

Birthers, Hand Signals, and Spirit Cooking: The Impact of Political Fake News Content on
Facebook Engagement during the 2016 U.S. Presidential Election

Grace Claire Wheaton

Thesis submitted to the Faculty of Virginia Polytechnic Institute and State University in partial
fulfillment of the requirements for the degree of

Master of Arts
In
Political Science

Caitlin E. Jewitt (Chair)
Eric H. L. Jardine,
Karen M. Hult

December 4th, 2018
Blacksburg, Virginia

Keywords: fake news, content analysis, 2016 election, Facebook, social media

Birthers, Hand Signals, and Spirit Cooking: The Impact of Political Fake News Content on
Facebook Engagement during the 2016 U.S. Presidential Election

Grace Claire Wheaton

Abstract:

Following the 2016 U.S. presidential election, scholarly attention became focused on the phenomenon of political fake news. Despite this interest, there is a lack of research examining the effect of fake news content on social media engagement during the 2016 general election period. My research seeks to address that gap through examining different content characteristics of fake news articles spread on social media in 2016, and testing the impact of those characteristics on Facebook engagement. To do so, I conducted a data collection of political fake news created between September 26st 2016 – and November 7th, 2016 gathering article text and Facebook engagement metrics. I then conducted content analysis on the article text collected, creating measures of four content characteristics, and test for the significance of those characteristics using parametric and non-parametric statistics.

I find political fake news circulated during the 2016 U.S. election is relatively homogeneous in content: it avoids policy discussion, is highly partisan, and negative in tone. Furthermore, personal content, policy discussion, partisan lean, and article tone have no detectable effect on the engagement received on Facebook. My findings provide avenues for future research, and seek to increase the understanding of the impact of political fake news.

Birthers, Hand Signals, and Spirit Cooking: The Impact of Political Fake News Content on
Facebook Engagement during the 2016 U.S. Presidential Election

Grace Claire Wheaton

General Audience Abstract:

Throughout the 2016 U.S. presidential election, public debate and media coverage was shaped by so called “fake news” – news articles which were intentionally false, and designed to influence opinion and policy. Although fake news itself is not a new concept, the way in which it was covered, and the way it was spread on social media platforms, was. Given this, scholarly literature examining fake news, and specifically the content or stylistic characteristics of fake news, is minimal. My research seeks to address that gap through examining different content characteristics of fake news articles spread on social media in 2016, and testing the impact of those characteristics on Facebook engagement (the number of likes or shares an article received).

I find political fake news circulated during the 2016 U.S. election is relatively homogeneous in content: it avoids policy discussion, is highly partisan, and negative in tone. Furthermore, personal content, policy discussion, partisan lean, and article tone have no detectable effect on the engagement received on Facebook. My research serves to provide avenues for future research, and increase our understanding of how fake news is spread. More importantly, given the negative influence fake news has on public discussion and democratic legitimacy, my research also increases our understanding of how to best combat the influence of fake news, and how to limit its spread.

Table of Contents

List of Figures: - v

List of Tables - vi

Chapter I: Introduction - 1

Chapter II: Literature Review - 3

Chapter III: Methods - 25

Chapter IV: Data - 35

Chapter V: Data Analysis – 38

Chapter VI: Conclusion - 68

Appendix A: Content Analysis Rubric – 74

Bibliography – 75

List of Figures

Figure 1: Model for Consumption and Production of Fake News Content on Social Media - 22

Figure 2: Dataset Characteristics - 37

Figure 3: Distribution of Normalized Shares Received - 45

Figure 4: Distribution of Normalized Likes Received, Dataset B - 46

Figure 5: Effect of Partisan Support on Median Engagement per Article - 55

Figure 6: Effect of Partisan Opposition on Median Engagement per Article - 55

List of Tables

- Table 1:** Fake News Content Classification Examples – 31
- Table 2:** Intercoder Reliability Check Statistic – 33
- Table 3:** Frequency of Fake News Articles and Narratives per Candidate – 40
- Table 4:** Percentage of Articles or Stories Discussing Scandal Content – 41
- Table 5:** Summary Statistics for Facebook Engagement Metrics – 43
- Table 6:** Summary Statistics for Impact of Personal Content on Engagement - 48
- Table 7:** Frequency Table for Articles Coded For/Against Candidates – 54
- Table 8:** Engagement Statistics, by Article Tone - 57
- Table 9:** Median Shares for Negative, Neutral, and Positive Content - 59
- Table 10:** Distribution of Article Tone, by Dataset – 60
- Table 11:** Table of Coefficients - 63

Chapter I: Introduction

The 2016 American presidential campaign differed from prior elections in many respects. One of the primary causes of these differences was the presence of political “fake news” on social media. 2016 represented the first open presidential election where a majority of Americans used a social media platform, and used social media as a news source (Pew Factsheet 2018, Shearer & Gottfried 2017). Furthermore, it represented the first widespread use of the term fake news to describe political content which is “intentionally and verifiably false, with the intent to mislead” (Allcott & Gentzkow 2017, 213).

Following the 2016 election, the attention paid to fake news and social media has increased. Research demonstrates that fake news is present and disseminated throughout social media (Allcott & Gentzkow 2017, Bessi & Ferrara 2017, Vicario et al. 2017). Moreover, reporting indicates that political fake news content is seemingly heterogeneous subject, with topics ranging from Pope Francis alleged endorsement of Trump, to Hillary Clinton’s weapons sales to ISIS (Silverman 2016). Perhaps most alarmingly, research shows that fake news producers sought to influence voter behavior through target exposure to false content, be it by individuals within the United States, or by external state actors. (Howard et al. Junk News 2017, Howard et al. Social Media 2017, Intelligence Community 2017).

Despite the increased attention and research on the subject, however, there is little research focused on the effect of fake news content itself; most research generalizes all false content, treating the article on Pope Francis endorsement the same as one discussing Clinton’s weapons sales. Furthermore, there is even less research about what impact fake news content – both subject matter and style - has on social media engagement (Horne & Adali 2017, Zannettou 2018).

In addition, the majority of research about fake news and social media is concentrated on Twitter, in spite of Facebook's greater popularity (Zannettou 2018). Approximately 68% of American adults use Facebook, and three-fourths of those individuals check the site at least once a day (Pew Factsheet 2018). Twitter, by contrast, is used by only about 24% of Americans, and Instagram (second most widely used platform) is used by only 35% of Americans (Pew Factsheet 2018). As the most popular social media platform in the United States, more research is needed on how and what type of political fake news is spread on Facebook.

