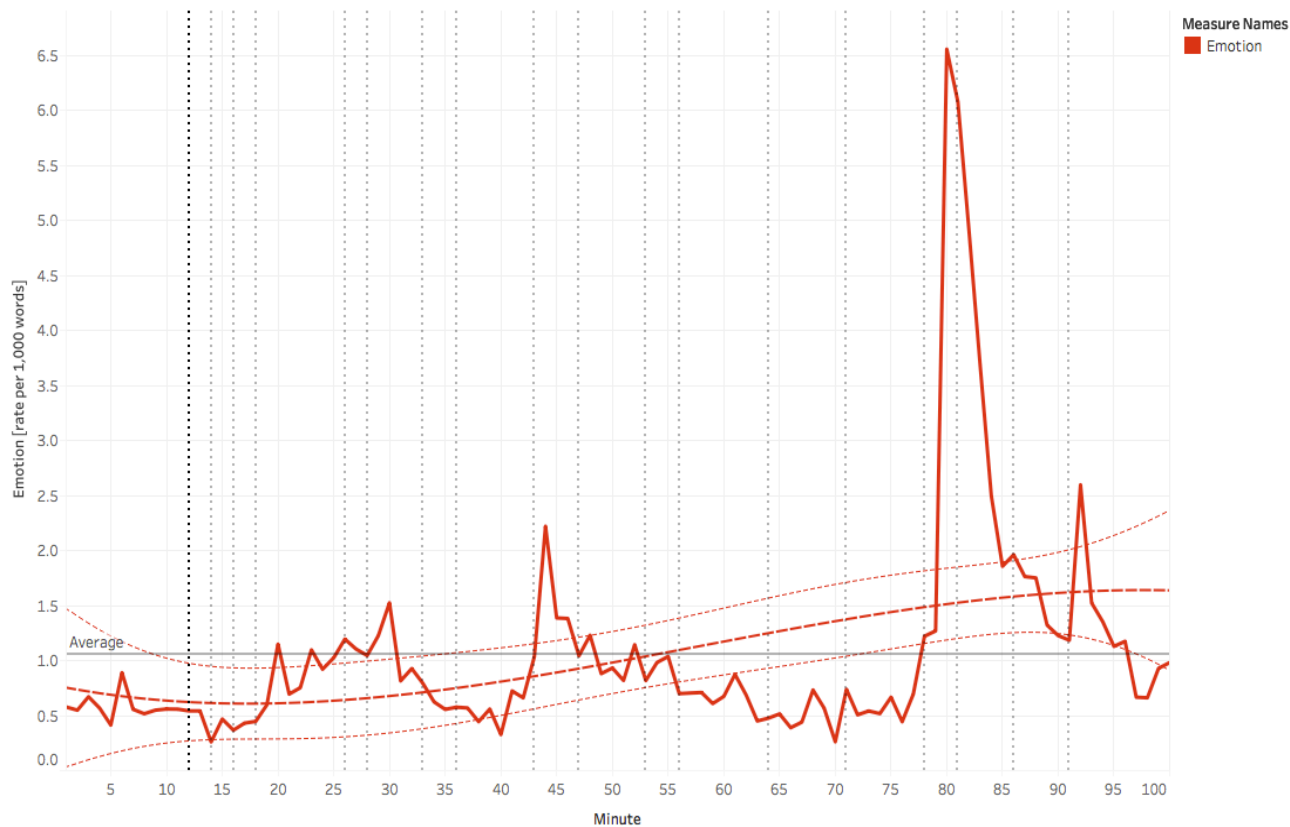Secondary Process vs. Emotion
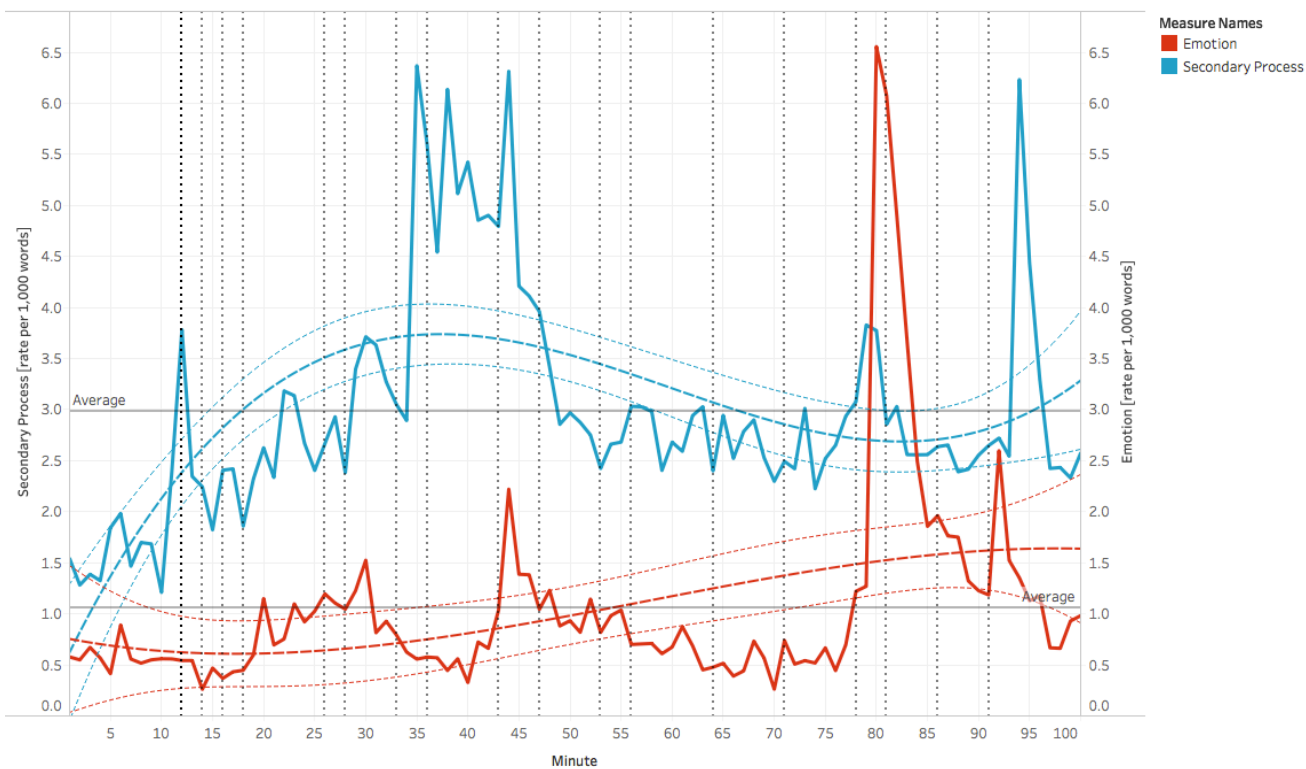
Sources: Twitter Live Stream, 9/26/2016 from 7:00pm to 10:30pm Central Standard Time;
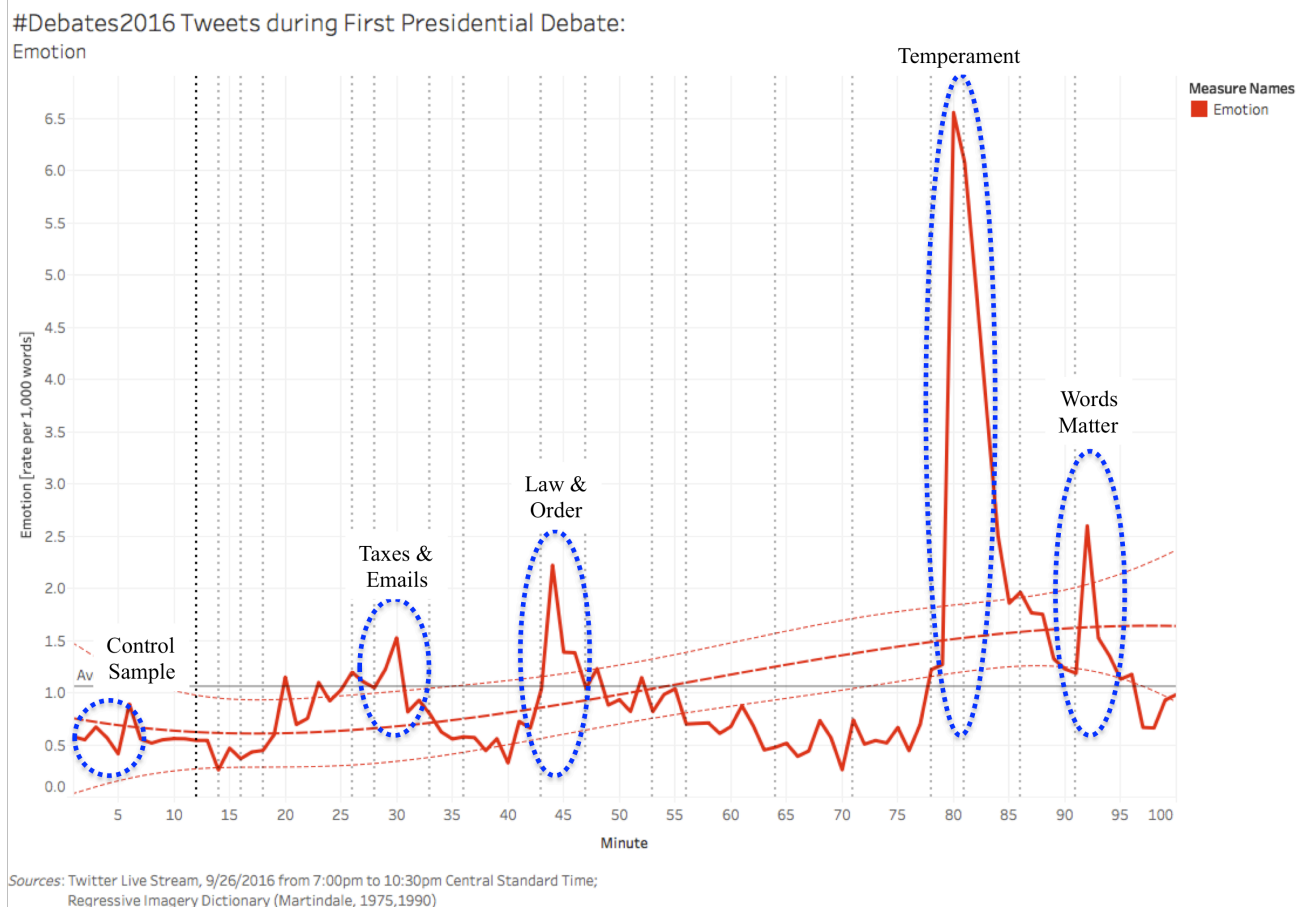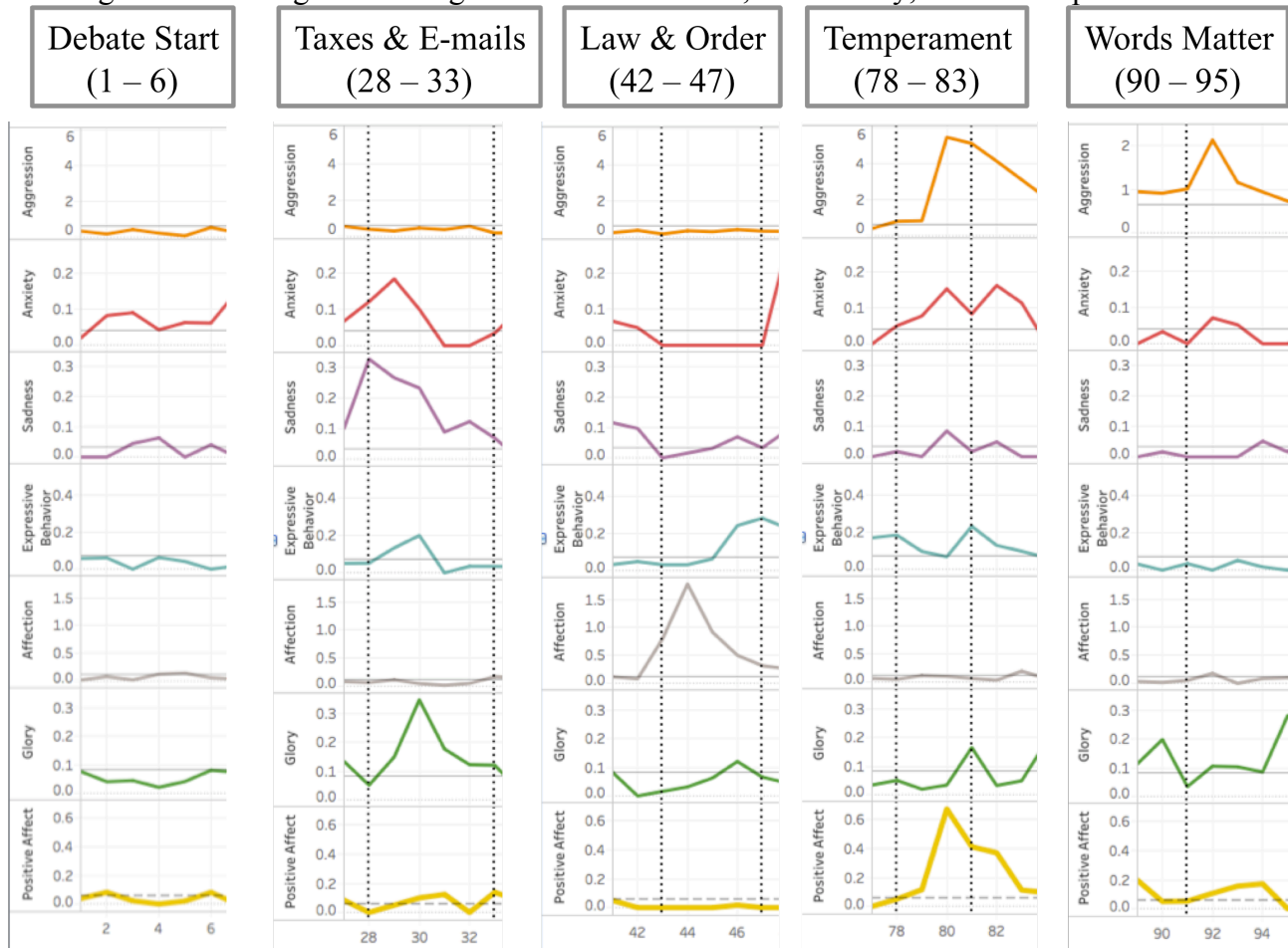Regressive Imagery Dictionary (Martindale, 1975,1990)

Graph 1.2: #Debates2016 Tweets during First Presidential Debate, Rates of Emotion

#Debates2016 Tweets during First Presidential Debate:
Emotion

Sources: Twitter Live Stream, 9/26/2016 from 7:00pm to 10:30pm Central Standard Time;
Regressive Imagery Dictionary (Martindale, 1975,1990)

Graph 1.3: #Debates2016 Tweets during First Presidential Debate, Rates of Emotion by Category

#Debates2016 Tweets during First Presidential Debate:
Emotion by Category

Graph 1.4: #Debates2016 Tweets during First Presidential Debate, Rates of Emotion
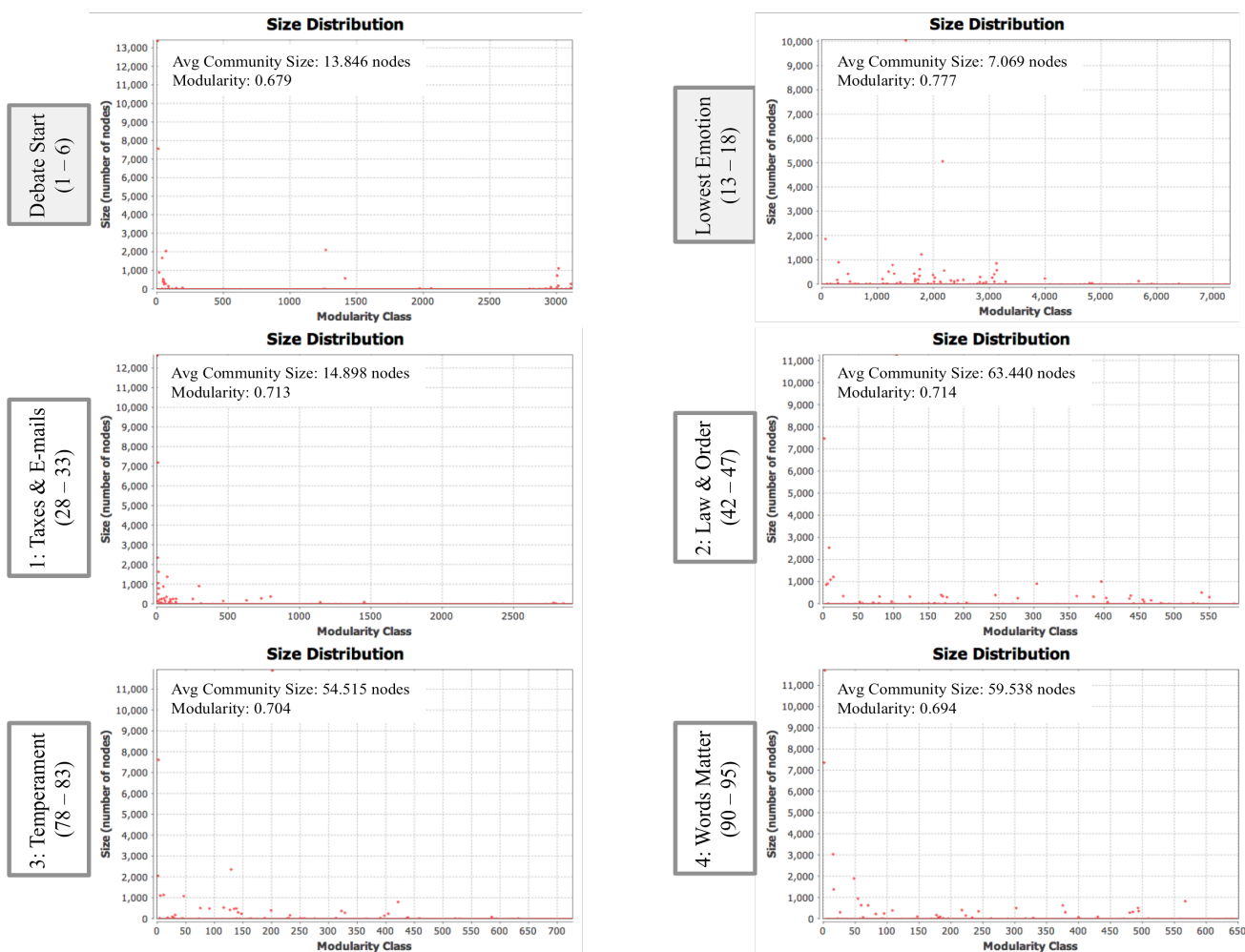Isolating six-minute segments of significant confrontation, controversy, and contempt