Understanding what fake content individuals "like" and "share" is important, because it indicates what type of fake news content has the greatest reach on social media. Moreover, characterizing widely liked and shared political fake news can help prevent its dissemination, and mitigate fake news' impact on public discourse and political events. In a democracy where government is based on the free and open debate of ideas, it is important that individuals form their opinions with truthful information. Public discourse founded on false or misleading claims can lead to ineffective or misguided policy, as representatives respond to the false beliefs of constituents. More importantly, political fake news undermines the legitimacy of government, leading individuals to think that government cannot be trusted. Although dissent is integral to democracy, dissent based upon fundamentally false claims is harmful, and leads to increased division and mistrust within society.

Thus, the objective of this thesis is to understand the impact that different fake news content has, and identify what fake news content receives the most engagement. Specifically, does the type of political fake news content published during the 2016 general election affect the engagement that fake content receives as measured by the number of likes and shares received on Facebook?

To answer this question, I conduct a content analysis of political fake news articles published from September 26, 2016 to November 7, 2016, and measure four different characteristics of fake news content. I then test the impact of this content on Facebook engagement, examining what, if any, differences exist between the type of fake news content and quantity of Facebook engagement. I find that political fake news overwhelmingly avoids policy-oriented content, is anti-Democrat in partisanship, and negative in tone. Furthermore, while I find no statistically significant relationship between different types of political fake news content and engagement, my results do indicate that fake news producers create content specifically curated towards consumer preferences. Additionally, I find that engagement levels in political fake content decreased dramatically after the 2016 election. Although there are limitations to my analysis, my results provide a useful guide for future studies and point to areas where more research is needed.

The rest of the paper is as follows. Chapter II overviews the literature on fake news and social media. There, I define fake news, discuss who is involved in consuming and producing fake news, and the relationship between those groups. I also include the theoretical framework for my thesis. Chapter III describes my methods and how I operationalized my key independent and dependent variables, along with overviewing which stories are included in analysis. Chapter IV overviews my data, and the sampling methods used. Chapter V provides my analysis and results, along with discussion of key findings. Chapter VI concludes my analysis, and points to areas of future research.

Chapter II: Literature Review

The primary conceptual question in this study is what, if any, difference exist between types of fake news content and engagement levels received on Facebook. In this chapter, I

review the relevant literature on fake news and social media during the 2016 U.S. presidential general election. I first define political fake news and characterize its role in the 2016 general election. Then I provide an overview of the groups involved in fake news creation and dissemination: fake news consumers, fake news producers, and the social media platforms which mediate the relationship between the two. Finally, I situate my thesis within political science literature, and demonstrate how my research increases our knowledge of fake news content and how that fake content operates on Facebook.

Defining Fake News

While the 2016 general election has increased interest in and coverage on fake news, political misinformation and fake news are not new concepts.¹ Rather, the type of fake news seen in 2016 represents a deviation from prior manifestations of fake news. Until the 2016 election, the term fake news was primarily used to describe political satire that mimicked the structure and composition of authentic reporting (e.g. *The Daily Show with Jon Stewart*, *The Onion*) (Balmas 2012, Brewer & Marquete 2007, McBeth & Clemons 2011). Others suggested using the term as a type of folklore, and incorporated concepts as urban legends and myths into their definition (Frank 2015). This list is not exhaustive, and fake news has taken on a variety of definitions over the years (Tandoc et al. 2017).

None of these definitions, however, describes the political misinformation spread during the 2016 general election. As a result, I use the definition of fake news proposed by Allcott and Gentzkow, in their article “Fake News, Social Media, and the 2016 Election.” This was one of the first studies to quantify the impact of fake news as manifested in the 2016 general election

¹ Within the United States, fake news has roots in the yellow journalism of the 1900s, when the press focused on publishing sensationalist content to increase sales. Furthermore, misinformation and fake news have been mainstays of military operations, as a form of psychological warfare and intelligence (Uberti 2016).

(Allcott & Gentzkow 2017). They define fake news as “intentionally and verifiably false, designed to mislead the reader” (Allcott & Gentzkow 2017, 213), a definition that best categorizes the fake news I am researching.

This definition offers several advantages when characterizing fake news in the context of the 2016 general election. First, this definition includes the intentionality of fake news; producers of fake news deliberately generate fake content. It also considers verifiability - there is reason to believe a fake news article is false, and there is evidence available to counter the fake content’s claim. Finally, the definition also contains the deception element of fake news. Fake news is intentionally misleading; as such, this definition incorporates articles and websites that make efforts to present fake content as credible. This definition of fake news excludes content classified as fake news by prior literature, such as political satire or misreporting and accidental errors made by traditional media. It also excludes the recent appropriation of fake news as a way to attack traditional media outlets or unfavorable reporting, a strategy commonly used by President Trump and his supporters (Yglesias 2016).

An example contrasting the different manifestations of “fake news” provides some clarity, and demonstrates the merits of my definition. For example, prior research has classified the “The Daily Show with Jon Stewart,” and the “Pizzagate” conspiracy as fake news. The Daily Show is a well-known political satire show, while the Pizzagate scandal originated on Reddit and attempted to link Hillary Clinton to a child trafficking ring based out of a Washington, D.C. pizza parlor (LaCapria 2016).² Both examples discuss exaggerated and/or false political content. However, whereas The Daily Show is clearly satire and is transparent about that premise, content related to the Pizzagate conspiracy is unironic and misleading in nature. My definition of fake

² This, of course, was a completely false and unfounded claim, but did lead to a shooting at the pizza parlor in question, and received ample news coverage as a result (La Capria 2016).

news categorizes Pizzagate as fake news and does not categorize the Daily Show as such. Hence, the definition I use is a better measure of fake news as seen during the 2016 election than prior definitions.

The literature is also full of related concepts to fake news that help explain variation in fake news content. For instance, research on conspiracy theories, defined as an “effort to explain some event or practice by reference to the machinations of powerful people, who have also managed to conceal their role” (Sunstein & Vermeule 2009, 205), has ample application to fake news; false content frequently makes use of conspiracy theories, such as the Pizzagate scandal described above. Other related terms include astroturfing (Ratkiewicz 2011, 295), “alternative narratives” (Starbird 2017), and rumors (DiFonzo & Bordia 2007, 13). These terms are not mutually exclusive – fake news can include any of those concepts, all of which could fit the “intentionally and verifiably false” definition of fake news.

In sum, while the conceptualization of fake news has shifted over time, I provide a suitable definition to address my research question. Defining fake news as “intentionally and verifiably false, with the intent to mislead,” best describes the political fake news circulated during the 2016 general election, which is the focus of my analysis. I now outline the research examining how and why fake news is disseminated on social media platforms, and how the relationship between consumers of fake news, producers of fake news, and social media platforms helps shape the type of false content spread on those platforms.