Graph 1.5: #Debates2016 Tweets during First Presidential Debate, Rates of Emotion by Category
Isolating six-minute segments of significant confrontation, controversy, and contempt


| Debate Start (1 – 6) | Taxes & E-mails (28 – 33) | Law & Order (42 – 47) | Temperament (78 – 83) | Words Matter (90 – 95) |

Graph 1.6: #Debates2016 Tweets during First Presidential Debate, Average Community Size and Modularity
Isolating six-minute segments of significant confrontation, controversy, and contempt

Graph 1.7: #Debates2016 Tweets during First Presidential Debate, Average Path Length and Closeness
Isolating six-minute segments of significant confrontation, controversy, and contempt
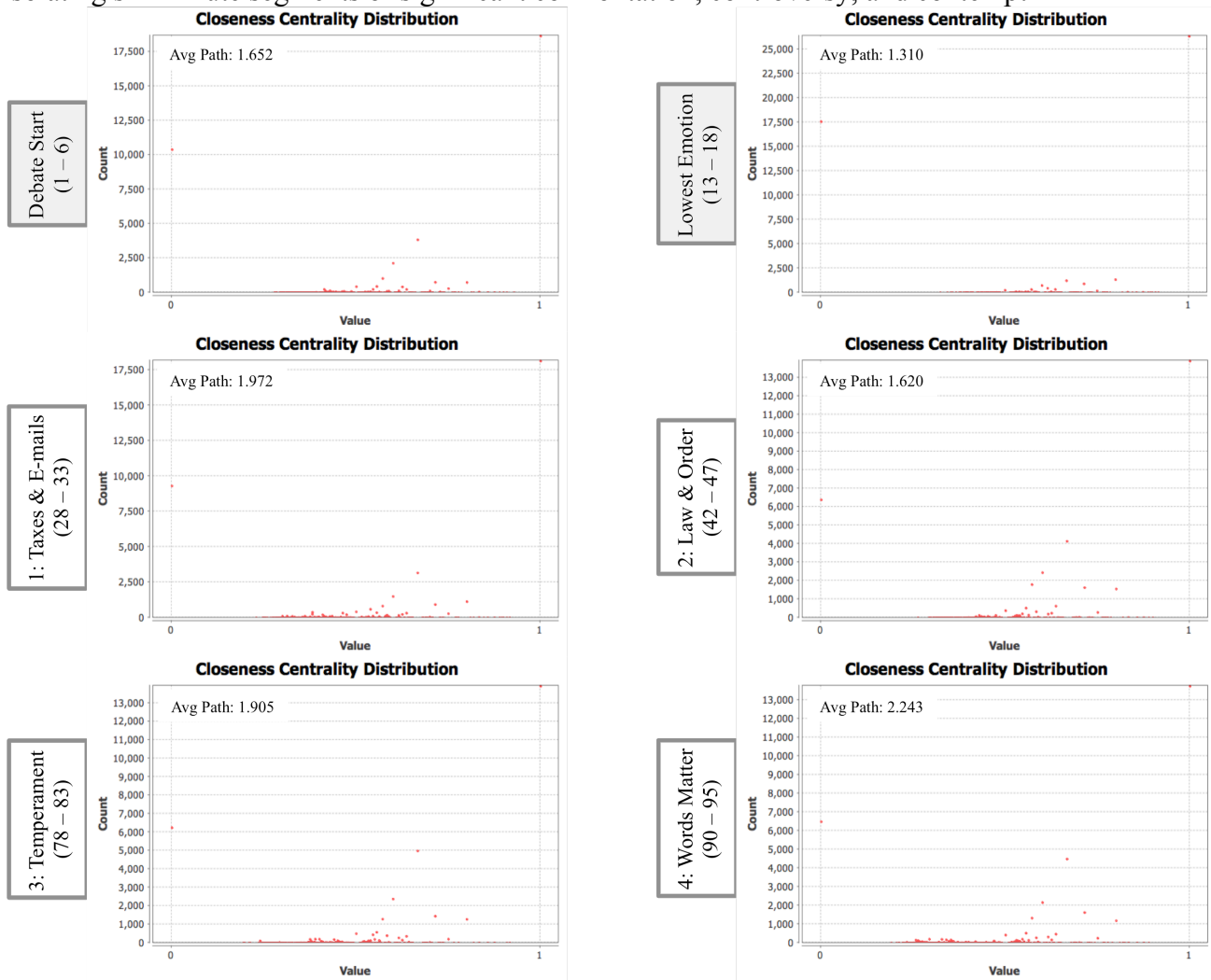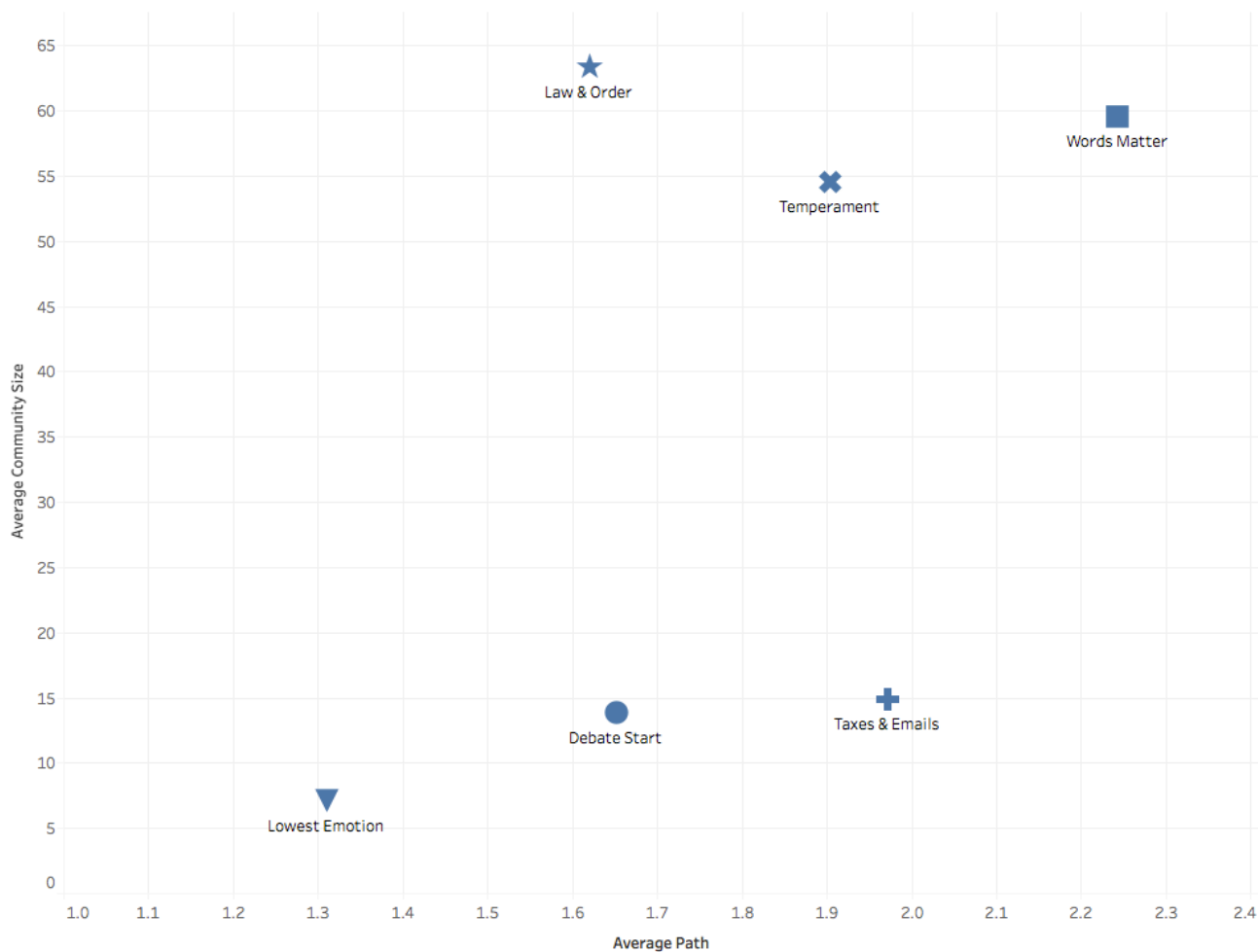
Graph 1.8: #Debates2016 Tweets during First Presidential Debate, Relative Community Polarization

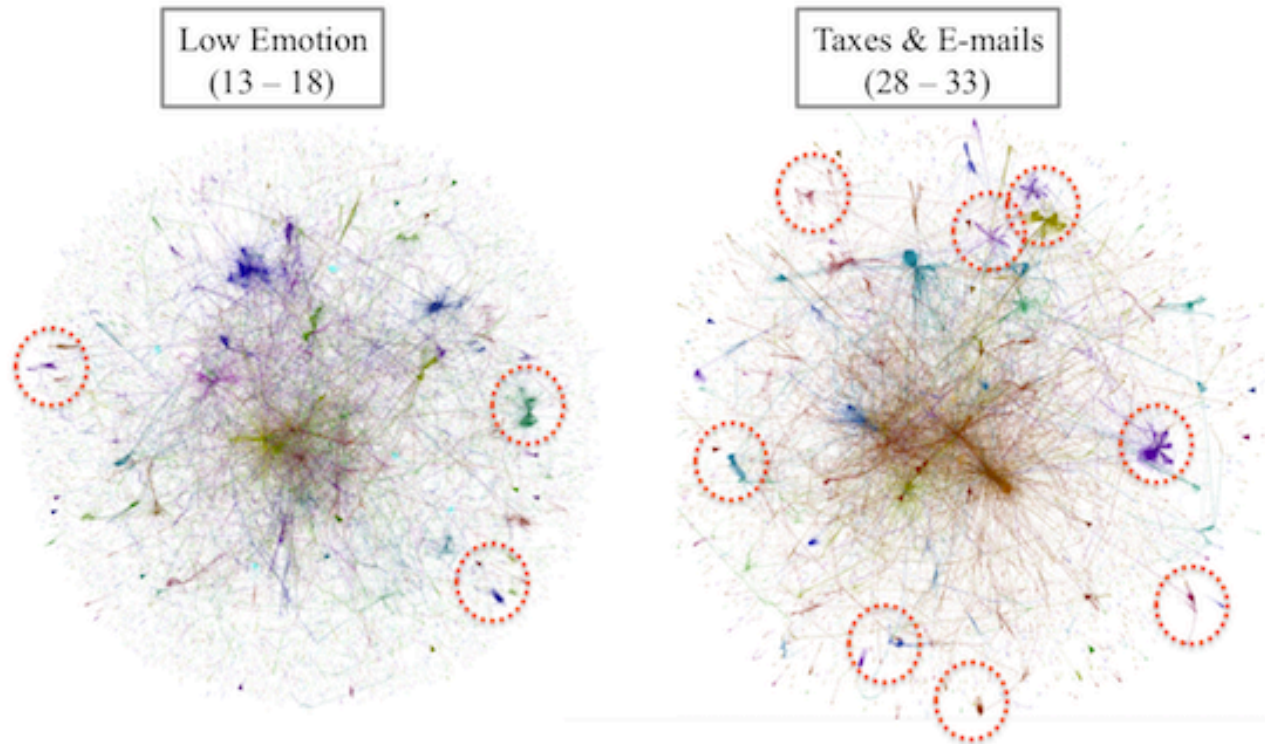Isolating six-minute segments of significant confrontation, controversy, and contempt

Moments of Confrontation, Controversy, and Contempt in the First 2016 Presidential Debate
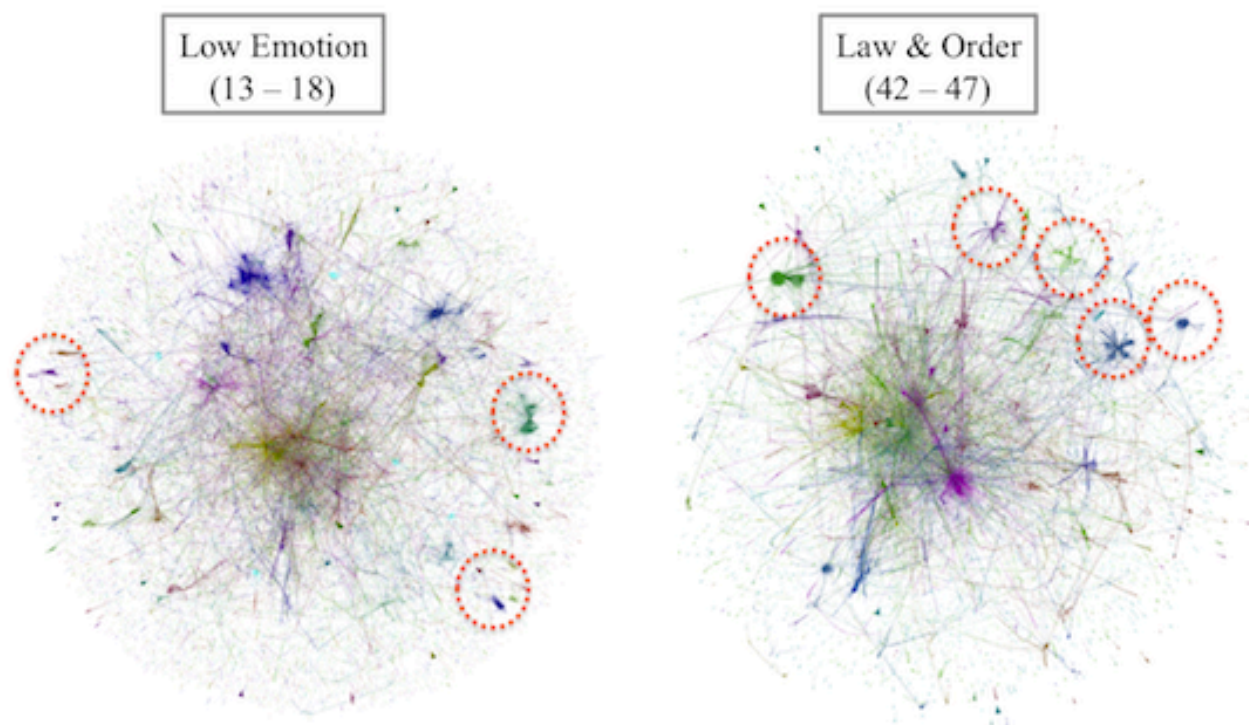#Debates2016 Relative Community Polarization and Intensity

Source: Twitter (9/26/2016)

Visualization 1.1: #Debates2016 "Low Emotion" segment vs. "Taxes & Emails" segment
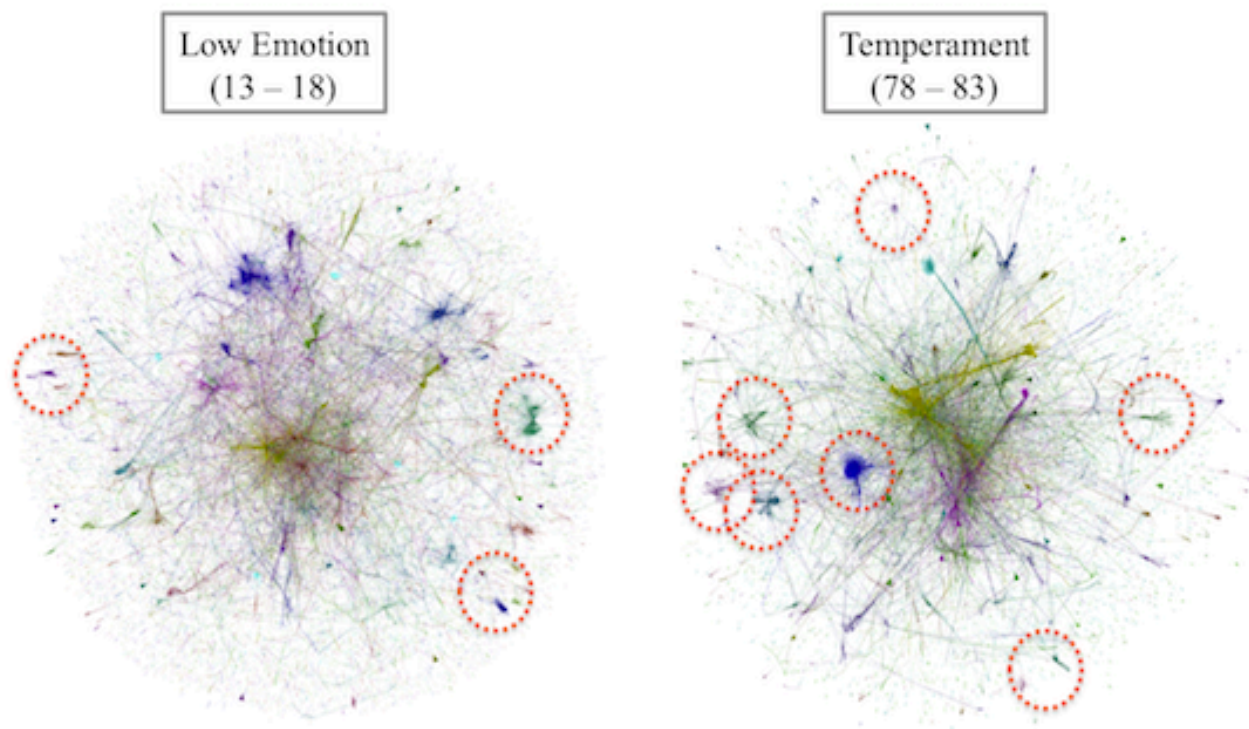

Source: Twitter (9/26/2016) and Gephi 0.8.2

Visualization 1.2: #Debates2016 "Low Emotion" segment vs. "Law & Order" segment
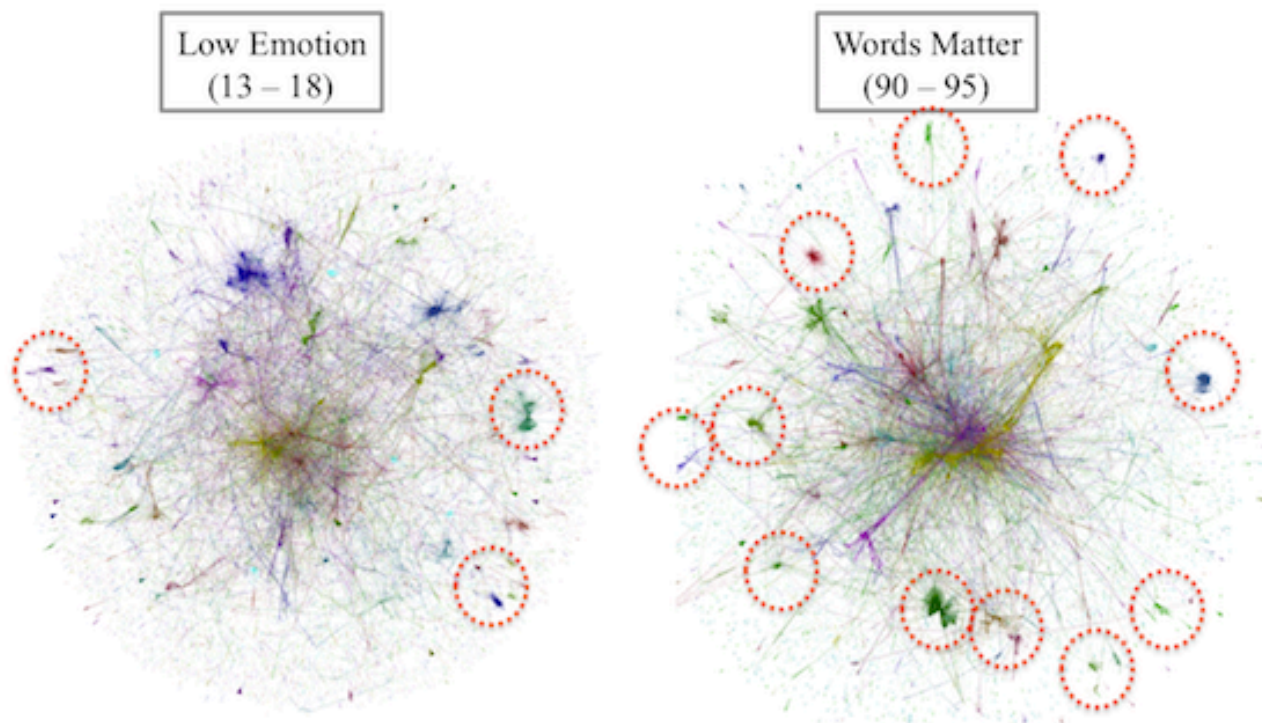

Source: Twitter (9/26/2016) and Gephi 0.8.2

Visualization 1.3: #Debates2016 "Low Emotion" segment vs. "Temperament" segment

Source: Twitter (9/26/2016) and Gephi 0.8.2

Visualization 1.4: #Debates2016 "Low Emotion" segment vs. "Words Matter" segment

Source: Twitter (9/26/2016) and Gephi 0.8.2

Dissertation Chapter 6

Mountains or Molehills?

Examining the Trump Effect on Twitter

Eric C. Vorst

Introduction

On April 13, 2016, the Southern Poverty Law Center published a report asserting Donald Trump's presidential campaign was "producing an alarming level of fear and anxiety" in our nation's schools while "inflaming racial and ethnic tensions" (Costello 2016). Labeled the "Trump Effect", this phenomenon received significant media attention and was cited by Hillary Clinton in her August 25, 2016, "alt-right" speech. The report raised important questions regarding how presidential candidates' patterns of speech may influence feelings of resentment, fear, or incivility. However, consistent with the Southern Poverty Law Center's own admissions, the report lacked scientific rigor in its methodology and reporting. Despite these flaws, the concerns voiced by the Southern Poverty Law Center have significant societal implications, especially when placed into the context of the United States presidency. As such, these concerns beg for scientific research to examine the relationship between messages delivered by political elites and the behavior of those who receive these messages.

This chapter seeks to determine the extent to which a "Trump Effect" manifests itself in social media. First, I approach this question by examining whether a temporal relationship exists between controversial statements made by Donald Trump during the span of his presidential campaign and a resulting rise in affective rhetoric in the form of aggression, anxiety, or negativity in discussion about Donald Trump on Twitter. Second, I examine whether or not these comments influenced social networks in a way that was more conducive to affective polarization, while also examining the extent to which these networks facilitated potential confrontation. When combined, evidence of such relationships would help identify the existence of a Trump Effect in social media (or,

"Does it happen?") while providing a measure of how deeply such an effect impacts civil

discourse (or, "Does it matter?").

Research Design

Research Questions

> **Research Question #1:** Do controversial comments by political elites affect
>
> social media discourse negatively?

The first section of this chapter addresses the question: "Does it happen?". I

explore this question by measuring whether a relationship exists between controversial

statements made by Donald Trump during the span of his presidential campaign and a

resulting rise in affective rhetoric in the form of aggression, anxiety, or negativity in

discussion about Donald Trump on Twitter. For my independent variables, I isolate

especially controversial and insensitive remarks made by Donald Trump over the span of

his primary and general election campaigns. For my dependent variable, I use levels of

aggression, anxiety, and negativity in tweets mentioning Donald Trump during this same

time span. Last, I use time series analysis to determine whether there is a relationship

between Donald Trump's comments and a resulting rise in affective rhetoric.

> **Research Question #2:** How do controversial comments by political elites in
>
> online spaces shape the network architectures that facilitate these discussions?

The second section of this chapter addresses the question: "Does it matter?". I explore this question by analyzing how networks of discussion behave on days where Donald Trump made especially controversial comments. During these days, I conduct network metrics analyses to measure the size of communities of discussion and the relative efficiency of the networks within which these communities form. These metrics provide evidence as to the relative levels of polarization within a given network at a given time. If evidence of polarized communities exist, this would suggest a possible causal link between Donald Trump's statements and polarized environments of discussion. Further, I seek evidence of how these networks facilitate discussion of Donald Trump's supporters as well as of his critics.

Hypotheses

Hypothesis 1: Controversial comments by Donald Trump during the 2016 U.S. presidential primary and general election caused a measurable increase in rates of anxiety, aggression, and negative affect on social media.

Hypothesis 2: Controversial comments by Donald Trump involving during the 2016 U.S. presidential primary and general election contributed to the creation of polarized and confrontational environments on social media.

Data and Methods

Data Collection

I test these hypotheses by using an original data set of several million tweets mentioning Donald Trump collected daily from 9/1/2015 through 3/1/2017 using the NodeXL template for Microsoft Excel (Social Media Research Foundation 2017).[5] I use content analysis and network metrics to measure the extent to which levels of anxiety, aggression, and negative sentiment in tweets mentioning Donald Trump align temporally with controversial public comments he has made during his primary and general election campaigns.

[Insert Table 1.1]

Content Analysis

As was the case in previous chapters, the primary tool for conducting content analysis was Lexicoder 3.0, a software application developed by Mark Daku, Stuart Soroka, and Lori Young at McGill University. This software was used in conjunction with the Regressive Imagery Dictionary (Martindale 1975, 1990) and the Lexicoder Semantic Dictionary (Daku, Soroka, and Young 2016). The Regressive Imagery Dictionary is comprised of approximately 3,000 words divided into three primary categories with 44 sub-categories. These categories are divided into Primary Processes, which include drive, sensation, defensive symbolization, regressive cognition, and Icarian imagery; Secondary Processes, which include abstraction, social behavior, instrumental behavior, restraint, order, temporal references, and moral imperative; and Emotions,

---

[5] See Chapter 2

which include positive affect, anxiety, sadness, affection, aggression, expressive

behavior, and glory. The Lexicoder Semantic Dictionary draws upon a dictionary of

approximately 5,000 words and is designed to measure the positive and negative

sentiment in political texts.[6]

[Insert Table 1.2]

The full text for each day of observation from 9/1/2015 to 3/1/2017 was processed

through Lexicoder 3.0 using both sentiment dictionaries, producing 548 individual data

sets containing raw counts for anxiety, aggression, and negativity. These raw counts

were then converted into rates by dividing them into the total number of words tweeted

during each day of observation. In order to make these numbers more usable, they have

been reported as "rate of anxiety per 1,000 words", "rate of aggression per 1,000 words",

and "rate of negativity per 1,000 words".

Network Metrics Analysis

Several network metrics were utilized to measure for signs of relative polarization

between networks. First, I tested for average community size, as this measure provides

an indication of how densely clustered neighborhoods of discussion are in the network.

Generally speaking, networks with high average community sizes indicate a tendency of

individuals in a particular network to cluster together around certain topics, themes, or

pieces of information. Second, I tested for average path length as a percentage of

network diameter, as this measure provides evidence of a network's efficiency and

tendency towards homophily. Generally speaking, networks featuring low relative

---

[6] See Chapter 2

average path lengths tend to be more efficient networks. As a network's average path length approaches its diameter, this indicates the network is becoming less efficient, as more "steps" are required for a message to reach one end of the network to the other. For example, if a network had a diameter of 5 and an average path length of 5, this would mean that all messages in the network needed to pass through 5 steps in order to be shared.[7]

Last, I used network visualization techniques to provide an interactive and "big picture" view of how different messages propagate throughout different networks' unique architectures. This allowed for a unique look at how social networks react to extremely controversial comments made on the part of political elites by demonstrating how certain types of language, themes, or narratives propagated within the unique structures formed by a social network over a discrete time frame. Most importantly, these techniques allowed for specific messages – such as those containing hashtags supportive (or critical) of Donald Trump – within the network to be highlighted so that their influence within the network could be assessed spatially.

Findings

"Does it happen?"

> **Hypothesis 1**: Controversial comments by Donald Trump during the 2016 U.S. presidential primary and general election caused an increase in rates of anxiety, aggression, and negative affect on social media.

---

[7] See Chapter 3

The first test of this hypothesis involved measuring how rates of anxiety, aggression, and negative affect increased and decreased over the 548-day span of observations. The main goal of this first test was to obtain a contextual assessment of fluctuations in the communications environment to see if there were any identifiable patterns in tweet content and, more importantly, to see if these patterns aligned with any of the preselected controversial remarks.

[Insert Graph 1.1]

Upon initial analysis, it was clear that there were significant increases and decreases in these types of language over time. It was not initially as clear whether or not these increases and decreases aligned with specific comments made by Donald Trump. This was likely due to the wide span of observation combined with several significant spikes which created scaling issues that may have obscured some hidden effects. For example, if the highest rate of anxiety measured was 14 per 1,000 words on one day, and the rates on all remaining 547 days were below 1 per 1,000 words, rate variations within those days would be hidden due to the predominance of a one-day spike. This is especially important to consider, given such a spike could – and likely should – be viewed as an outlier. As such, the time frame of observation was subdivided into four-month periods in order to gain a clearer focus. These subdivisions are reported in the following four graphs:

[Insert Graph 1.2]

[Insert Graph 1.3]

[Insert Graph 1.4]

[Insert Graph 1.5]

Despite decreasing the time frame of observation into four-month periods, there were no clear patterns evident between any of the 20 pre-selected controversial remarks by Donald Trump and a resulting rise in aggression, anxiety, or negativity in tweets mentioning Donald Trump. This is not to say that increases in such language on Twitter did not align with some of these comments by Donald Trump. Levels of aggression rose following Donald Trump's accusations that Hillary Clinton was playing the "woman card", after tweeting "I love Hispanics!", after his reference to "Second Amendment people" being able to do something about stopping Hillary Clinton's election, after the release of the "Access Hollywood" tape, and after his claim that "millions of people voted illegally" in the election. Similarly, levels of negativity rose following Donald Trump's comments referencing Carly Fiorina's face, him asking "How stupid are the people of Iowa?", his claim that he could "shoot someone on 5th Avenue" and not drop in the polls, his failure to immediately disavow David Duke, his claim that "Islam hates us", Hillary Clinton "playing the woman card", comments following the Pulse Nightclub shooting in Orlando, and the release of the "Access Hollywood" tapes.

However, rates in aggression and negativity also dipped on other dates where Donald Trump made controversial remarks. Further, there were significant spikes in aggression and negativity on dates where Donald Trump did not make especially controversial remarks. This suggests that while there may be a causal link between Donald Trump and increases in aggression and negativity, such a relationship may not be driven by his controversial remarks – at least not in social media. Thus, I did not find compelling evidence to either fully support or reject Hypothesis #1.

While a Trump Effect may exist, it was not readily evident using this method of testing. However, this could have been due to a number of flaws in the research design. For example, there may have been an error in selecting independent variables. Given that there were spikes in affective rhetoric on certain days, this suggests that something was causing a reaction among Twitter users as they were discussing Donald Trump. It is possible that there were other actions, comments, or interactions by Donald Trump which led to these reactions. Given that Donald Trump made a large number of campaign speeches where a variety of extracurricular events occurred, it is possible that one of these events could have elicited negative reactions among Twitter users.

Further, it is possible that there were better choices for measuring the dependent variable. While rates of aggression, anxiety, and negativity seem to fit the Southern Poverty Law Center's notion of a Trump Effect, there could be other ways of performing content analysis that would be more effective in drawing such sentiments out of Tweets. For example, aggression, anxiety, and negativity do not necessarily equate to indicators of "bullying" behavior. In sum, it is entirely possible that a more applicable and relevant content dictionary would lead to different results in the content analysis.

"Does it matter?"

> **Hypothesis 2**: Controversial comments by Donald Trump involving during the 2016 U.S. presidential primary and general election contributed to the creation of polarized and confrontational environments on social media.

The time series analysis did not provide much evidence to support a Trump Effect on social media as defined by a temporal relationship between controversial comments and forms of mass affective rhetoric. As previously noted, the content analysis portion of this study could very easily have used the wrong words for analysis and, in turn, may have missed evidence of a Trump Effect. However, it is still possible that elite cues in the form of controversial comments have a polarizing effect on the networks in which discussion occurs. Further, it is possible that the creation of such networks could facilitate confrontation more effectively than others. If so, evidence of a *type* of Trump Effect may still exist. For the second section of this analysis, I used network analysis metrics to gain a better understanding regarding the nature and structure of the discrete networks that form in response to controversial comments made by political elites and, more importantly, facilitate potentially deleterious sentiment in individuals' responses. Additionally, I used hashtag analysis and network visualizations to examine the extent to which competing narratives clash in the wake of Donald Trump's most controversial comments.

Network Analysis

[Insert Table 1.3]

Just as it was necessary to limit the number of Donald Trump's controversial statements for use in the time series analysis, it was also necessary to limit these statements even further in order to conduct a focused set of network analyses. As such, I narrowed these statements down to six instances where Donald Trump's controversial comments dealt with issues of gender, religion, race, and disabilities. I selected Hillary

Clinton's "alt-right speech" (8/25/2016) as a control measure to measure whether or not networks developed differently when Donald Trump made these controversial comments compared to when he was being attacked or criticized for making these comments.

Unfortunately, I experienced corruption in my data set for the day when the Access Hollywood tape was released (12/7/2016). As of the time of this writing, I was not yet able to repair the corrupted data so it was not able to be included in this study. Future research will include this data, as this comment created a major firestorm which nearly derailed Donald Trump's presidential campaign and led to widespread protests following his election.

[Insert Graph 2.1]

The first network test was designed to identify whether networks on certain days were more likely to facilitate confrontation than networks on other days. This test used average path length as a proportion of network diameter as a measure of homophily, as this can be an indicator of polarized communities. Such a measure is important, as discussion within polarized communities tends to be more homogenous in nature and, in turn, less confrontational. This test also used average community size as a measure of cluster intensity. Generally speaking, networks with higher average community sizes tend to have a larger number of people discussing similar issues or interests. Whereas average path length as a proportion of network diameter suggests polarized communities, average community size suggests the intensity of this polarization.

It should be stressed that these measures are relative to each other and lose their descriptive power if read as absolute independent values. Network analysis is often a process of comparison, where conclusions regarding the nature of one network gain

strength based upon the ability to compare that network's characteristics with another network's characteristics. Using such a comparative approach when analyzing this first network test, a network with a comparatively high average path and average community size is more likely to facilitate highly populated polarized communities than a network with a comparatively low average path and average community size. Likewise, a network with a comparatively low average path and high community size is more likely to facilitate highly populated connected communities, while a network with a comparatively high average path and low community size is more likely to facilitate diffuse and more sparsely inhabited communities.

This initial network analysis found that the control network on the day of Hillary Clinton's "alt right speech" was the most connected (or least polarized) network containing the second highest population density per community. These results suggest that a comparatively high amount of discussion between communities was occurring during Hillary Clinton's "alt right speech", and that these individual communities were more highly populated.

The network on the day of Donald Trump's "millions of people voted illegally" comment was the second most connected (or least polarized) network containing the highest population density per community. However, this network's level of polarization was significantly higher than the network on the day of Hillary Clinton's "alt right speech", placing it in the middle of the polarization pack with networks on the days of Donald Trump's comments about a disabled reporter, his accusation that Hillary Clinton was playing the "woman card", and his proposal for a temporary ban on Muslims. These latter three networks also ranked comparatively lower in average community size which,

in turn, suggested lower community population density. Last, networks that formed in the wake of Donald Trump's comments stating "Islam hates us" and "I love Hispanics!" easily ranked as the most polarized networks with the 2nd and 3rd lowest community population densities; communities of discussion in these networks were far less concentrated and far less connected than the networks that formed on other days.

When analyzing these results within the context of the existence of a Trump Effect, questions arise as to whether network polarization can be a "good" thing as well as a "bad" thing. For example, if the predominant language in a given network is inflammatory, aggressive, intemperate, or counterproductive to civil discourse, one could argue that network polarization would be a "good" thing; polarized networks are less connected and less efficient, thus moderating the effect of deleterious discourse. Conversely, tightly-knit (or less polarized) networks could be argued to be a "bad" thing, given the same conditions.