Fake News and Social Media

As described previously, fake news is not a new idea or concept. Fake news and misinformation online have existed ever since the internet was invented (Fitzgerald 1997). What distinguishes fake news during the 2016 general election, rather, is how aspects of the election,

combined with social media, created an environment which facilitated the sharing of fake news. To more fully understand political fake news during the 2016 election, I examine the literature on the primary actors related to fake news: fake news consumers, fake news producers, and social media platforms. This discussion will provide the context necessary in order to hypothesize what types of content might receive more or less engagement on Facebook.

Fake News Consumers

Before considering fake news content, I first overview the literature on why individuals consume fake news or are misinformed more generally. Understanding what motivates an individual to consume fake news is an important step to hypothesizing what content they prefer, and what content might receive more engagement on Facebook.

First, research has long known that individuals base decisions on correct and incorrect information. However, the literature has only recently included a third form of information – misinformation - first characterized by Kuklinski as including individuals who are “not just in the dark, but wrongheaded” (Kuklinski et al. 2000, 793) when consuming information. These misinformed individuals hold factually incorrect views, and firmly believe those views are correct (Flynn et al. 2017, Hochschild & Einstein 2015, Kuklinski et al. 2000). Providing that third category of information allows for an important distinction, and enables researchers to study individuals who are incorrect in their views rather than simply being ignorant. More importantly, this distinction also applies to individuals who engage with fake news, as fake news is a type of misinformation; it contains “factual beliefs that are false or contradict the best available evidence in public domain” (Flynn et al. 2017, 2).

Following this distinction, research has sought to understand what makes an individual susceptible to misinformation, and why individuals seek out misinformation. One of the most

widely applied paradigms describing individuals and fake news is that of “directionally motivated reasoning.”³ When searching for political information, individuals search with the intent of achieving a specific outcome. Rather than looking for factual knowledge to become correctly informed, individuals seek information to reinforce their identity or current beliefs, and are hence “directionally motivated” in their reasoning: they wish to find information (motivated) and they wish to find information which confirms to their priors (directional) (Flynn et al. 2017, Kunda 1990, Taber & Lodge 2009).

Because this reasoning is intended to reinforce what an individual already believes, individuals seeking political information are susceptible to several cognitive biases, including confirmation bias, disconfirmation bias, prior attitudinal effect, among many others. Along with cognitive biases, partisanship and prior opinions are two of the most influential bases for directionally motivated reasoning. This means individuals are particularly prone to consuming content that conforms to their biases based on partisanship or prior opinion, which can lead to consuming misinformation (i.e., political fake news) as a result (Allcott & Gentzkow 2017, Berinsky 2017, Flynn et al. 2017, Taber and Lodge 2009).

In short, individuals believe and consume political information not because they want to be informed, but because it strengthens their identity and worldview. This makes individuals susceptible to multiple biases, and can cause them to consume and believe false information. Political fake news exploits these biases and type of reasoning, producing content that confirms the priors of the individual. Thus, fake news content is read and disseminated as long as it can cater to, and reflect, the preferences of those individuals reading it.

³ “Directionally motivated reasoning” is a type of reasoning – individuals are “directional” or “motivated” in their reasoning in other settings besides when seeking political information. Although these other situations are interesting, my thesis is focused on the outcome of this reasoning when it comes to political information, and will focus on that specific context.

Given that individual consumption of fake news is driven by directionally motivated reasoning, research has sought to determine who consumes fake news and in what quantity. In regards to who consumes fake news, the literature is increasingly showing that fake news is consumed by the ideological extremes of the political spectrum, and specifically by the far-right (Faris et al. 2017, Guess et al. 2018, Lazer et al. 2017, Narayanan et al. 2018).⁴ In contrast, there is little and scattered evidence regarding how much fake news individuals consume. Although research supports the idea that online social networks of individuals which consume fake news are highly concentrated (that is, a few individuals consume a majority of fake news), there is less information regarding casual encounters with fake news – e.g., scrolling past a headline on your NewsFeed, hearing a story on the radio, etc., and the impact that might have on individual perceptions – or how much fake news on average and individual saw (Allcott & Gentzkow 2017, Faris et al. 2017, Fournay et al. 2018, Guess et al. 2018, Starbird et al. 2017).

Despite the lack of research discussing how much fake news individuals consume, research is clear that individuals struggle to discern between fake and legitimate news, despite most feeling confident in their ability to do so (Barthel et al. 2016, Domonoske 2016, Ipsos & BuzzFeed 2016, Horrigan 2016). In addition, misperceptions supported by fake news are very challenging to correct. Since individuals consume misinformation and fake news because of directionally motivated reasoning, changing a misperception requires that the individual fundamentally change some aspect of themselves, making such a correction challenging. Furthermore, direct intervention is at best ineffective, and at worse backfires, making an individual hold more tightly to their misperceptions (Einstein & Glick 2014, Nyhan et al. 2010,

⁴ Although the political science literature generally understands the ideological composition of who consumes fake news (as individuals who mainly are far-right), much less is known about why this is the case. The predominance of far-right individuals likely comes from a variety of reasons, such as economic divisions, rural/urban divides, feelings towards globalization, racial tensions, etc. which warrant further attention and study from scholars.

Thorson 2015). Current research is showing that the best way to correct misperceptions is to prevent them from occurring in the first place, by teaching information literacy, and “inoculating” individuals to fake news content (Cook et al. 2017, Hannack et al. 2014).

In sum, individuals consume fake news through directionally motivated reasoning. They seek out (or are exposed to, as I discuss later) political information that confirms their identities and priors. Whereas the literature is clear regarding who consumes fake news (mainly tight networks of far-right individuals), literature on the frequency which fake news is seen or how best to correct misperceptions is less definitive.

Understanding the characteristics of fake news consumers’ preferences is important, as those motivations guide what content they engage in – if a certain type of fake news content is particularly appealing to an individual, they are more likely to engage with it, and thus increase the number of likes and shares that content receives. These differences indicate what types of content consumers prefer, and are possibly more susceptible to. My research is focused on identifying differences in content as measured by engagement on Facebook. Thus, understanding why individuals consume fake news guides my research, and indicates possible differences in content engagement. I discuss this idea further in my theory chapter, and test for some of these differences later on.

Additionally, the specific reason for why individuals consume misinformation and political fake news influences the type of fake news content produced. Given that, I now describe the literature on fake news producers, and how they use knowledge of consumption patterns to achieve their organizations’ goal.

Producers of Fake News

There are several different types of fake news producers, each with different rationales for creating and disseminating fake news. Of those, there are two principal objectives for creating and spreading fake news: profit and influence. I discuss each of these in turn, as they are important to understanding why and what type of fake news is produced.