Answering such questions requires an understanding that measures of network polarization only indicate the structure of the networks facilitating discussion. They take on new meaning when we take into account the nature of the discussion occurring within these networks. Content analysis performed while testing the first hypothesis found that rates of aggression in tweets (11.11 per 1,000 words) mentioning Donald Trump reached their highest point (out of 548 days) on the day which he claimed "millions voted illegally" in the election. The second highest rate of aggression (10.92 per 1,000 words) was measured the day after these comments.

Putting these rates in perspective, this meant that more than 1% of all words in tweets mentioning Donald Trump on these days could be categorized as aggressive.

These two rates were significantly higher than the 3rd highest rate of aggression (7.94 per 1,000 words) and far higher than the average rate for all 548 days of observation (0.66 per 1,000 words). Given that the network on this date was relatively connected and populated by communities with relatively high population densities, it could be concluded that a *lack* of network polarization was somewhat of a "bad" thing. The network was facilitating efficient communication between large communities and the discussion was notably aggressive. Further, it provides compelling evidence that a form of a Trump Effect could, in fact, exist when observed in a network analysis frame of reference.

Interestingly, the content analysis performed while testing the first hypothesis found that rates of negativity (5.44) in tweets mentioning Donald Trump reached their 8th highest value (out of 548 days) on the day of Hillary Clinton's "alt-right speech". This rate was roughly twice as high as the average rate of negativity (2.64) for the entire 548 days of observation. As was the case with the network on the day of Donald Trump's "millions voted illegally" comment, it could be concluded that a *lack* of network polarization was a "bad" thing, in that the network was efficient in facilitating negativity between large communities.

The battery of network analyses strongly suggest that Donald Trump's "millions voted illegally" comment created a network structure that efficiently spread aggression, while Hillary Clinton's "alt right speech" created a network that efficiently spread negativity. These results provide evidence of structural network preconditions for a type of Trump Effect to exist.

Hashtag Analysis and Network Visualizations

Given that evidence was found of networks that were capable of efficiently facilitating aggression and negativity, a final set of tests were conducted to approximate the individuals to whom these sentiments were directed. Rather put, in order to determine the extent to which a Trump Effect exists, it was necessary to determine at whom this aggression and negativity was being directed. To this end, I performed a hashtag analysis on tweets mentioning Donald Trump on the day of his "millions of immigrants voted illegally" comment and on the day of Hillary Clinton's "alt right speech".

[Insert Graph 2.2]

[Insert Graph 2.3]

[Insert Table 1.4]

The hashtag analysis for the network of tweets mentioning Donald Trump on the day he claimed "millions voted illegally" contained several hashtags that could be categorized as "pro Trump" and several that could be categorized as "anti Trump". The "pro Trump" hashtags were identified as #maga (115), #gop (41), and #tcot (34). While it is true that there was significant resistance to Donald Trump amongst Republicans (#gop) and conservatives (#tcot, or "true conservatives on Twitter), given the context of the date for this network, I assumed that neither Republicans nor conservatives would more likely to support Hillary Clinton over Donald Trump in a dispute over the results of the presidential election. The "anti Trump" hashtags were identified as #recount2016

(85), #auditthevote (81), #amjoy (77), #recount16 (52), #notmypresident (43),

#votersuppression (32), #imstillwithher (28), and #resisttrump (22).

In the aggregate, "anti Trump" hashtags (420) far outnumbered "pro Trump"

hashtags (190). Further, this network featured by far the highest rate of aggression (11.11

per 1,000 words) in tweets mentioning Donald Trump out of the 548 days of observation.

Donald Trump's claim that "millions voted illegally" in the election produced a great

deal of aggression and, most importantly, this aggression was expressed in *opposition* to

Donald Trump, rather than in *support* of Donald Trump. This suggests that if a Trump

Effect exists in social media, it does not exist in a form where it is aggression directed

towards those who oppose Donald Trump, but rather it is aggression directed towards

Donald Trump. However, aggression in the aggregate does not necessarily mean this

aggression was influential. In order to determine the extent to which aggression

contributed to a confrontational environment, I employed a second test using network

visualizations.

[Insert Graph 2.4]

[Insert Graph 2.5]

[Insert Graph 2.6]

For these visualizations, tweets with "pro Trump" hashtags were highlighted in

red, while "anti Trump" hashtags were highlighted in blue. As these visualizations

demonstrate, "anti Trump" hashtags were much more successful at engaging the center of

discussion far more frequently than "pro Trump" hashtags and, as such, were more

influential. In sum, "anti Trump" sentiment was more frequent, more aggressive, and

more influential than "pro Trump" sentiment on the day Donald Trump alleged that

"millions voted illegally" in the election. Rather than resulting in a Trump Effect where

Donald Trump's controversial comments fostered aggression among his supporters, it

appears that these comments fostered aggression in his opponents.

[Insert Graph 2.7]

[Insert Graph 2.8]

[Insert Table 1.5]

An identical pair of network tests were performed on the "control" network

containing tweets mentioning Donald Trump on the day of Hillary Clinton's "alt right

speech". The "pro Trump" hashtags on this day were identified as #trump2016 (15),

#makeamericagreatagain (14), #maga (13), #latinosfortrump (8), #trumptrain (8),

#leadright (6), #tcot (6), #sickhillary (5), and #trumppence16 (5). The "anti Trump"

hashtags were identified as #imwithher (31), #nevertrump (24), #toxictrump (21),

#uniteblue (8), and #voteblue (6).

In the aggregate, "anti Trump" hashtags (90) slightly outnumbered "pro Trump"

hashtags (75). This network featured the 8th highest rate of negativity (5.44 per 1,000

words) in tweets mentioning Donald Trump out of 548 days of observation. Just as

Donald Trump's claim that "millions voted illegally" in the election produced a great

deal of aggression, Hillary Clinton's "alt right speech" produced a significant amount of

negativity. This negativity occurred in an atmosphere that was only slightly (20%) more

"anti Trump" than "pro Trump", rather than being decidedly more "anti Trump" as was

the case following Donald Trump's "millions voted illegally" comment. However,

negativity in the aggregate does not necessarily mean such sentiment was influential. In

order to determine the extent to which negativity contributed to a confrontational

environment, I employed an identical set of network visualizations to those performed on the network following Donald Trump's "millions voted illegally" comment.

[Insert Graph 2.9]

[Insert Graph 2.10]

[Insert Graph 2.11]

For these visualizations, tweets with "pro Trump" hashtags were highlighted in red, while "anti Trump" hashtags were highlighted in blue. These visualizations indicate that "anti Trump" hashtags were much more successful at engaging the center of discussion far more frequently than "pro Trump" hashtags and, as such, were more influential. As was the case with the "millions voted illegally" network, "anti Trump" sentiment in the "alt right speech" network was more frequent, more negative, and more influential than "pro Trump" sentiment. Rather than resulting in a Trump Effect where Donald Trump's controversial comments fostered negativity among his supporters, it appears that these comments fostered negativity in his opponents.

Conclusions and Discussion

In sum, this chapter did not find compelling evidence to support the existence of a Trump Effect in social media. After using time series analysis spanning 548 individual days of observation, the findings suggested controversial comments made by Donald Trump during this time frame did not align temporally with increases in aggression, anxiety, or negativity in the resulting Twitter discussion about Donald Trump. When applying a network analysis approach, evidence was found of a phenomenon similar to the alleged Trump Effect; however, this evidence seemed to point more to a

"Clinton/Trump Effect". Rather put, the network analysis suggested that controversial comments and expressions of elite polarization do have the effect of inciting aggression and negativity in the general public. As such, the findings provide compelling evidence of a causal link between elite polarization and mass affective polarization, which should provide for ample opportunities for future research.

Questions surrounding the existence of a "Trump Effect" are, at the very least, based in noble intentions and grounded in the goals of reducing incivility, anger, and resentment towards others that can lead to very negative real world consequences. For example, we know that the problem of bullying in schools is a serious issue that often leads to a type of destruction to the innocence of youth which can have deep impacts on a child's development, both immediately as well as into the future. There is also evidence that adults are influenced by inflammatory rhetoric designed to target the lesser advantaged or the more vulnerable amongst us. If it is true that our leaders are engaging in actions that lead to an increase in such predatory behavior, then it is incumbent upon our civil society to identify, condemn, and seek corrections to such actions.

At the same time, care must be taken to confirm such a causal relationship exists before even the first moral sanction begins. The "Trump Effect" first gained notoriety following a survey conducted by an organization who openly admitted to its lack of scientific rigor. This does not negate the good intentions of such an effort, nor does it diminish the need to question whether such an effect exists. It does, however, highlight the need to maintain a pragmatic and scientific mindset when investigating questions with such important implications. This is especially true considering discussions about and references to the "Trump Effect" often blurred the line between moral imperative and a

strategic political tool. Maintaining a pragmatic and scientific mindset can indeed be a challenge when one attempts to maintain objective neutrality while analyzing aspects of one of the most contentious and nasty presidential elections in modern history. Regardless, this research has been conducted in manner that placed such an approach as the highest priority.

The goal of this chapter has been to test for a Trump Effect in a small slice of American society: discussions about Donald Trump on Twitter. It should be stressed that this slice of American society was not representative of the broader American citizenry, nor were the behaviors engaged in this environment representative of the wide range of behaviors human beings in which human beings are capable of engaging. Further, this research applied a unique mixed methods approach using empirical tests that may benefit from further methodological refinement.

The findings presented in this chapter also raised the possibility that forms of mass polarization – whether ideological or affective in nature – may not always necessarily be a "bad" thing. For example, in networks where there are high levels of aggression or negativity, perhaps it is more desirable for these discussions to be confined within clusters of likeminded individuals. In such a scenario, the likelihood of open and aggressive confrontation with others is less likely, leaving misery to enjoy company. Such questions beg for additional research into some of the potential network effects highlighted in this chapter.

Table 1.1: Selections from Donald Trump's Most Controversial Statements as Presidential Candidate and President Elect

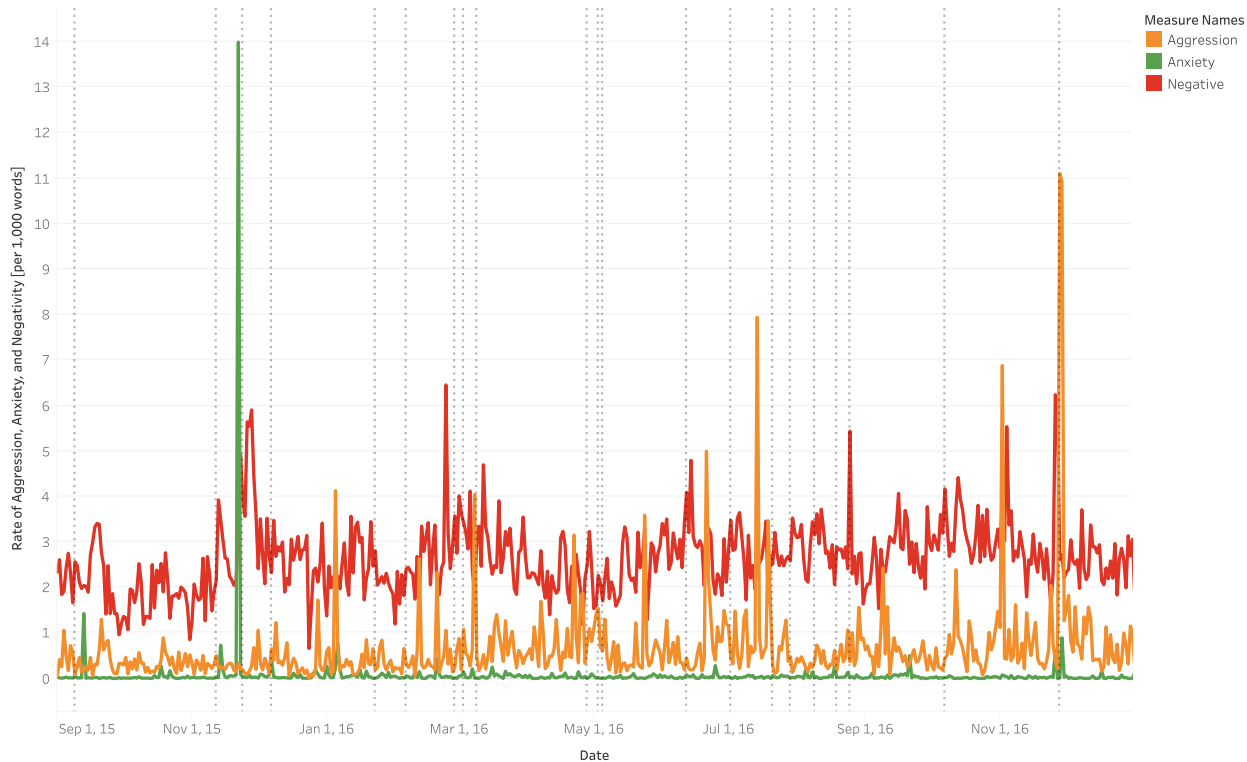| Date	Comment
9/9/15	Rolling Stone interview, Carly Fiorina's face
11/12/15	"How stupid are the people of Iowa?"
11/24/15	Disabled Reporter
12/7/15	Temporary Muslim Immigration Ban
1/23/16	"I could shoot someone on 5th Avenue"
2/6/16	Bring back things a "hell of a lot worse" than waterboarding
2/28/16	David Duke
3/3/16	Size of his "something else" (Rubio and hands spat)
3/9/16	"I think Islam hates us."
4/28/16	Clinton playing the "woman card"
5/3/16	Ted Cruz's Dad and JFK Assassination
5/5/16	"I love Hispanics!"
6/12/16	"appreciate the congrats for being right on radical Islamic terrorism"
7/2/16	The "star" tweet
7/21/16	RNC Speech
7/29/16	Khan
8/9/16	"Second Amendment people"
8/19/16	to black voters: "What do you have to lose?"
8/25/16	Hillary's "alt-right" speech
10/7/16	Access Hollywood tape
11/28/16	Millions of people voted illegally

Table 1.2: Emotion and Secondary Process Categories, Regressive Imagery Dictionary

SECONDARY PROCESS	
Abstraction	Know, may, thought
Social Behavior	Say, tell, call
Instrumental Behavior	Make, find, work
Restraint	Must, stop, bind
Order	Simple, measure, array
Temporal References	When, now, then
Moral Imperative	Should, right, virtue
EMOTIONS	
Positive Affect	Cheerful, enjoy, fun
Anxiety	Afraid, fear, phobic
Sadness	Depression, dissatisfied, lonely
Affection	Affectionate, marriage, sweetheart
Aggression	Angry, harsh, sarcasm
Expressive Behavior	Art, dance, sing
Glory	Admirable, hero, royal

Source: Provalis Research, https://provalisresearch.com/products/content-analysis-software/wordstat-dictionary/regressive-imagery-dictionary/

Graph 1.1: Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
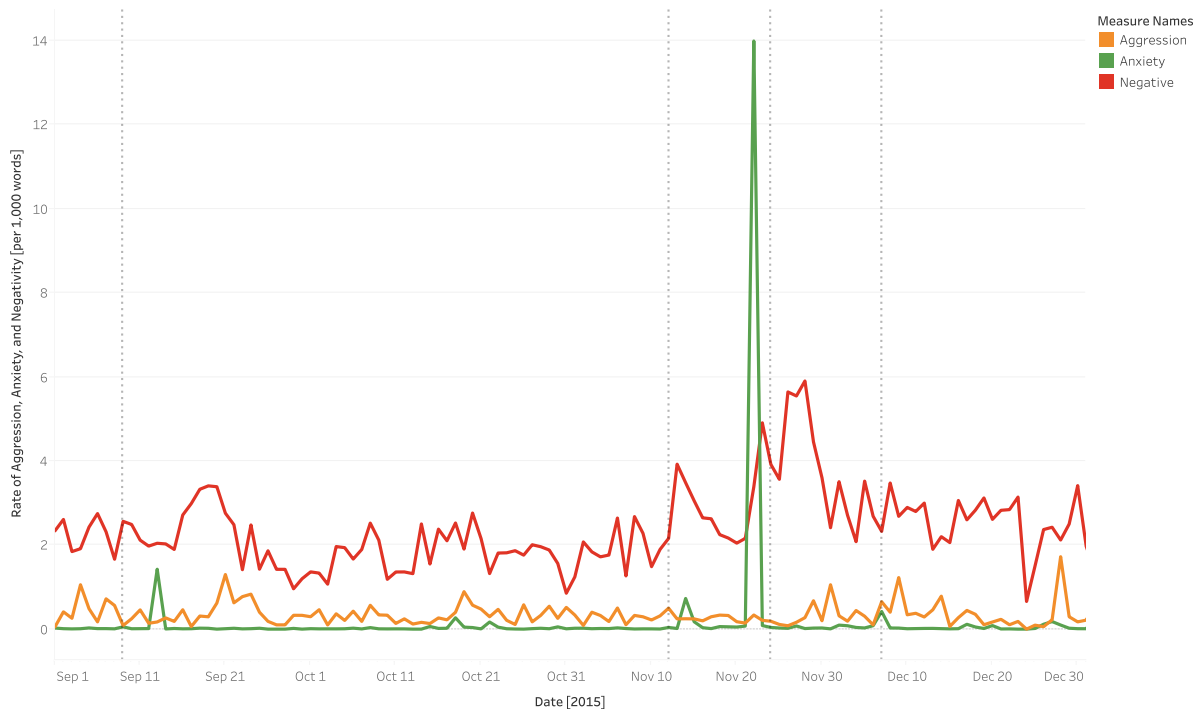9/1/2015 – 3/1/2017

Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump

Source: Twitter (9/1/2015 - 3/1/2017)
9/9/15: Carly Fiorina's face, 11/12/15: "How stupid are the people of Iowa?", 11/24/15: Disabled reporter, 12/7/15: Temporary Muslim Immigration Ban, 1/23/16: "I could shoot someone on 5th Avenue", 2/6/16: Bring back "hell of a lot worse" than waterboarding, 2/28/16: David Duke, 3/3/16: Rubio's hands, 3/9/16: "I think Islam hates us", 4/28/16: Clinton playing "woman card", 5/3/16: Ted Cruz's father, 5/5/16: "I love Hispanics!", 6/12/16: "appreciate the congrats for being right", 7/2/16: The "star" re-tweet, 7/21/16: RNC Speech, 7/29/16: Khan, 8/9/16: "Second Amendment people", 8/19/16: to black voters, "What do you have to lose?", 8/25/16: Clinton's "alt-right" speech, 10/7/16: Access Hollywood tape, 11/28/16: "Millions of people" voted illegally

Graph 1.2: Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
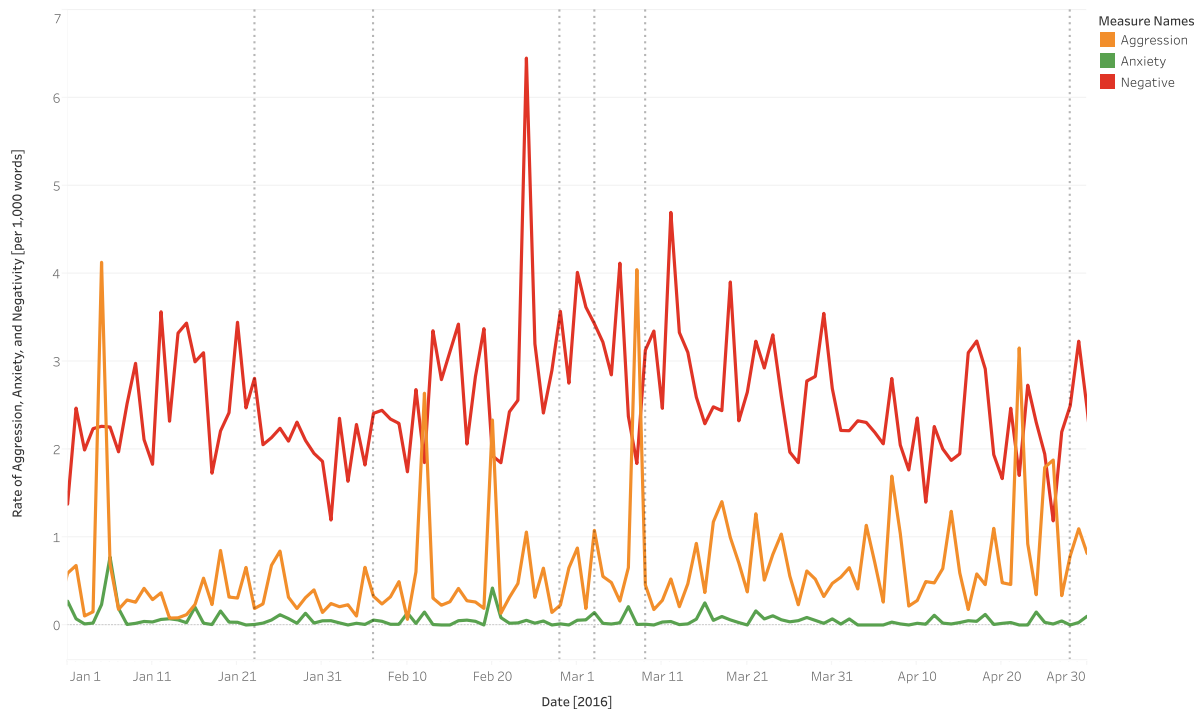9/1/2015 – 12/31/2015

Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
9/1/2015 - 12/31/2015

Source: Twitter (9/1/2015 - 3/1/2017)
9/9/15: Carly Fiorina's face, 11/12/15: "How stupid are the people of Iowa?", 11/24/15: Disabled reporter, 12/7/15: Temporary Muslim Immigration Ban, 1/23/16: "I could shoot someone on 5th Avenue", 2/6/16: Bring back "hell of a lot worse" than waterboarding, 2/28/16: David Duke, 3/3/16: Rubio's hands, 3/9/16: "I think Islam hates us", 4/28/16: Clinton playing "woman card", 5/3/16: Ted Cruz's father, 5/5/16: "I love Hispanics!", 6/12/16: "appreciate the congrats for being right", 7/2/16: The "star" re-tweet, 7/21/16: RNC Speech, 7/29/16: Khan, 8/9/16: "Second Amendment people", 8/19/16: to black voters, "What do you have to lose?", 8/25/16: Clinton's "alt-right" speech, 10/7/16: Access Hollywood tape, 11/28/16: "Millions of people" voted illegally

Graph 1.3: Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
1/1/2016 – 4/30/2016

Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
1/1/2016 - 4/31/2016

Graph 1.4: Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
5/1/2016 – 8/31/2016

Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
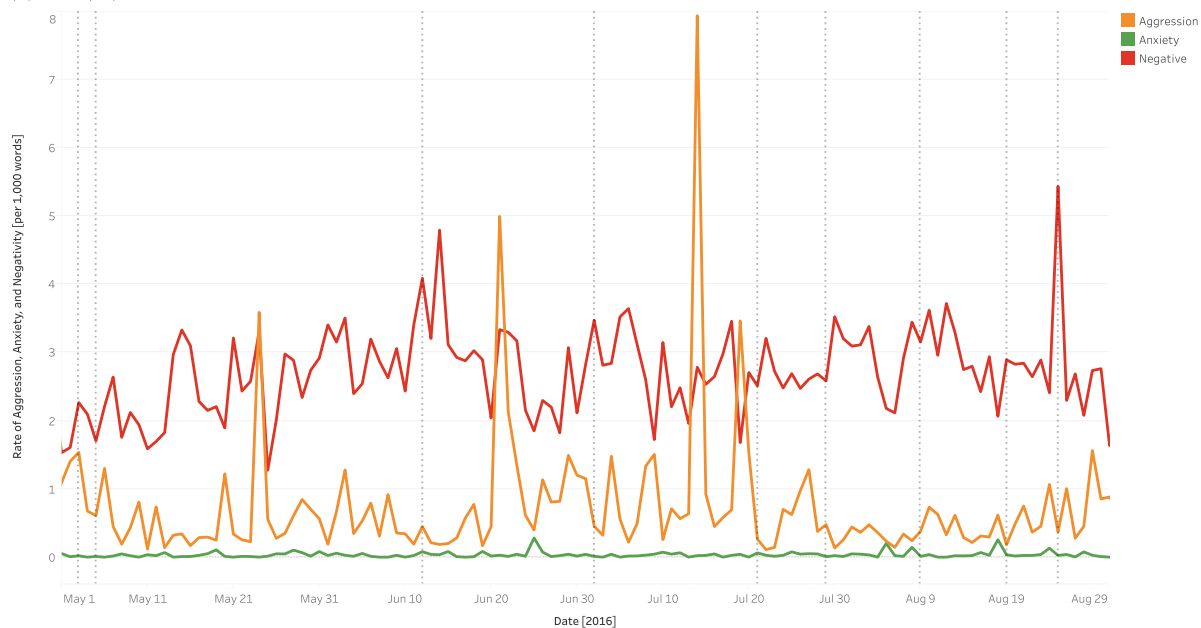5/1/2016 - 8/31/2016

Source: Twitter (9/1/2015 - 3/1/2017)
9/9/15: Carly Fiorina's face, 11/12/15: "How stupid are the people of Iowa?", 11/24/15: Disabled reporter, 12/7/15: Temporary Muslim Immigration Ban, 1/23/16: "I could shoot someone on 5th Avenue", 2/6/16: Bring back "hell of a lot worse" than waterboarding, 2/28/16: David Duke, 3/3/16: Rubio's hands, 3/9/16: "I think Islam hates us", 4/28/16: Clinton playing "woman card", 5/3/16: Ted Cruz's father, 5/5/16: "I love Hispanics!", 6/12/16: "appreciate the congrats for being right", 7/2/16: The "star" re-tweet, 7/21/16: RNC Speech, 7/29/16: Khan, 8/9/16: "Second Amendment people", 8/19/16: to black voters, "What do you have to lose?", 8/25/16: Clinton's "alt-right" speech, 10/7/16: Access Hollywood tape, 11/28/16: "Millions of people" voted illegally

Graph 1.5: Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
9/1/2016 – 12/31/2016

Aggression, Anxiety, and Negativity in Tweets Mentioning Donald Trump
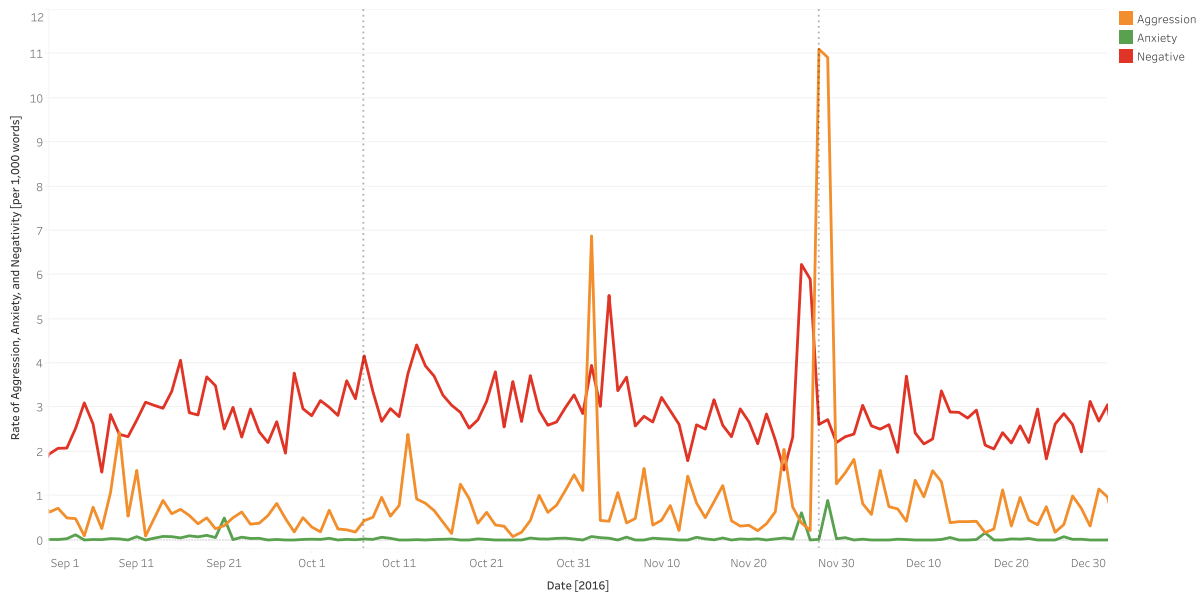9/1/2016 - 12/31/2016

Source: Twitter (9/1/2015 - 3/1/2017)
9/9/15: Carly Fiorina's face, 11/12/15: "How stupid are the people of Iowa?", 11/24/15: Disabled reporter, 12/7/15: Temporary Muslim Immigration Ban, 1/23/16: "I could shoot someone on 5th Avenue", 2/6/16: Bring back "hell of a lot worse" than waterboarding, 2/28/16: David Duke, 3/3/16: Rubio's hands, 3/9/16: "I think Islam hates us", 4/28/16: Clinton playing "woman card", 5/3/16: Ted Cruz's father, 5/5/16: "I love Hispanics!", 6/12/16: "appreciate the congrats for being right", 7/2/16: The "star" re-tweet, 7/21/16: RNC Speech, 7/29/16: Khan, 8/9/16: "Second Amendment people", 8/19/16: to black voters, "What do you have to lose?", 8/25/16: Clinton's "alt-right" speech, 10/7/16: Access Hollywood tape, 11/28/16: "Millions of people" voted illegally

Table 1.3: Donald Trump's Controversial Statements as Presidential Candidate and President Elect - Issues of gender, religion, race, and disabilities

11/24/15	Disabled Reporter
12/7/15	Temporary Muslim Immigration Ban
3/9/16	"I think Islam hates us."
4/28/16	Clinton playing the "woman card"
5/5/16	"I love Hispanics!"
8/25/16	Hillary's "alt-right" speech (control measure)
11/28/16	Millions of people voted illegally

Graph 2.1: Controversial Statements by Donald Trump involving Race, Gender, Religion, and Disabilities – Network Efficiency and Intensity
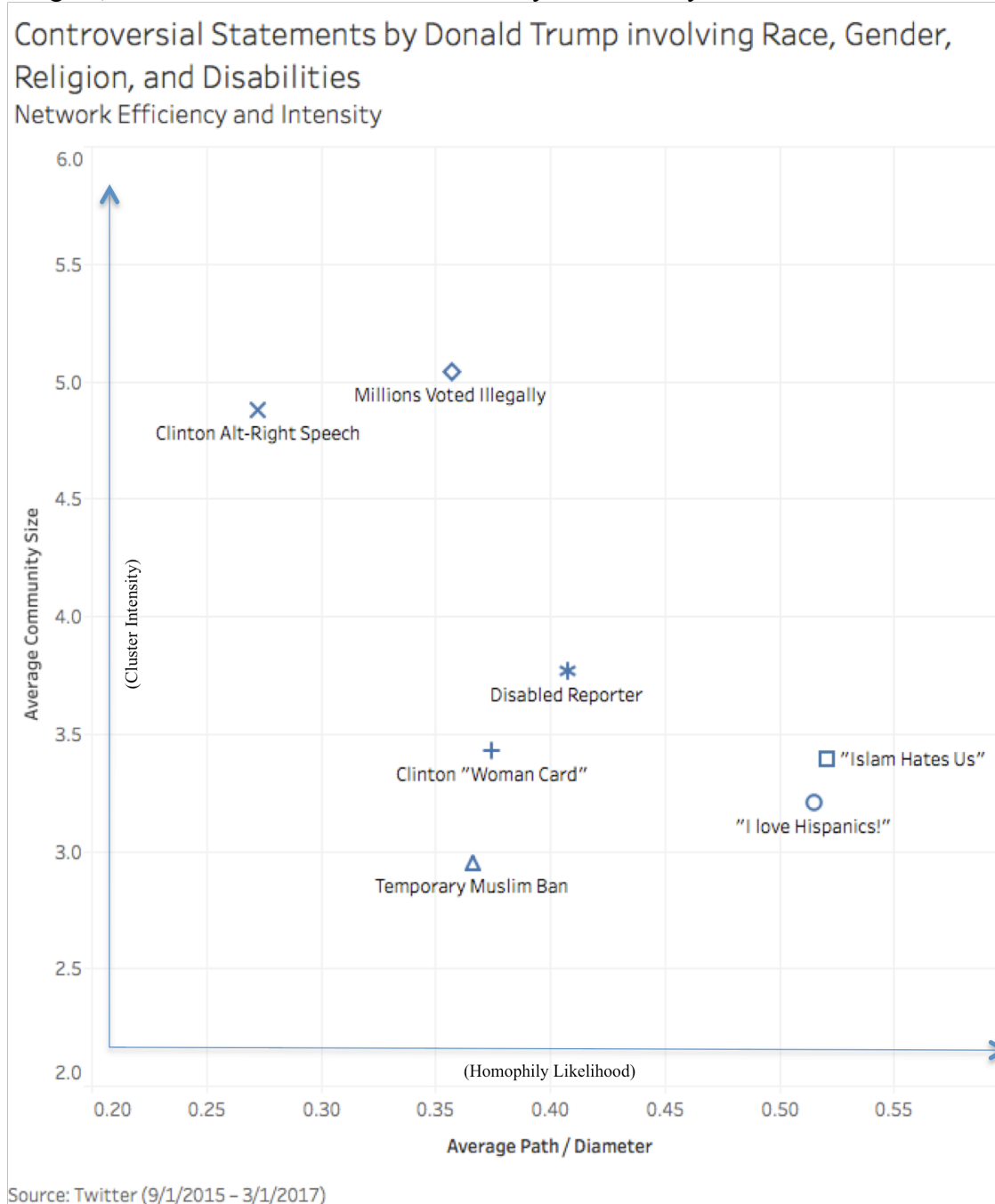
Graph 2.2: Hashtag Analysis of Tweets Mentioning Donald Trump on 11/28/2016 – "Millions of people voted illegally"

Top Hashtags in Tweets Mentioning Donald Trump
"Millions Voted Illegally"

Word	
trump	222
cuba	173
breaking	162
election2016	138
maga	115
recount2016	85
auditthevote	81
amjoy	77
news	58
donaldtrump	54
recount16	52
notmypresident	43
gop	41
tcot	34
eeuu	33
nc	32
votersuppression	32
donaltrump	31
guncontrol	31
climate	30
time	30
imstillwithher	28
noway	28
crown	27
manypeoplearesaying	26
resisttrump	22
climatemayors	21
parisagreement	20
trumpconflicts	20
whitehouse	20
electoralcollege	19
fidelcastro	19
withoutconsequencesiwo..	19
actonclimate	18
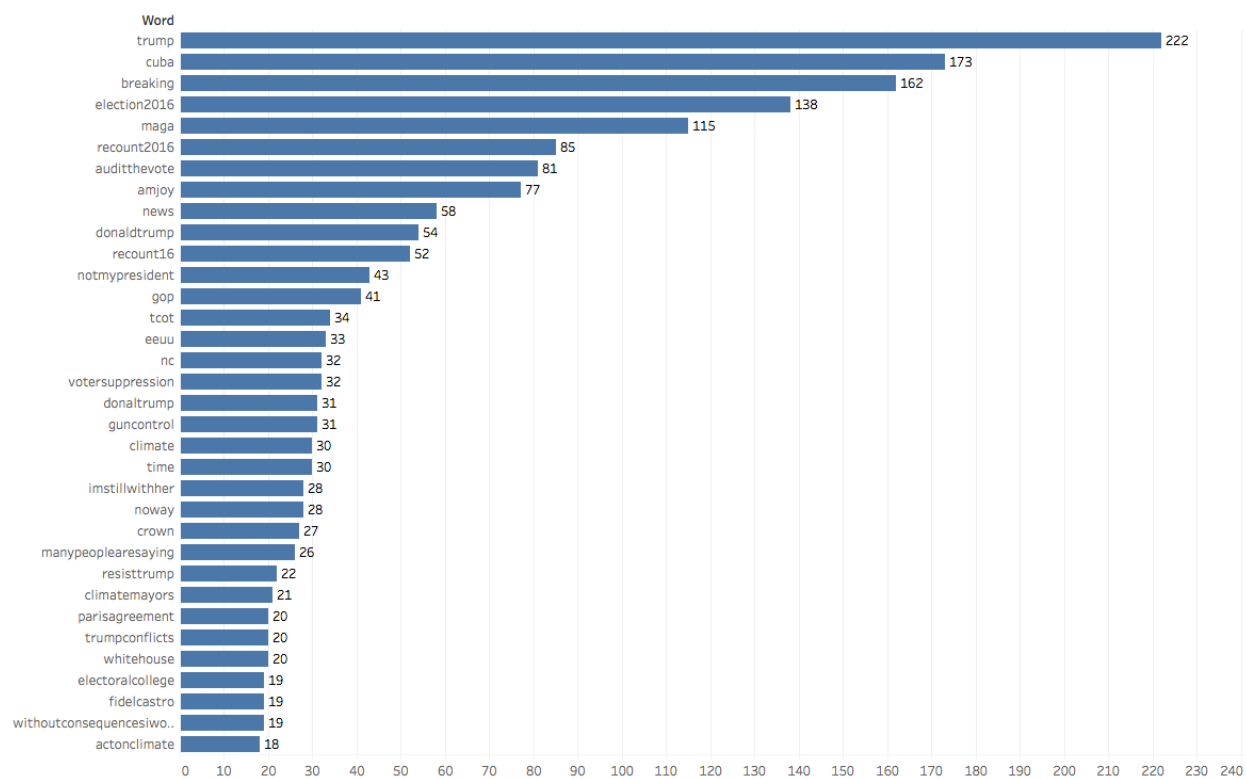
Source: Twitter (11/28/2016)