The first of these objectives is profit collection via advertising revenue. Advertising is a common form of revenue for websites and social media platforms; companies pay platforms and websites to host their advertisements, with the hopes of reaching consumers who frequent them. Given this revenue model, fake news sites and articles are commonly used as a form of “clickbait” - articles or websites with attention grabbing headlines designed to increase reader engagement and maximize ad revenue (Chakraborty et al. 2016, Chen et al. 2015, Evans 2009). The most notable example of this practice during the 2016 general election was reporting on a Macedonian village that is thought to have produced over 100 pro-Trump websites (and several of the fake news articles analyzed later). The individuals producing the fake news professed no interest in the actual election result and were simply exploiting the environment surrounding the 2016 election to collect revenue (Silverman & Alexander 2017, Subramanian 2018). However, this is not the only instance of using fake news as a revenue source, and there is plenty of research and reporting or describing this phenomenon (Sydell 2016, Vojak 2017).

The second objective is political influence; that is, individuals, groups, and other actors that produce fake news to gain support for their cause. Literature on politically motivated fake news production divides into two categories based on the actor in question: individuals and groups internal to the United States, and actors external to it, particularly state actors. Fake news production motivated by political influence has generally received more press coverage than fake

news production for profit, and is the focus of most federal governmental action up to this point (particularly in reference to external state actors).

Internal politically motivated fake news production is driven by the far-left and far-right disseminating false content to influence citizens. As implied by the characteristics of fake news consumers, most political fake news generated in 2016 was targeted at and produced by the far-right (Faris et al. 2017, Guess et al. 2018, Lazer et al. 2017, Narayanan et al. 2018). In particular, the conservative website Breitbart served as a major hub for fake news dissemination in conservative networks during the 2016 election cycle, particularly on issues related to immigration (Faris et al. 2017).

The foremost external actor to use fake news as a means of political influence during the 2016 general election was the Russian government. There is definitive and conclusive evidence that the Russian government conducted an active misinformation campaign with the intent of undermining U.S. electoral institutions (FBI 2018, Intelligence Community 2017, Weisburd et al. 2016). The Russian government used social media as a means to disseminate convincing and hyper-partisan content to both the far-right and far-left. The objective of this operations was to increase partisan tensions and undermine U.S. institutions through promoting false content in public discourse (Coldewey 2017, Keating et al. 2017, Sullivan & Byers 2017). Although the literature on this area is still evolving, external political actors use of fake news is important to consider when examining who produces fake news.⁵

Additionally, there are overlaps between both internal and external far-right fake news producers and “trolls,” or individuals who produce and disseminate fake news content to see

⁵The use of misinformation as a weapon is not novel, and Russia was not the only political actor to do this. Recent evidence suggests Iran has conducted a similar campaign, and Russia has used the misinformation strategy many different ways, most recently in its annexation of the Crimea in Ukraine (Centre for Applied Research, Selyukh & Mak 2018).

what reaction it provokes. (Heikkilä 2017, Malmgren 2017). This especially so during the 2016 election, where some far-right fake news was produced as a way to troll traditional media outlets and disrupt political discourse. Furthermore, Donald Trump's candidacy – which some argued demonstrated characteristics of trolling – seems to have amplified trollish far-right political fake news or at least increased media coverage on it (Applebaum 2016, Kardas 2017, Ohlheiser 2016, Silver 2016).⁶ This phenomenon was also leveraged by Russia, which employed trolls and used trolling behavior and content during its influence campaign (Weisburd et al. 2016).

Thus, there are varied rationales for producing fake news. Furthermore, these are not binary distinctions; for example, the Russian government funded an independent troll farm which made revenue off of ads fake news articles as mentioned above. The categories I pose here are useful distinctions for classifying prior literature and describing the broad objectives of fake news producers.

More importantly, describing producers' reasons for creating fake news provides insight as to what content they produce. Both categories of producers benefit from engagement in their fake content, as more engagement will increase ad revenue and/or content reach. Producers will therefore generate fake news that maximizes engagement by creating content that caters to consumer preferences. Consumers engage with the content that fits their preferences, which therefore allows fake news producers achieve their specific aims.

To have content reach consumers, however, fake news producers must navigate social media platforms and the structures created by them. As I describe in the next section, research

⁶ It's important to note that I am not describing the objectives of all trolls. The type of content trolls produce is varied, and most of it is not overtly political. Furthermore, trolling has existed for as long as the internet has, so this is not a new phenomenon. However, the distinct nature of Trump's candidacy has increased attention in this area of research. (Burroughs 2013, Donath 2006, Philips 2016)

shows that social media platforms themselves play a crucial role in spreading fake news, and mediate the relationship between fake news producers and consumers.

Social Media and Fake News

Social media platforms, or “social networking sites,”⁷ play an important role in disseminating fake news. Whereas these sites have existed and played a role in political events since 1997, the 2016 general election represented the first open presidential election in which a majority of Americans used at least one social media site (Boyd & Ellison 2007, Pew Factsheet 2018). The 2016 election also was the first instance in which scholars could observe the role of social media in the context of an open seat presidential election. While some scholars were hopeful that such platforms and the internet more generally would facilitate cross cutting political dialogue, this idea was not borne out in the 2016 general election.⁸

Social media was a key vector in the transmission of fake news during the 2016 election (Guess et al. 2017). This is because social media platforms create the structures that connect producers of fake news to consumers of fake news. The literature provides two mechanisms for how social media interacts with fake news consumers and producers. First, social media platforms regulate what content consumers can interact with. Secondly, because social media platforms select what content reaches consumers, they also set the incentive structure for

⁷ Social network sites are defined as “web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system” (Boyd & Ellison 2007). This definition includes social media sites like Facebook and Twitter, but excludes other well-known platforms such as Snapchat (where there is no way to view another individual’s “list of connections”), and Reddit (which does not offer a well-used way to articulate a list of connections).

⁸ There is a fair amount of disagreement regarding if social networking sites and the internet have increased or decreased exposure to opposing ideological views and cross cutting political opinions. However, the current research is focused on 1) the dissemination of true content and 2) the impact of non-political networks as a vector for cross cutting opinions— so, interactions that just simply happen, not people looking for political news. The research does not account for the definition of fake news used here, or the political networks developed on social media platforms (Mitchelstein & Boczkowski 2015, Wojcieszak & Mutz 2009).

producers in their content creation. I discuss each of these points in detail, outlining what characteristics of social media create this relationship between consumers and producers of fake news.

i) Direct and indirect control

The first aspect of social media relevant to the discussion of fake news is how these platforms directly and indirectly regulate what content individuals consume. Social media platforms directly govern what consumers see through the process of “algorithmic curation.” This process uses an algorithm that predicts and selects what content an individual will find most relevant, and thereby maximizes the user’s time on the platform (Rader & Gray 2015). Facebook’s NewsFeed algorithm is an example of algorithmic curation, though it is by no means the only social media platform which uses this process.

In regards to Facebook’s NewsFeed algorithm – the primary interaction a user has with an algorithm on the site - little is known about the precise mechanisms used to curate content. Although Facebook is public about the general factors used in its NewsFeed, the exact calculations which rank content are closely guarded. Moreover, Facebook frequently changes how the algorithm ranks content, and does not consistently publicize alterations (Facebook 2018, Tien 2018). Most Facebook users do not explicitly know how their NewsFeed is curated, that an algorithm chooses what content they see, or what information Facebook collects to inform the NewsFeed algorithm (Agwin 2016, Eslami et al. 2015, Rader & Gray 2015). While a discussion of Facebook’s transparency issues is outside the scope of this thesis, suffice it to say that ignorance on the part of Facebook users does nothing to prevent the flow of misinformation.

More important than the direct control of content, however, is how social media indirectly controls what individuals see and engage with. This indirect control is driven by the

recursive relationship between how a user engages with content on Facebook, and how Facebook's algorithm selects what content appears in their NewsFeed. When users engage with content in their NewsFeed, each engagement provides a signal to the algorithm as to what content the user prefers. The algorithm then takes that signal and incorporates it into its calculation of how it ranks a user's NewsFeed, using prior knowledge to better predict future user interest (Seaver 2014, Gillespie 2014).⁹ For example, if a user consistently "likes" content related to the NFL, then Facebook will learn to include more NFL - related content in the user's NewsFeed.

In the context of social media and Facebook, this recursive relationship will limit the introduction of content that is counter to a user's preferences. The recursive relationship between users and their NewsFeed amplifies the preferences already present in individuals, and reduces the number of instances where these preferences can be altered, or new preferences can be introduced. To refer to the prior example, a user who consistently engages with NFL content might find that their NewsFeed becomes full of NFL related articles, videos, etc., possibly excluding other content they are interested in.

This process is not problematic in and of itself. However, this relationship can be very problematic in relation to political content. Here, users are already vulnerable to prefer and consume content in a biased manner as motivated reasoners (e.g., through selective exposure, confirmation bias). As users engage with political content according to specific biases, Facebook's NewsFeed will classify these biased choices as the user's preferences. The algorithm then places more biased content into the user's NewsFeed and initiates the recursive relationship

⁹ I describe the Facebook NewsFeed algorithm here, since it is most relevant to my research. However, it is important to note that this recursive feedback loop between user interaction and algorithmic curation applies across social media platforms, including Twitter, Instagram, etc.

described above. Similar to the preference for NFL content, if a user engaged with a fake news article questioning, for instance, the Clinton Foundation, the user would then have seen more extreme far-right content in their NewsFeed. Thus, Facebook's NewsFeed amplifies the political biases and misperceptions present in individuals and supplies content that supports these biases.

Furthermore, if a group of individuals with similar beliefs and opinions form a network (which is more likely, as Facebook's NewsFeed prompts individuals to make connections to likeminded individuals), this collective belief system can form an "echo chamber," a network where almost no new information is introduced, and where there are few checks to correct biases for misperceptions present in individuals. This process is well documented and well understood in the scholarship of political science and other disciplines (Bakshy et al. 2015, Bessi et al. 2015, Lewis 2012, Vicario et al. 2016, Aiello et al. 2012).

As a result of this direct and indirect regulation of user content and social networks, social media can magnify and intensify the biases present in individuals seeking political information. Furthermore, it generates an environment in which misinformation congruent to group values can disseminate quickly and broadly, regardless of its veracity – given that the group is composed of likeminded and similarly misinformed individuals, no one will stop the flow of misinformation, or introduce new information to counter it (Sunstein & Vermeule 2008).

Thus, fake news can spread quickly in a network, and possibly beyond one if there are sufficient connections outside of the homophilic group. These systems are what enable fake news to play such a prominent role in highly partisan, homophilic networks and influence public discourse as a result.

However, while the underlying interactions between users and social media platforms are important in understanding how fake news is spread, they do not address what characteristics of

fake news articles – their content - makes them more or less effective in manipulating this recursive relationship. To that end, I now look at the relationship between fake news producers and social media platforms, and describe how this relationship influences what fake content is created and seen by users.

ii) User Preferences and the Relationship with Producer Content

The second important influence of social media on content is the recursive relationship between user engagement and Facebook's NewsFeed which shapes what content producers create. Fake news producers understand the relationship between user engagement and Facebook's NewsFeed algorithm. Therefore, producers generate content that appeals to consumers' tastes (knowing that present user engagement will increase future user engagement) and thereby maximize ad revenue and/or reach. For example, if a fake news producer is trying to target highly partisan individuals, they will produce and spread heavily partisan content, knowing that this content will appeal to a user's preferences. If the user engages with their content, they increase the chance they see similar content from the producer later on, as that content is considered relevant in a user's NewsFeed. Additionally, other individuals in that user's network are more likely to see that content, further increasing engagement.

Furthermore, social media platforms provide means by which content producers can tailor content specifically to users' preferences, which upends the limitations of traditional media. This is done in a variety of ways, all of which are based in the structure of social networking sites.

First, social media allows producers to bypass the traditional content limits of earlier time periods. For instance, social media removes the traditional gatekeepers (e.g., newscasters, editors) that curated what content users saw; no editors control or check the accuracy of user

content spread on social media (Shaw & McCombs 1972, White 1950). This allows fake news producers to generate content that is highly salacious or partisan, and plays specifically to an individual's political biases (Vorst 2017).

In addition, the size and scope of social media platforms are much greater than earlier outlets. For example, Facebook has about two billion users, while Instagram has about 200 million (Chaykowski 2017). This means that along with having almost no limits on content type, there are large and diverse networks available for any producer to access, including fake news producers.

Furthermore, along with large available networks, social media provides tools to precisely target audiences with specific content. Producers can use Facebook's ad manager tools, Twitter's promotion campaigns, or other similar products which enable them to more effectively target their content (Facebook Business 2018, Twitter 2018). Through these mechanisms, producers can select which users they want to target based on specific characteristics, and specifically design content to cater to their interests.

The actions taken by the Russian government in order to influence the 2016 election provide a useful demonstration of how producers could bypass these prior limitations on content dissemination, through the use of social media to direct what and how they targeted content. As discussed earlier, the Russian government created content to appeal specifically to both the far-left and right; while extreme in nature, this content was widely shared in Russia's targeted audiences (FBI 2018, Hart 2017, Keating et al. 2017). Russia could create this content without limitations due to the lack of content regulation on Facebook. It could then take that content, and spread it among a huge network of individuals due to the size of Facebook. Finally, the Russian

government could take that large number of individuals, and target those who would likely believe and spread their misinformation through the targeting tools available on Facebook.