Graph 2.3: Hashtag Analysis of Tweets Mentioning Donald Trump on 11/28/2016 – "Millions of people voted illegally"

Top Hashtags in Tweets Mentioning Donald Trump
"Millions Voted Illegally"
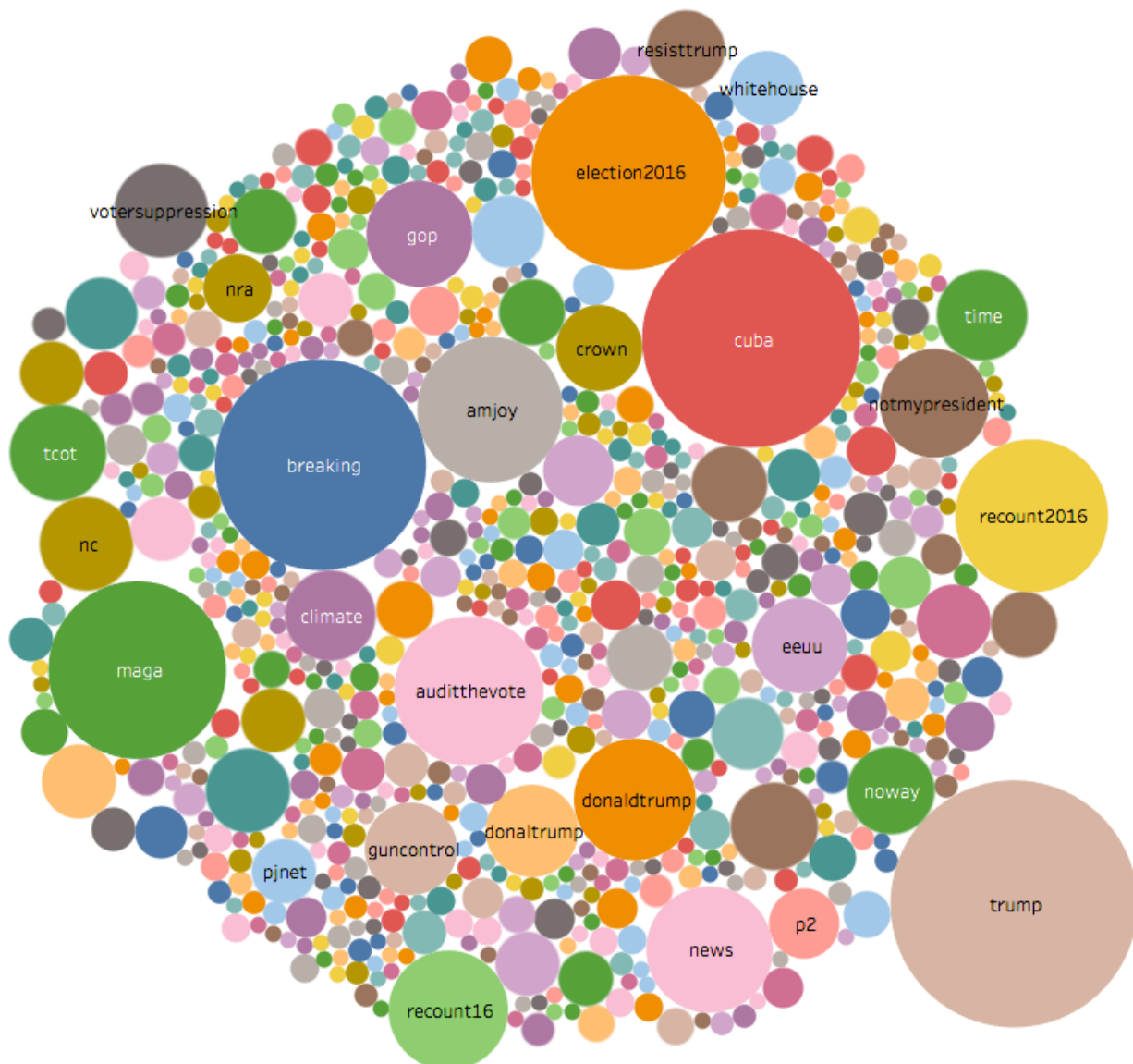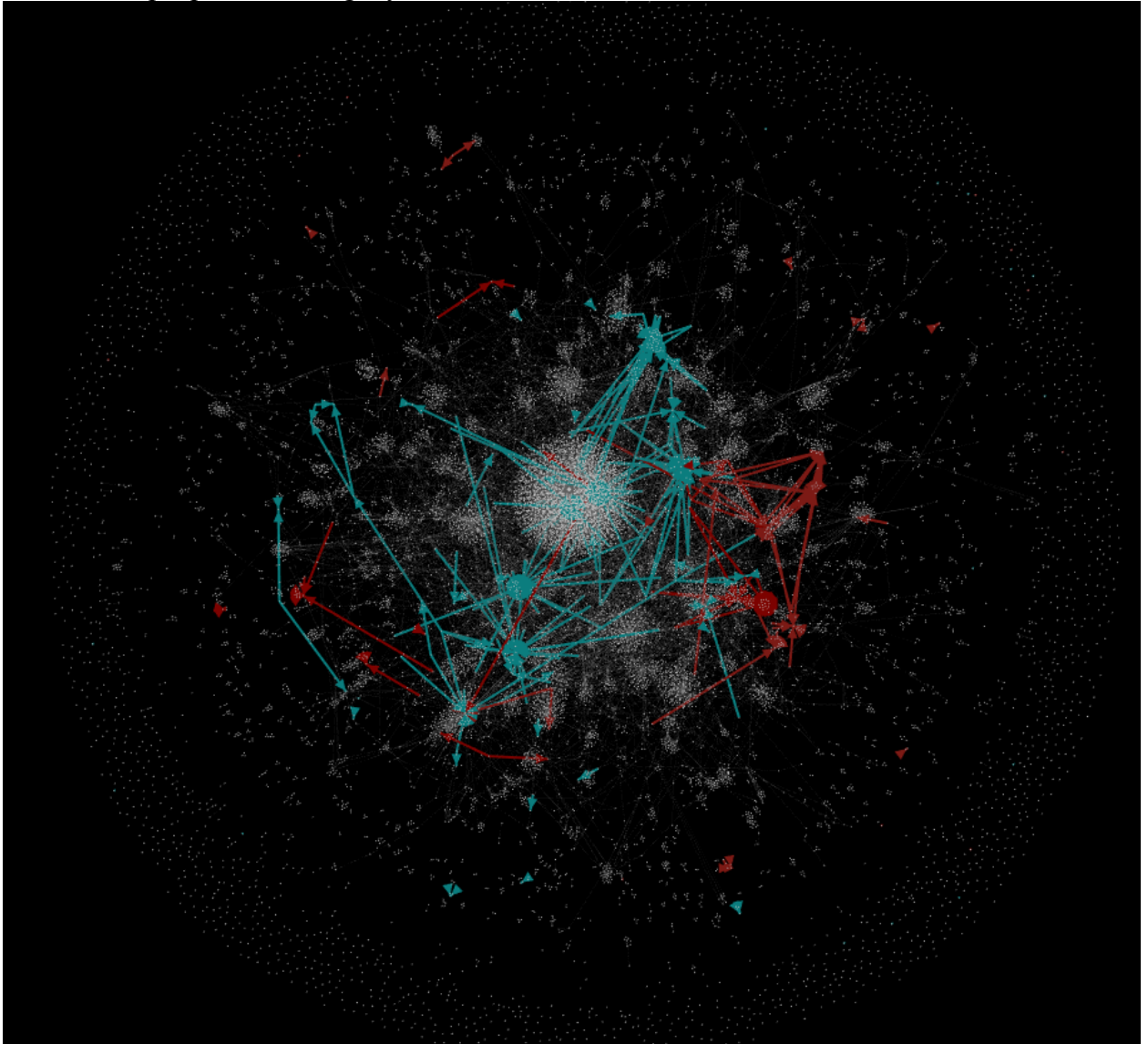
Source: Twitter (11/28/2016)

Table 1.4: "Pro Trump" and "Anti Trump" Hashtags in Tweets Mentioning Donald Trump on 11/28/2016 – "Millions of people voted illegally"

"Pro Trump"	"Anti Trump"
#maga	#recount2016
#gop	#auditthevote
#tcot	#recount16
	#amjoy
	#notmypresident
	#votersuppression
	#imstillwithher
	#resisttrump

Graph 2.4: Network Visualization of Tweets Mentioning Donald Trump on 11/28/2016
–
"Millions of people voted illegally"

Source: Twitter (11/28/2016) and Gephi 0.8.1
Note: "Pro Trump" Hashtags in Blue, "Anti Trump" Hashtags in Red
Note: Supplemental short 3d video fly-through of network at: http://bit.ly/2ppC4Hq

Graph 2.5: Network Visualization of Tweets Mentioning Donald Trump on 11/28/2016 – "Millions of people voted illegally"

Source: Twitter (11/28/2016) and Gephi 0.8.1
Note: "Pro Trump" Hashtags in Blue, "Anti Trump" Hashtags in Red
Note: Supplemental short 3d video fly-through of network at: http://bit.ly/2ppC4Hq

Graph 2.6: Network Visualization of Tweets Mentioning Donald Trump on 11/28/2016 –
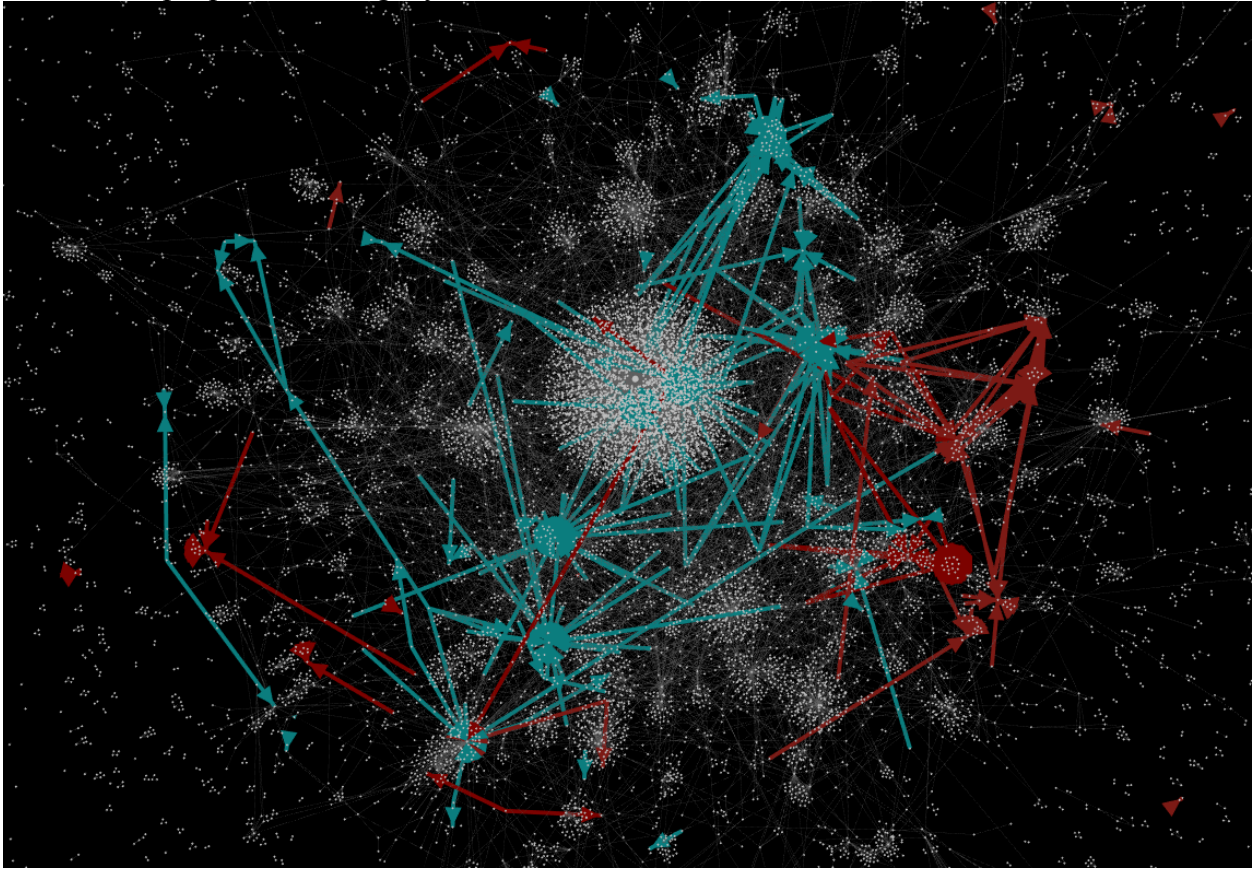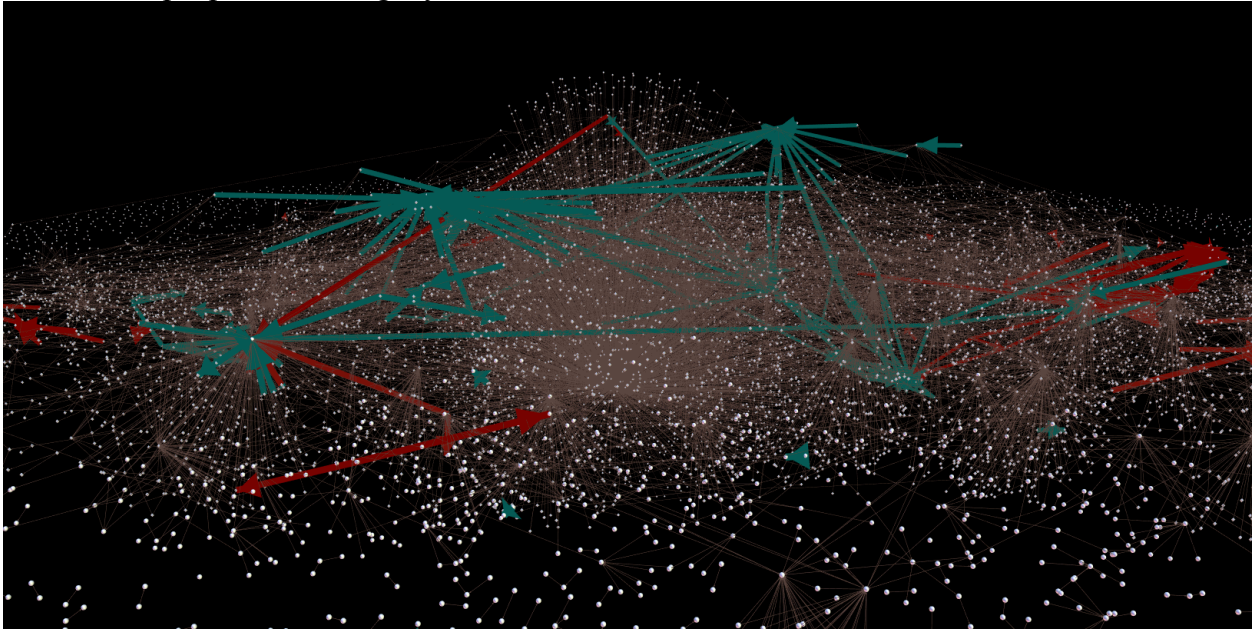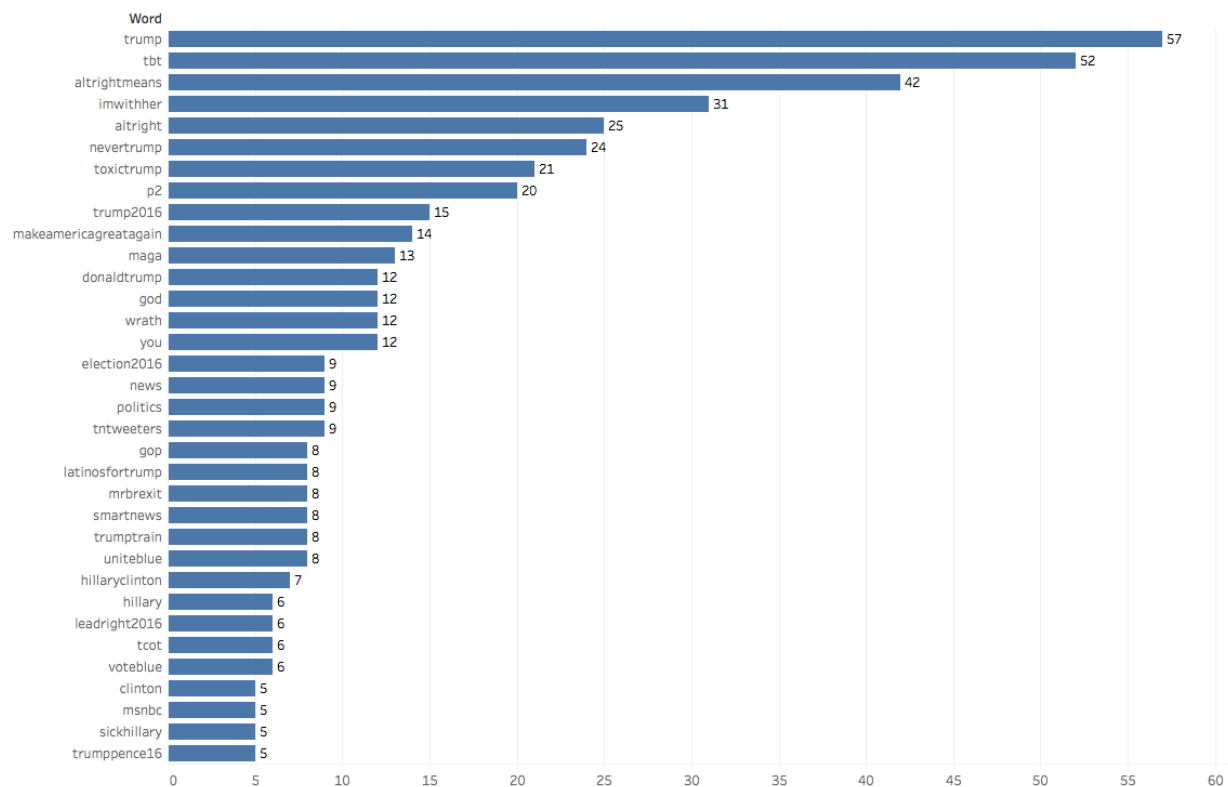
"Millions of people voted illegally"