The Russian government knew that individuals would want to believe and engage with the messages they were disseminating, regardless of veracity. Furthermore, it also knew that Facebook's NewsFeed algorithm would continue to feed those individuals extreme content the more they engaged with it. Thus, by producing content catering to the targeted audiences' tastes, combined with the nature of social media, the government was able to widely disseminate messages attempting to undermine U.S. institutions.

In sum, there is ample literature describing the mechanisms for how social media platforms regulate what content a user sees through direct means (i.e., algorithmic curation), and indirect means (i.e., the recursive relationships between users and social media algorithms). Furthermore, content producers use and exploit this direct and indirect control of user content in order to achieve their specific goals, whether they be to maximize profits, or influence users.

Moreover, it is evident that fake news content itself is a crucial component in determining the relationship between social media users, content producers, and social media platforms. Despite the importance of fake news content, however, the literature has not thoroughly examined or tested differences between specific types of fake news content. Whereas research understands how and why fake news spreads, and who (in general) disseminates fake content, there is a gap in the literature regarding the specific subject matter of fake news articles (Horne & Adali 2017, Zannettou et al. 2018). Moreover, there is a disproportionate focus on fake news and Twitter, as opposed to other social media platforms, despite that Facebook was a known vector for fake news and is more widely used (Zannettou et al. 2018, Pew Factsheet 2018).

As such, my investigation provides a better understanding of what political fake news content received the most engagement on Facebook in the immediate months before the 2016 general election. By looking at fake news content, and specifically fake news content on Facebook, I hope to contribute to the understanding of what, if any, false content received more engagement than other types. The next section describes my theoretical framework for what content receives more engagement on Facebook, and how consumer tastes ultimately drive what content is produced, as mediated by social media platforms.

Theory

Having provided an overview of the literature around fake news and the actors related to it, I now establish my theoretical framework for how and why fake content is spread on Facebook. I do this by describing the relationships among consumers of fake news, producers of fake news, and social media platforms. In doing so, I suggest that content that receives the most engagement on Facebook satisfies the preferences of consumers, producers, and social media platforms more than other false or mainstream content.

Figure 1: Model for Consumption and Production of Fake News Content on Social Media

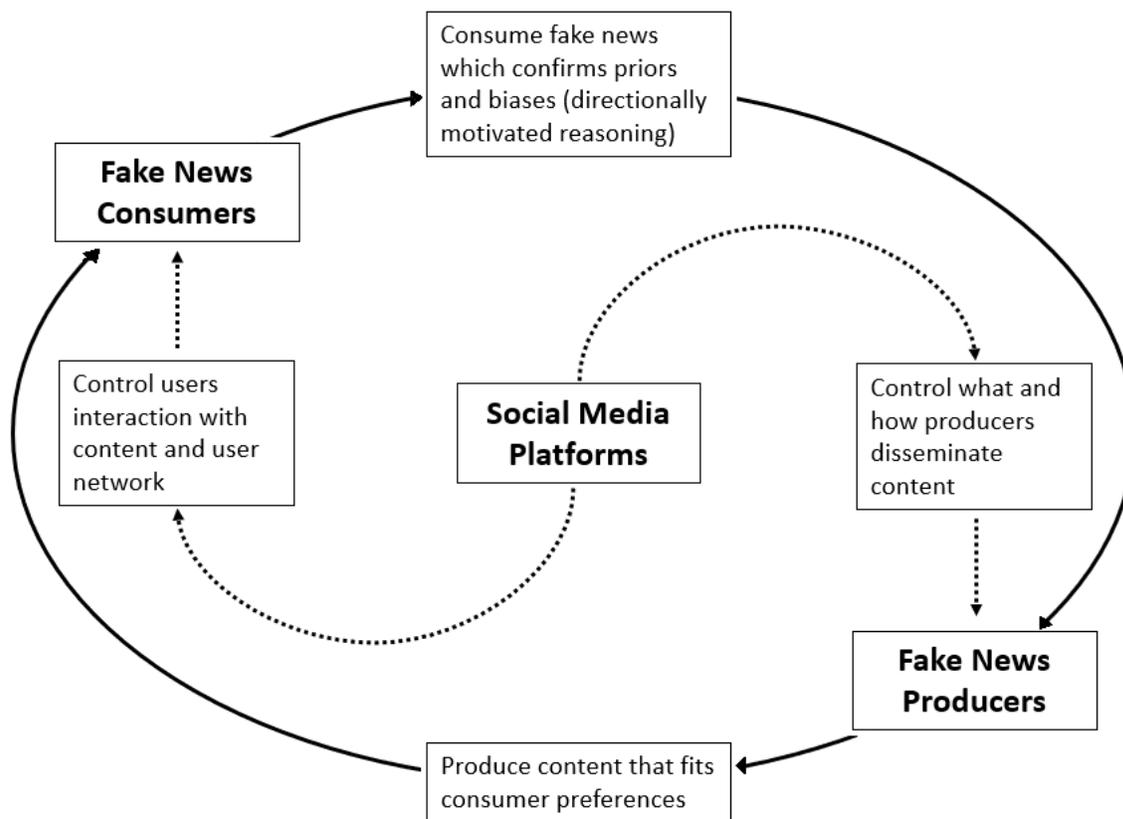


Figure 1 provides a representation of the relationships among consumers of fake content, producers of fake content, and social media platforms. Consumers of fake news content consume false content according to directionally motivated reasoning; consuming information that validates their identity or worldview, rather than to become better informed. Producers of fake content generate content to suit those individuals' preferences, as doing so maximizes engagement, leading to greater profit and/or influence. Social media platforms are what connect consumers to producers. They control what content consumers see through direct and indirect regulation of their NewsFeed and social network. They also control how producers disseminate content and establish what content is successful (i.e., can exploit Facebook's NewsFeed algorithm) and what content is not.

Given the above, fake news content is important because of this feedback loop – consumers pick what news they want to see via directionally motivated reasoning, and producers make content to meet consumer demand. Social media is the market where these preferences and goods meet, and sets the rules of the game as to what content from a producer gets seen and engaged with by a consumer.

I base my analysis on the idea that producers will create content individuals want to consume and that consumers will consume content according to directionally motivated reasoning as shown in Figure 1. As a means for determining what content consumers of fake news on social media prefer, I base social media users' content preferences on offline content preferences. That is, I believe individuals will react to and be drawn towards the same type of 2016 election content on and off social media sites, and react to content following traditional political media and behavioral preferences. These preferences will also mirror content preferences from other social media sites.