Source: Twitter (11/28/2016) and Gephi 0.8.1
Note: "Pro Trump" Hashtags in Blue, "Anti Trump" Hashtags in Red
Note: Supplemental short 3d video fly-through of network at: http://bit.ly/2ppC4Hq

Graph 2.7: Hashtag Analysis of Tweets Mentioning Donald Trump on 8/25/2016 – Hillary Clinton's "alt right speech"

Top Hashtags in Tweets Mentioning Donald Trump
"Clinton Alt-Right Speech"

Word	Count
trump	57
tbt	52
altrightmeans	42
imwithher	31
altright	25
nevertrump	24
toxictrump	21
p2	20
trump2016	15
makeamericagreatagain	14
maga	13
donaldtrump	12
god	12
wrath	12
you	12
election2016	9
news	9
politics	9
tntweeters	9
gop	8
latinosfortrump	8
mrbrexit	8
smartnews	8
trumptrain	8
uniteblue	8
hillaryclinton	7
hillary	6
leadright2016	6
tcot	6
voteblue	6
clinton	5
msnbc	5
sickhillary	5
trumppence16	5

Source: Twitter (8/25/2016)

Graph 2.8: Hashtag Analysis of Tweets Mentioning Donald Trump on 8/25/2016 –
Hillary Clinton's "alt right speech"

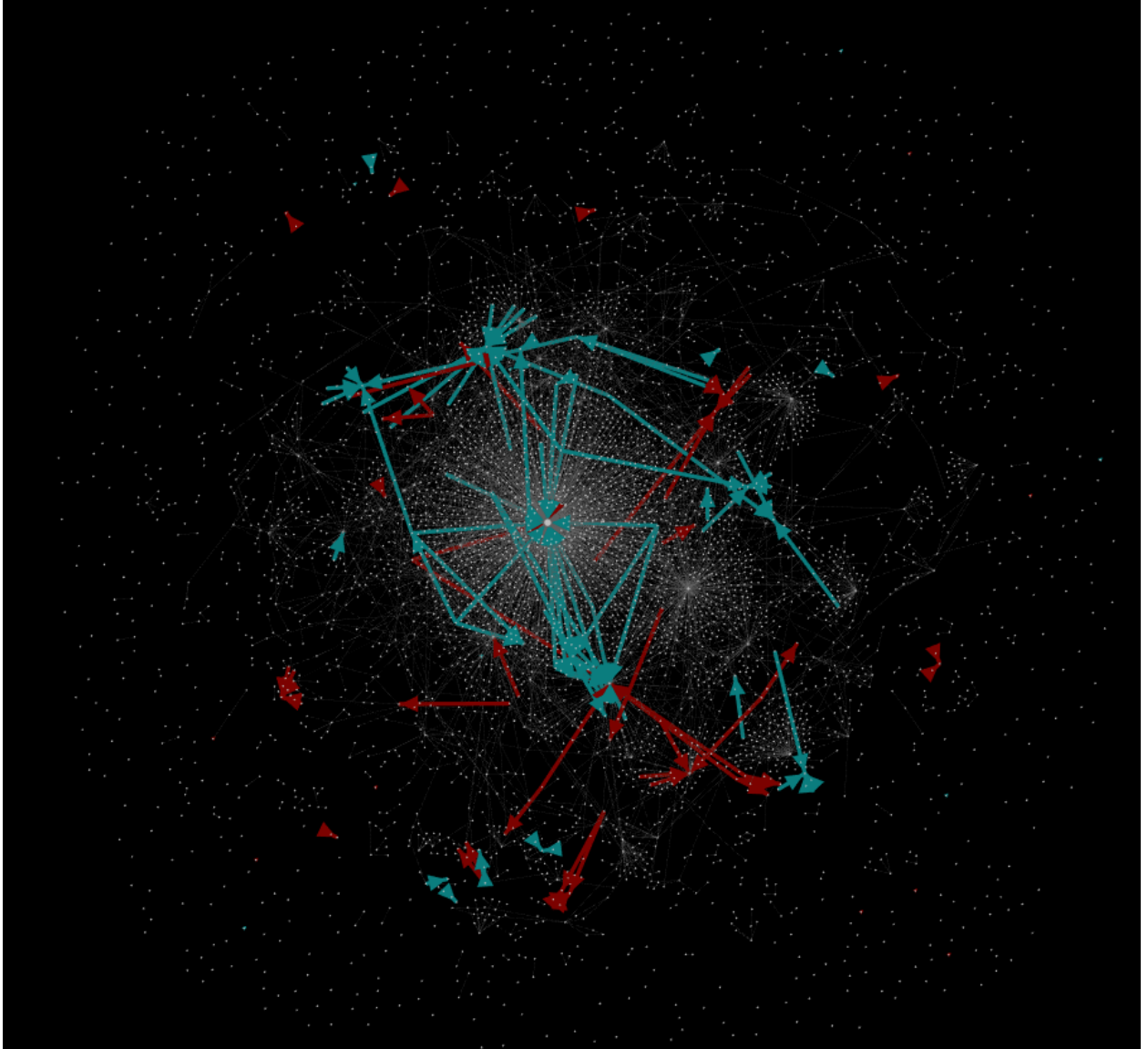Top Hashtags in Tweets Mentioning Donald Trump
"Clinton Alt-Right Speech"

Source: Twitter (8/25/2016)

Table 1.5: "Pro Trump" and "Anti Trump" Hashtags in Tweets Mentioning Donald
Trump on 8/25/2016 – Hillary Clinton's "alt right speech"

"Pro Trump"	"Anti Trump"
#trump2016	#imwithher
#makeamericagreatagain	#nevertrump
#maga	#toxictrump
#latinosfortrump	#uniteblue
#trumptrain	#voteblue
#leadright	
#tcot	
#sickhillary	
#trumppence16	

Graph 2.9: Network Visualization of Tweets Mentioning Donald Trump on 8/25/2016 – Hillary Clinton's "alt right speech"

Source: Twitter (8/25/2016)
Note: "Pro Trump" Hashtags in Blue, "Anti Trump" Hashtags in Red
Note: Supplemental short 3d video fly-through of network at: http://bit.ly/2o3xD4I

Graph 2.10: Network Visualization of Tweets Mentioning Donald Trump on 8/25/2016
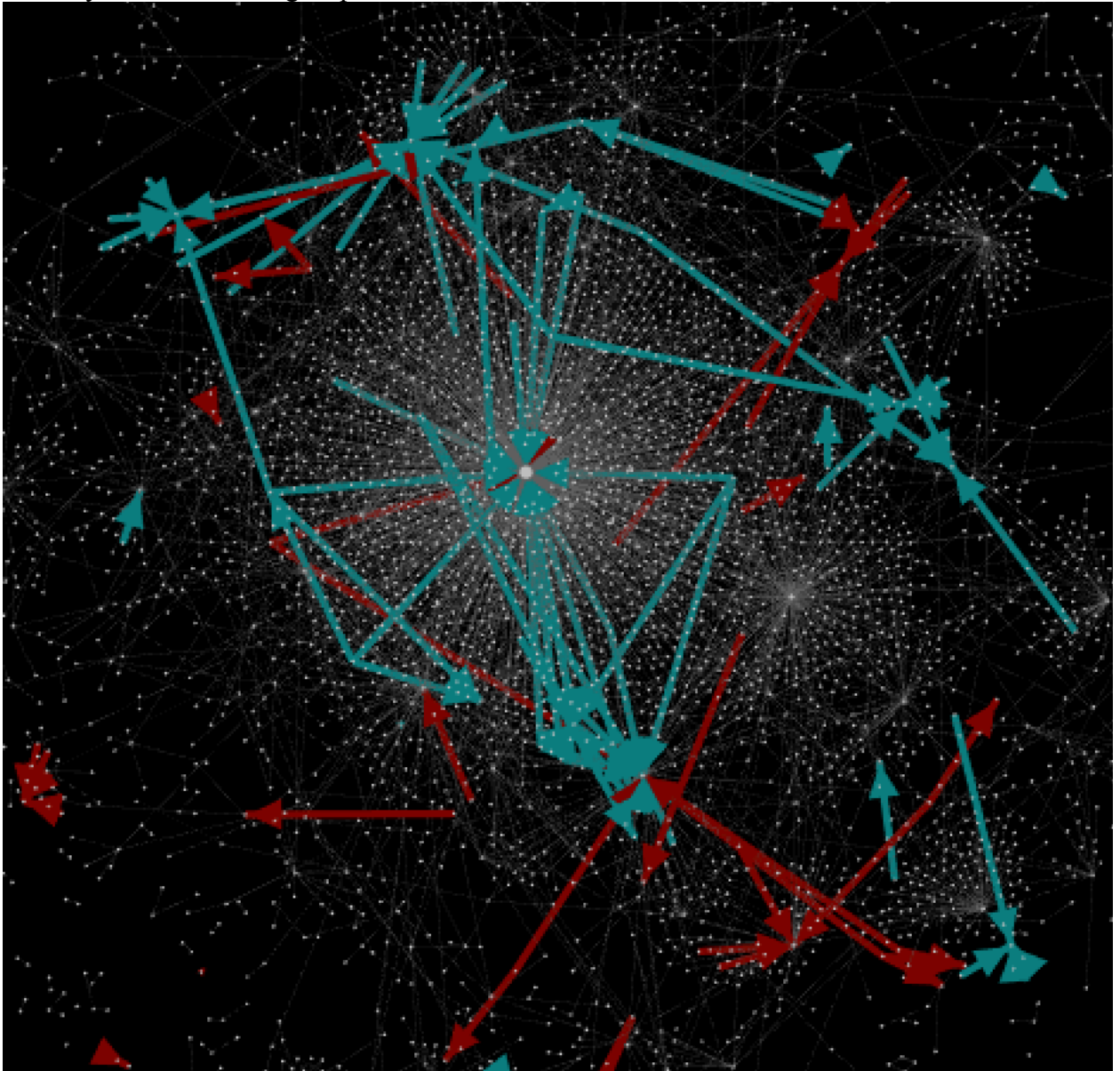– Hillary Clinton's "alt right speech"

Source: Twitter (8/25/2016)
Note: "Pro Trump" Hashtags in Blue, "Anti Trump" Hashtags in Red
Note: Supplemental short 3d video fly-through of network at: http://bit.ly/2o3xD4I

Graph 2.11: Network Visualization of Tweets Mentioning Donald Trump on 8/25/2016 – Hillary Clinton's "alt right speech"
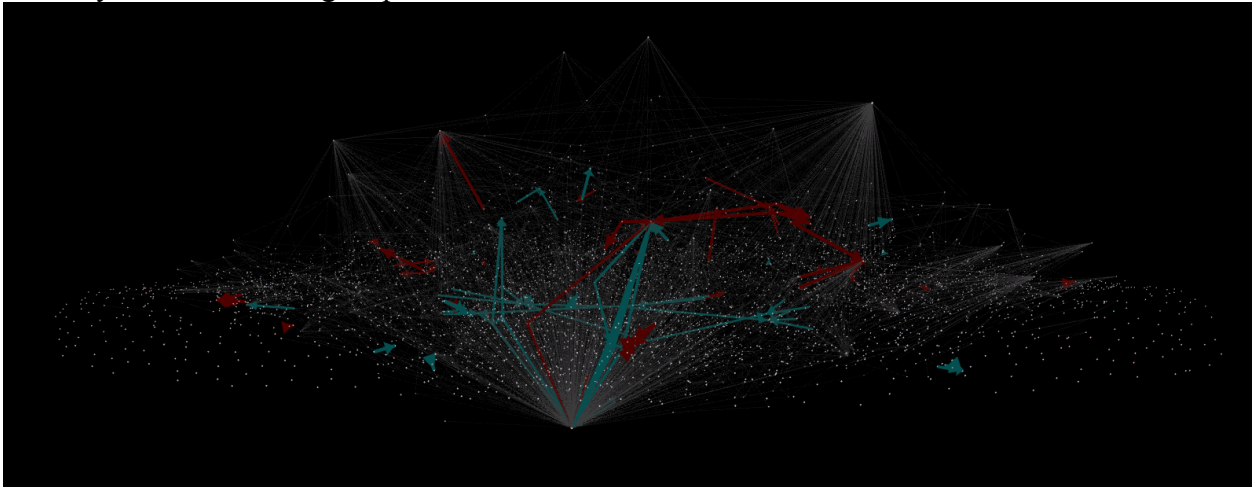

Source: Twitter (8/25/2016)
Note: "Pro Trump" Hashtags in Blue, "Anti Trump" Hashtags in Red
Note: Supplemental short 3d video fly-through of network at: http://bit.ly/2o3xD4I

Dissertation Chapter 7

Trolling Twitter: Conclusion

Eric C. Vorst

At the dawn of the Information Age in the early 1990's, the Internet was viewed in largely utopian frames as a powerful new tool that would allow people to share and connect with each other in ways previously unimaginable. This utopian perspective extended into the world of politics, where many believed the low cost and easy accessibility of the Internet would encourage people to engage in the democratic process more frequently. Given the availability of such a powerful tool, it was also assumed that citizens worldwide would seize upon the opportunity to organize collectively to promote causes important to them, while also organizing dissent in order to hold their leaders responsible.

Perhaps the most visible example of such organized engagement by citizens was seen in the "Arab Spring" protests that spread throughout the Middle East and Northern Africa from late 2010 through early 2011. It is difficult to deny the extent to which the Internet – and especially social media – provided citizens with the tools to organize massive protests that would likely have not been achievable otherwise. The end results were visibly striking, with images of massive public protests in multiple town squares across several middle-eastern nations being broadcast around the clock on television sets and computer screens around the world. To many, the ability for massive public demonstrations to rise from the grass-roots and demand democratization in the face of oppressive regimes was proof positive of the positive nature of social media.

At the same time, the Arab Spring also had a far less publicized downside. Just as social media equipped citizens with the power to organize coordinated opposition to ostensibly unjust regimes, social media also equipped these regimes with the power to identify, isolate, and retaliate against the political threats. The Arab Spring provides a

powerful example of the dual nature of social media and how it can be used as a strategic tool for achieving political goals, whether those goals are perceived to be good or bad for democratic ideals such as egalitarianism or the right to free political speech.

The dual nature of social media matters within the scope of democratic participation because it frames our understanding of affective rhetoric and the relationship between elite polarization and mass polarization in social media. Social media can be used to organize public dissent or to share information regarding important policies, just as it can be used to target political threats or to spread false and misleading information. Likewise, social media can be used as a vehicle for citizens to engage in rational and reasoned debate, just as it can be used to spread hatred, fear, anger, and aggression. One could argue social media is a reflection of human nature. However, rather than merely reflecting human nature, in many ways social media magnifies it. Such power is critical when considering the impact of social media on the democratic process, as this is a process that depends upon healthy participation.

It is broadly accepted that polarization amongst the political elites in the United States is extremely pronounced. The extent to which this has manifested itself in the form of mass polarization is less clear. In part, this clarity has been elusive due to various possible causes of mass polarization, whether they be due to political elite cues, sorting, psychological impulses, or something else. Regardless, mass polarization can be used by politicians in order to mobilize support and to win elections. Unfortunately, when it is time for those politicians to transition into elected officials, the politics of fear and loathing become incompatible in an environment that requires negotiation and compromise (Kimball et al. 2013).

As social media continues to evolve as a platform for individuals to gather, process, share, and debate political issues, its potential for affecting mass polarization also grows. Measuring how the frequency of potentially polarizing language on social media changes over time with relation to the occurrence of a political event is an important step in better understanding how elite cues influence mass polarization in different information environments. More importantly, measuring how social networks treat polarizing language provides valuable insight into whether increases in such language will have a net negative effect on constructive debate, or whether it will be pushed to the fringes where its audience will be smaller and less influential.

Future Research

This dissertation presents a number of opportunities for future research, both on the front end with data collection as well as the back end with hypothesis testing. First, future research could employ notable improvements in the data collection process. As noted in previous chapters, data for this dissertation was gathered using outdated computers and an almost-free alternative for gathering daily random samples of Twitter data within the last seven days. With more robust funding, higher quality data could be obtained directly from the historical Twitter archives. Such an option would allow researchers to search for specific terms within a clearly defined set of parameters, including specific time and date ranges, tweet volume per day, and so on. In turn, this would lead to more consistent data sets for use in time-series analysis, rather than requiring research to be conducted on data sets consisting of random samples varying from 1,000 to 30,000 tweets per day. Further, more robust funding would allow for

several more powerful computers, thus allowing researchers to conduct live-streaming of Twitter messages as they occur. Such data collection methods were tested during the writing of this dissertation and provided rich and descriptive data sets that would be invaluable to creating more complete data sets for future research.