For instance, prior research finds that voters, particularly voters who do not affiliate strongly with one party or another, vote based on candidate impressions – that is, the personal aspects of a candidate that appeal to them (Fridkin & Kenney 2011, Lodge 1989). Furthermore, individuals are generally apathetic and ignorant when it comes to political issues (Pew 2015, Somin 2004, Somin 2010). These findings imply preferences as to what content individuals prefer on other media platforms. The heavy reliance on candidate impressions (i.e., voting for the candidate an individual likes more, not based on policy position) suggest that individuals prefer content discussing the candidates themselves, as opposed to other political news. In contrast, the generally low level of public knowledge implies that individuals do not consume policy related or “heavy” news items as a general preference.

Individuals consume political fake news content based on directionally motivated reasoning, and producers generate fake news content to fit individual preferences. Content that conforms to individuals' preferences will garner more engagement on social media, which is the immediate goal of content producers and social media platforms. Given that individuals have established preferences in traditional political content (i.e., describing political figures' personal characteristics appeals to individuals, and individuals tend to have low levels of political knowledge), one might assume that preferences in fake news content will at least mirror these established preferences, if not magnify them (e.g., if individuals who consume fake news have lower levels of political knowledge than average, it would increase, not decrease, the strength of these preferences). As such, fake news articles that focus on political figures' personal characteristics will be liked and shared more on Facebook than other narrative themes (H1). Furthermore, articles that contain a lower level of policy discussion will receive more likes and shares than articles that contain more policy discussion (H2).

Similarly, factors specific to the 2016 election cycle might also influence individuals' content preferences. For example, partisanship played a central role in the 2016 election. The 2016 election was more polarized and partisan than recent elections, as demonstrated in the rhetoric surrounding campaign and election coverage (Gentzkow 2016, Pew May 2017, Tyson & Maniam 2016). In regards to my framework, partisanship is a primary influence in an individual's view of media and is a key driver of directionally motivated reasoning. This pattern is demonstrated in individuals' view of media content, which falls on partisan lines, and how partisan content receives more engagement on social media than neutral content (Barthel & Mitchell 2017, Hughes & Lam 2017, Van Bavel & Pereira 2018).

Similarly, the 2016 election was also markedly more negative than recent elections. While this manifested itself in a number of ways, the most notable of them is the historically low candidate ratings (Bump 2016, Gieger 2016, Saad 2016). Furthermore, it is well established that content with a clear affect, and specifically a negative affect, receives more engagement on social media (Brady et al. 2017, Stieglitz & Deng-Xuan 2012, Ryan 2012). These two findings follow my theoretical framework – as individuals express their preferences for negative content (on and off social media), producers respond by producing more of it, and social media platforms perpetuate this cycle.

Given this environment, partisan tension will act to increase reactions to specific content; or, that the level of partisanship in content will have a positive correlation with the number of likes and shares an article receives (H3). Additionally, the level of negativity in tone in content will have a positive correlation with the number of shares of an article (H4). In short, I would expect that the partisanship and negativity demonstrated in public, offline, discussion will also demonstrate itself on social media through the reactions of individuals on Facebook.

Thus, I provide four hypotheses to test for variation in engagement between different types of fake news. The following chapter describes my variables my operationalization of them.

Chapter III: Methods

The purpose of my study is to assess the effect of fake news article content on Facebook engagement. As such, my primary independent variable is fake news article content. I operationalize this by measuring four specific dimensions in fake news content: article subject, presence of policy discussion, partisan bent, and overall tone.

I use content analysis to measure these four dimensions of my independent variable. I chose holistic grading as my specific method of content analysis, and used manual coding to

categorize my variables. These choices allowed me to best address my research question, and operationalize my independent variable. I elaborate on the merits of my method of content analysis below, first discussing content analysis generally, then holistic grading, and finally manual interpretation.

There are several advantages offered by content analysis. Content analysis “involves replicable and valid methods for making inferences from observed communications to their context” (Krippendorff 1980, 61). This method involves a researcher coding a qualitative trait (e.g., negativity of tone) by assigning it a value according to a predetermined and specific rubric of analysis (“1 = very negative, 2 = somewhat negative,”) for all of the content in the study sample. The rubric creates a uniform standard that leads to replicable categorization – any researcher who uses that rubric on the same content should produce identical results. When done appropriately, content analysis produces valid and reliable measures of a text’s content, measuring attributes and characteristics which are otherwise difficult to analyze in an objective manner. Furthermore, transforming qualitative traits into quantitative variables opens up the use of specific statistical methods, thereby facilitating more rigorous analysis of my research questions.

The method of content analysis used here - holistic grading - is well suited for analyzing fake news article content. This method “asks readers to interpret whole texts rather than count content at the level of words or sentences” (Hawkins 2009, 1049). Holistic grading looks at the fake news article in context, considering the images, font, and other visual attributes of the article which are not captured when only examining the textual elements of an article. Additionally, holistic grading has a researcher consider the interactions of an article’s characteristics. For example, while no individual element might be sufficient to make an article “highly negative” in

tone, a combination of an article's characteristics (e.g., format of title, use of unflattering visuals, derogatory nicknames, syntax, etc.) could interact to produce a negative impression. Examining the characteristics of an article as a whole, rather than isolating individual elements, generates a more accurate measure of an article's content and is an appropriate means for testing my research question.

Finally, I chose to use manual coding (as opposed to using machine learning or other automated methods) to categorize article content as it is the best method for conducting content analysis using holistic grading. Human coders are capable of understanding the complex writing syntax found in fake news; for instance, a human coder can discern language elements such as sarcasm or irony, while a computer struggles to do so. Furthermore, a human coder can include images and visual syntax elements (e.g., bolded text, associated pictures) that are commonly used in fake news articles when coding content. Although my analysis focuses on the textual elements of an article's content, the visual elements of fake news can significantly impact the overall impression of an article by increasing or decreasing the intensity of tone, partisanship, or other variables. Additionally, although manual coding is more labor intensive, my dataset was not so large as to make the method impractical. Given these considerations, manual coding was the superior means of conducting my content analysis.

Research on automated content analysis supports my method selection. While undoubtedly increasing in capability, computer-based coding struggles to analyze language at or above the sentence level (which is the focus of most literature on fake news). Automated coding is best suited for well archived digital material in a predetermined library, which my data are certainly not (Horne & Adali 2017, Lacy et al. 2015, Zamith & Lewis 2015). More importantly,

AI and machine learning coding techniques are a poor measure for measuring the overall tone or attributes of an article, which are the characteristics I wished to measure and test.

In sum, content analysis provides a valid and reliable way to transform the qualitative characteristics of an article's content into a quantitative measure more suitable for some sorts of analysis. Moreover, holistic grading best captures the overall tone and feel of an article, providing the most accurate measure of an article's content in context. Given the complexity and nature of fake news content and writing, and the relatively small nature of the dataset, manual coding was used to categorize article content.