Second, future research could build upon this dissertation by improving upon the classification of data, specifically in the area of content analysis. While unsupervised content analysis provides a powerful tool for making sense of incredibly large amounts of data, it has some shortcomings in that it is dependent upon the quality of the content classification dictionary. The first content classification dictionary used in this dissertation was the Lexicoder Semantic Dictionary, which is widely considered to be an accurate and dependable tool for measuring sentiment in political texts. The second classification dictionary (the Regressive Imagery Dictionary) used in this dissertation was not quite as well-known or well-used in the social sciences, although it was valuable in identifying specific types of affect in messages, such as anger, aggression, and fear. However, a modified version of this dictionary specially tailored for analyzing political texts would likely provide more accurate results. Further, future research could better define extremely uncivil language. This dissertation selected the "F Word" as an example of extremely uncivil language, although this word can sometimes be used to express joy, excitement, anger, fear, or many other emotions. Additionally, there are other extremely uncivil words that were not measured in this dissertation. Again, this provides a number of opportunities for future researchers to create and refine an original content classification dictionary.

Third, future research could build upon the explanatory power of the methodology employed in this dissertation by better identifying exactly what is happening inside the communities of discussion identified through network analysis. While the methods used in this dissertation identified areas where polarization was more likely, more could be done with respect to determining whether or not the polarization was involving mostly positive or mostly negative types of discussion. Further, more could be done in the area of identifying basic characteristics of the users involved in these discussions. For example, future research could be conducted on measuring Twitter account data such as the account age, number of followers, number of people followed, whether or not the account has a profile picture, the types of websites the account tends to share, the users the account follows, and so on. Such information would be invaluable for better describing the ideological leanings of users in the network, while also providing evidence of whether the user was a real person versus a bot.

Last, the findings in this dissertation could be refined by future research into what their implications are in the real world, by questioning what happens outside of Twitter when people put down their phones or walk away from computers and decide to either participate politically – or not. One way of potentially providing such answers could involve the use of a second wave survey of members of the networks, using a traditional question-and-answer format. An immediate benefit of a second wave survey would be that it would instantly filter out bots. A more valuable benefit would be that it would provide researchers with a clearer picture of how behavior in social networks (and, possibly, polarized social networks) translates into political participation. Additionally, such data would be quite useful in describing differences in Twitter users when compared

to the broader population. Perhaps most importantly, a second wave survey would allow the more novel approach of network analysis to be supplemented by more traditional (and better known) methodologies like regression analysis.

Implications

A sense of urgency is needed in developing a better understanding of human behavior in networked environments. This urgency is warranted given the rapid pace at which everyday objects are being merged into networked environments. Commonly referred to as the "Internet of Things", this phenomenon raises a host of important questions with significant political implications regarding privacy, individual freedoms, and the proper role of government – not to mention the deeper philosophical questions regarding the ideal relationship between humans and technology.

While estimates vary, the size of the Internet of Things is expected to grow from approximately 6 billion devices in 2016 to as many as 50 billion devices in 2020. Understandably, the worlds of technology and business are hard at work to ensure the Internet of Things is implemented in a manner that is timely, efficient, and profitable. At the same time, it is essential for social scientists to keep pace with these rapid developments, especially with respect to how the Internet of Things will undoubtedly impact the relationship between citizens, business, and government. These areas have serious normative implications in a democratic republic with citizens who expect openness, transparency, and responsiveness in their institutions, while simultaneously valuing freedom, privacy, and the power of participation. Further, such rapid changes in humans' relationships with technology begs new questions with respect to government's

role in stimulating, moderating, and regulating these relationships in order to foster a civic culture that encourages participation, promotes responsible citizenship, and protects individuals' privacy rights.

Conclusions

Does elite polarization contribute to mass polarization in social media? This dissertation has presented strong evidence of a causal relationship between the two. First, there is a positive relationship between the temporal proximity to certain types of political events and a corresponding rise in affective rhetoric in social media, as well as a corresponding rise in certain types of extremely uncivil language. This evidence was presented within the scope of the 2016 Republican Presidential Primary, and suggests that elite polarization expressed through affective rhetoric and other cues can elicit similar rises in affect among the mass public. However, despite rises in affective rhetoric in the aggregate, social media networks exhibit the ability to isolate extremely unproductive language, thus limiting its reach and impact.

Second, there is positive relationship between specific types of emotional responses in the mass public when exposed to elite affective rhetoric. This evidence was presented within the scope of a live broadcast of the first 2016 Presidential debate between Hillary Clinton and Donald Trump, and suggests that the types of emotional responses in the mass public (such as anger, aggression, or fear) are dependent upon the context of the elite cue (such as controversy, contempt, or confrontation). This evidence is especially useful because it was obtained in a manner designed to capture the instant

effects of elite polarization on the mass public, thus providing a glimpse into people's base reactions before they have had time to digest, contemplate, and reflect.

Third, this dissertation used an objective and scientific approach to test for the existence of a so-called "Trump Effect" in social media, as initially defined by an unscientific report by the Southern Poverty Law Center. This purported phenomenon was widely reported in the media and on the presidential campaign trail as having contributed to a rise in aggression towards women, people with disabilities, and people in the racial, ethnic, and religious minority. However, little evidence was found to support the claim that controversial comments made by Donald Trump during the primary and general election led to a corresponding rise in anger, aggression, fear, or anxiety in social media. Interestingly, the most observable rises in such emotions corresponded with Hillary Clinton's "Alt-Right" speech, which consisted primarily of strong criticisms towards Donald Trump and a segment of his supporters. Not only did these findings shed doubt on the existence of a "Trump Effect" in social media, but they provided additional strong evidence of a link between elite polarization and mass polarization in social media. Specifically, high frequencies of affective rhetoric in Hillary Clinton's "Alt Right" speech corresponded with notable increases in affective rhetoric on social media, thus creating an atmosphere more conducive to the growth of mass polarization.

One central argument forwarded in this dissertation has been that the relationship between elite and mass polarization in social media cannot be measured effectively or accurately using the same set of tools that are used within traditional media. Because of the significant differences between the traditional and social media environments, the nature of elite influence must be defined contextually when considering its power in

social media. Simply put, elite cues are processed differently in social media than they are in traditional media sources. Some expressions of affective rhetoric can be transmitted, processed, and disseminated within social media in a manner that has a far greater impact than it would have had through traditional sources. Similarly, other expressions of affective rhetoric can be sequestered by social media networks through self-policing behavior that shuns extremely uncivil messages, thus blunting the reach and impact of otherwise high-volume messages.

Rather, such questions must be answered in a way that accounts for the unique context of the media, lending credence to Marshall McLuhan's 1964 observation that "the medium is the message". If it is true that the characteristics of the medium are just as important as the content delivered over the medium, one must take into account that the nature of the social media communications environment is structurally different than the traditional media environment. In this respect, social media plays a part in modifying the relationship between political elites and the mass public, at the very least.

Spiro Agnew once lamented that the state of media was being pervaded by "nattering nabobs of negativity". Saffire-crafted clever alliterations aside, there was wisdom in this warning. Political elites – whether they are elected officials or members of the media – hold significant power when it comes to influencing mass beliefs. Affective rhetoric – whether it is positive or negative in nature – has an impact on shaping political discourse in the mass public. Most importantly, the medium matters when seeking to determine the reach and impact of these cues. Social scientists must use a set of tools appropriate to the type of communications medium being observed.

This dissertation has sought to measure the extent to which elite polarization influences mass polarization in social media by employing a unique mixed methods set of tools. In doing so, the research presented herein has contributed to the political science literature by revealing a relationship between elite and mass polarization that would not have been observable using the traditional tools of social science. However, the immediate utility of these findings does not represent this dissertation's full contribution to political science. Rather, the most durable contribution may be defined by the novel methodological approaches employed while addressing the main research question, as they provide a flexible blueprint for future researchers who seek to better understand how networked technologies shape human interactions.

References

"Political Polarization in the American Public." Pew Research Center,

> http://www.people-press.org/2014/06/12/political-polarization-in-the-american-

> public/.

"Social Media Usage: 2005-2015." Pew Research Center,

> http://www.journalism.org/2014/10/21/political-polarization-media-habits/.

"Facebook Help Center: What Names Are Allowed on Facebook?". Facebook,

> https://www.facebook.com/help/112146705538576.

Abramowitz, Alan. 2013. *The Polarized Public?* New York: Longman.

Albert, Réka, and Albert-László Barabási. 2002. "Statistical Mechanics of Complex

> Networks." *Reviews of Modern Physics* 74: 47-97.

Andris, Clio; David Lee, Marcus J. Hamilton, Mauro Martino, Christian E. Gunning,

> John Armistead Selden. 2015. "The Rise of Partisanship and Super-Cooperators

> in the U.S. House of Representatives." *PLoS ONE* 10.

Artime, Michael. 2016. "Angry and Alone: Demographic Characteristics of Those Who

> Post to Online Comment Sections." *Social Sciences* 5: 68.

Artime, Michael Robert. "New Media, New Politics the Emergence of the Internet in

> American Politics." Ph.D., University of Missouri - Saint Louis, 2012.

Barberá, Pablo, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau.

> 2015. "Tweeting from Left to Right: Is Online Political Communication More

> Than an Echo Chamber?". *Psychological Science* 26: 1531-42.

Bennett, W. Lance, and Shanto Iyengar. 2008. "A New Era of Minimal Effects? The

> Changing Foundations of Political Communication." *Journal of Communication*

58: 707-31.

Blumler, Jay G. 2015. "Core Theories of Political Communication: Foundational and Freshly Minted." *Communication Theory* 25: 426-38.

Borah, Porismita. "Does It Matter Where You Read the News Story? Interplay of News Frames and Incivility in the Political Blogosphere." Ph.D., The University of Wisconsin - Madison, 2010.

———. 2013. "Interactions of News Frames and Incivility in the Political Blogosphere: Examining Perceptual Outcomes." *Political Communication* 30: 456-73.

Boyd, Danah. 2010. "Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications." In *Networked Self: Identity, Community, and Culture on Social Network Sites*, ed. Zizi Papacharissi: Routledge. 39 - 58.

Brandes, Ulrik. 2001. "A Faster Algorithm for Betweenness Centrality." *The Journal of Mathematical Sociology* 25: 163-77.

Brooks, Deborah Jordan, and John G. Geer. 2007. "Beyond Negativity: The Effects of Incivility on the Electorate." *American Journal of Political Science* 51: 1-16.

Converse, Philip E. 1962. "Information Flow and the Stability of Partisan Attitudes." *The Public Opinion Quarterly* 26: 578-99.

Druckman, James N., and Kjersten R. Nelson. 2003. "Framing and Deliberation: How Citizens' Conversations Limit Elite Influence." *American Journal of Political Science* 47: 729-45.

Druckman, James N., and Michael Parkin. 2005. "The Impact of Media Bias: How Editorial Slant Affects Voters." *The Journal of Politics* 67: 1030-49.

Entman, Robert M. 1989. "How the Media Affect What People Think: An Information

Processing Approach." *The Journal of Politics* 51: 347-70.

Farina, Cynthia R. 2015. "Congressional Polarization: Terminal Constitutional

Dysfunction?". *Columbia Law Review* 115: 1689-738.

Fiorina, Morris P., and Samuel J. Abrams. 2008. "Political Polarization in the American

Public." *Annual Review of Political Science* 11: 563-88.

Freeman, Linton C. 2004. *The Development of Social Network Analysis : A Study in the

Sociology of Science*. Vancouver, BC; Empirical Press

Gainous, Jason, and Kevin M. Wagner. 2011. *Rebooting American Politics : The

Internet Revolution*. Lanham, Md.: Rowman & Littlefield Publishers.

———. 2014. *Tweeting to Power : The Social Media Revolution in American Politics*.

New York: Oxford University Press.

Gruzd, Anatoliy, and Jeffrey Roy. 2014. "Investigating Political Polarization on Twitter:

A Canadian Perspective." *Policy & Internet* 6: 28-45.

Harary, Frank. 1959. "On the Measurement of Structural Balance." *Behavioral Science*

4: 316-23.

Hetherington, Marc J. 2009. "Putting Polarization in Perspective." *British Journal of

Political Science* 39: 413-48.

Himelboim, Itai, Stephen McCreery, and Marc Smith. 2013. "Birds of a Feather Tweet

Together: Integrating Network and Content Analyses to Examine Cross-Ideology

Exposure on Twitter." *Journal of Computer-Mediated Communication* 18: 40-60.

Hwang, Hyunseo. "Why Does Incivility Matter When Communicating Disagreement?:

Examining the Psychological Process of Antagonism in Political Discussion."

Ph.D., The University of Wisconsin - Madison, 2008.

Iyengar, Shanto, and Kyu S. Hahn. 2009. "Red Media, Blue Media: Evidence of

> Ideological Selectivity in Media Use." *Journal of Communication* 59: 19-39.

Joseph, Kenneth, Peter M. Landwehr, and Kathleen M. Carley. 2014. "Two 1%S Don't

> Make a Whole: Comparing Simultaneous Samples from Twitter's Streaming
>
> Api." In *Social Computing, Behavioral-Cultural Modeling and Prediction: 7th*
>
> *International Conference, Sbp 2014, Washington, Dc, USA, April 1-4, 2014.*
>
> *Proceedings*, eds. William G. Kennedy, Nitin Agarwal and Shanchieh Jay Yang.
>
> Cham: Springer International Publishing. 75-83.

Kavanaugh, Andrea, Ankit Ahuja, Samah Gad, Sloane Neidig, Manuel A. Pérez-

> Quiñones, Naren Ramakrishnan, and John Tedesco. 2014. "(Hyper) Local News
>
> Aggregation: Designing for Social Affordances." *Government Information*
>
> *Quarterly* 31: 30-41.

Kimball, David; Bryce Summary, Eric C. Vorst. 2014. "Political Identity and Party

> Polarization in the American Electorate." In *The State of the Parties : The*
>
> *Changing Role of Contemporary American Parties*, ed. John Clifford Green,
>
> Coffey, Daniel J., Cohen, David B., 7th ed. Lanham: Rowman & Littlefield. 37 -
>
> 54.

Levendusky, Matthew S. 2013. "Why Do Partisan Media Polarize Viewers?". *American*

> *Journal of Political Science* 57: 611-23.

Levi-Strauss, Claude. 1951. "Language and the Analysis of Social Laws." *American*

> *Anthropologist* 53: 155-63.

Lin, Yu-Ru. 2014. "Rising Tides or Rising Stars?: Dynamics of Shared Attention on

> Twitter During Media Events." *PLoS ONE* 9.

Marchionni, Doreen Marie. 2013. "Journalism-as-a-Conversation: A Concept Explication." *Communication Theory (10503293)* 23: 131-47.

Mark, Elliot, Fairweather Ian, Olsen Wendy, and Pampaka Maria. "Content Analysis." 'Oxford University Press'.

McCombs, Maxwell E., and Donald L. Shaw. 1972. "The Agenda-Setting Function of Mass Media." *The Public Opinion Quarterly* 36: 176-87.

Meraz, Sharon. 2009. "Is There an Elite Hold? Traditional Media to Social Media Agenda Setting Influence in Blog Networks." *Journal of Computer-Mediated Communication* 14: 682-707.

Morales, A. J., J. Borondo, J. C. Losada, and R. M. Benito. 2015. "Measuring Political Polarization: Twitter Shows the Two Sides of Venezuela." *Chaos* 25: 033114.

Moreno, J. L. 1934. *Who Shall Survive? A New Approach to the Problem of Human Interrelations*. Washington D.C.: Nervous and Mental Disease Publishing Co.

Mutz, Diana. 2006. "How the Mass Media Divide Us." In *Red and Blue Nation?*, ed. David Brady and Pietro Nivola. Vol. 1. [Stanford, Calif.]: Brookings Institution Press.

Mutz, Diana C., and Reeves Byron. 2005. "The New Videomalaise: Effects of Televised Incivility on Political Trust." *The American Political Science Review* 99: 1-15.

Newman, M. E. J. 2003. "The Structure and Function of Complex Networks." *SIAM Review* 45: 167-256.

Pariser, Eli. 2011. "The Filter Bubble (What the Internet Is Hiding from You)." *Medijska Istrazivanja/Media Research* 17: 202-05.

Rogers, Everett M. 1962. *Diffusion of Innovations*. New York,: Free Press of Glencoe.

Rowe, Ian. 2015. "Civility 2.0: A Comparative Analysis of Incivility in Online Political Discussion." *Information, Communication & Society* 18: 121.

Russell Bernard, H. 2005. "The Development of Social Network Analysis: A Study in the Sociology of Science, Linton C. Freeman. Empirical Press, Vancouver, Bc (2004)." *Social Networks* 27: 377-84.

Russell Neuman, W., Lauren Guggenheim, S. Mo Jang, and Soo Young Bae. 2014. "The Dynamics of Public Attention: Agenda-Setting Theory Meets Big Data." *Journal of Communication* 64: 193-214.

Santana, Arthur D. "Civility, Anonymity and the Breakdown of a New Public Sphere." Ph.D., University of Oregon, 2012.

Sebastião, Sónia Pedro. 2014. "Web Social Media, Politics, and Civism." In *Social Media in Politics*, ed. Bogdan; Monica Pătruţ Pătruţ. Vol. 13. Switzerland: Springer International Publishing. v - ix.

Sobieraj, Sarah, and Jeffrey M. Berry. 2011. "From Incivility to Outrage: Political Discourse in Blogs, Talk Radio, and Cable News." *Political Communication* 28: 19-41.

Sunstein, Cass R. 2001. *Republic.Com*. Princeton, N.J: Princeton University Press.

Vaccari, Cristian, Andrew Chadwick, and Ben O'Loughlin. 2015. "Dual Screening the Political: Media Events, Social Media, and Citizen Engagement." *Journal of Communication* 65: 1041-61.

van Dijck, José; Thomas Poell. 2013. "Understanding Social Media Logic." *Media and Communication* 1: 2-14.

Vincent, D. Blondel, Guillaume Jean-Loup, Lambiotte Renaud, and Lefebvre Etienne.

2008. "Fast Unfolding of Communities in Large Networks." *Journal of Statistical Mechanics: Theory and Experiment* 2008: P10008.

Wolf, Michael R., J. Cherie Strachan, and Daniel M. Shea. 2012. "Incivility and Standing Firm: A Second Layer of Partisan Division." *PS: Political Science and Politics* 45: 428-34.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. Cambridge [England] ;New York, NY, USA: Cambridge University Press.