Operationalization

Independent Variable

To operationalize my independent variable, I created a coding scheme which independently classified all four attributes of an article's content (subject, policy discussion, partisanship, and tone). I then coded every fake news article using this pre-determined schema, assigning each article a value for all four characteristics. I discuss the different scales and criteria used to code the four attributes below.¹⁰

First, all articles were coded based on subject matter. Articles were coded as referring to either a scandalous or substantive subject material – articles that contained content discussing clearly designed to provoke outrage against a perceived offense were categorized as “scandalous,” while articles that did not fit this definition were categorized as substantive.¹¹

¹⁰ A full rubric is included in the appendix.

¹¹ For this distinction, I use the categorization created from Faris et al. Although I could classify fake news as “substantive” v “not substantive” subject matter, distinguishing between content that is scandalous or substantive is a more accurate categorization of the fake news articles I analyzed. As I discuss later, many of the fake news articles I read discussed subject matter that was by definition “scandalous” – the article contained content that was meant to provoke outrage. Furthermore, the nature of my data was such that almost all articles fit neatly into one category or another – there were few articles which did not fit the binary I pose. Using this distinction emphasizes the differences between traditional and fake news content, and better characterizes the nature of fake news article in my dataset.

Then articles were coded based on the specific subject matter discussed (so, what scandal was referred to, or what substantive issue was discussed) (Faris et al. 2017). All articles were also coded according to which candidate the article referred to, which took on three values: “Clinton,” “Trump,” or “other” as a catch all category. Almost all stories (94%) were coded as discussing one of the two major party candidates.¹²

Following this, I coded articles based on whether content discussed the personal characteristics of a candidate. I used two criteria to code content as discussing a candidate’s personal characteristic. First, the article must reference a candidate’s private life (e.g. matters relating to the candidate and their family or close relationships) or personal condition. Second, an article must reference the candidate themselves or an immediate family member. To count as discussing a personal characteristic, an article must meet both criteria. Articles were coded as either containing or not containing content referring to a personal characteristic, and therefore appeared as a dummy variable in my analysis.

Policy discussion was also coded as a dummy variable, measuring the presence or absence of any policy discussion, defined as referencing a specific act, law, policy, etc. An example of policy content is an article discussing some aspect of the Affordable Care Act (“Obamacare”) or then candidate Trump’s “Border Wall.”

I used three separate variables to measure partisanship: a dummy variable for presence of partisanship, a three-point scale for partisanship in support of a candidate, and a three-point scale for partisanship in opposition to a candidate. All articles were coded for each of the three variables. This is because though nearly all articles analyzed were partisan, there was variation in how that partisanship was demonstrated - an article that contains anti-Democratic content does

¹² The “other” stories focused on then president Obama. These articles are aligned with the Democratic party, but stood out among all the articles which directly discussed Trump or Clinton.

not make it pro-Republican, and vice versa. To measure this variation, I assigned each article a separate value measuring support and opposition for each candidate and/or party. Content was coded as being “supporting Democratic/Clinton,” “neutral” or “supportive Republican/Trump,” and also as being “opposing Democratic/Clinton,” “neutral” or “opposing Republican/Trump.”¹³ By separating between positive and negative partisanship, I was able to perform a more complete analysis of partisanship in the datasets.

For example, an article titled “CAN HILLARY LIE HER WAY OUT OF THIS ONE? PHYSICIAN Says Hillary Has Parkinson’s Disease...Hillary Admits She Couldn’t Even “Get Up” After Convention,”(100 Percent, 2016) contains obvious partisan content. Based on my coding method, I would first code this article as partisan. Then, I would look to see if the article contained any content indicating partisan support. The article contains no reference to Trump, the Republican party, or any acknowledgement of an opposing side. Therefore, I coded this article as “neutral” for supporting a candidate. Finally, I would code an article based on the presence of oppositional partisan content. This article contains obvious anti-Clinton content; therefore, I coded this article as “opposing Democrat/Clinton” for the negative partisanship variable.

Finally, articles were coded for overall tone according to a three-point scale, with the categories as “positive,” “mixed/neutral,” and “negative.” Articles were classified by the presence or absence of professionalism (e.g., the presence of typos, neutral language, word choice) in their reporting, the use of derogatory names, punctuation, article subject matter, and other metrics to gauge tonality. To refer to the example above, I coded the article alleging

¹³ The terms Clinton/Democrat and Republican/Trump are used interchangeably throughout the article. While other parties do exist and influence American politics, and while Clinton and Trump are two individuals in large organizations, the time frame of analysis (the run up to the 2016 general election) makes distinguishing between the candidate and their party difficult to do in a meaningful way.

Clinton suffered from Parkinson’s Disease as “negative.” The capitalized “CAN HILLARY LIE HER WAY OUT OF THIS ONE,” along with the strong verb “lie” create an urgent, anxious tone. This, combined with the rest of the article’s content, indicate that this article has a negative tone, and was coded as such.¹⁴

Table 1: Fake News Content Classification Examples

Classification	Example	
<i>Personal Story</i>	http://100percentfedup.com/can-hillary-lie-way-one-physician-says-hillary-parkinsons-disease-hillary-admits-couldnt-even-get-convention/	
	CAN HILLARY LIE HER WAY OUT OF THIS ONE? PHYSICIAN Says Hillary Has Parkinson’s Disease...Hillary Admits She Couldn’t Even “Get Up” After Convention	
<i>Policy Discussion</i>	http://www.greenvillegazette.com/p/143404/	
	"Republican presidential nominee Donald J. Trump’s plan to eliminate the United States Department of Education could lead to more than 500,000 teacher layoffs"	
<i>Partisan</i>		
	pro Clinton	http://www.bipartisanreport.com/2016/10/12/breaking-hillary-clinton-vindicated-after-wikileaks-caught-fabricating-fake-emails-details/

¹⁴ Partisanship and tone were classified initially according to a five-point scale, but that scale was consolidated to improve inter coder reliability. Although the five point scale provided a more detailed level of measurement, coding articles as positive, neutral, or negative produced a much greater agreement between coders than when distinguishing between “partly” and “fully” negative or positive articles.

		"BREAKING: Hillary Clinton VINDICATED After WikiLeaks Caught FABRICATING Fake Emails (DETAILS)"
	pro GOP	http://freedomfinalstand.com/shoking-its-over-hillarys-isis-email-just-leaked-its-worse-than-anyone-could-have-imagined/
		PLEASE SHARE THIS IF YOU SUPPORT TRUMP IN HIS MISSION TO MAKE AMERICA GREAT AGAIN!!
	anti dem	http://www.thegatewaypundit.com/2016/10/awful-ap-shills-hillary-clinton-calls-bernie-supporters-bucket-losers-goldman-sachs-speech/
		AWFUL! AP Shills for Hillary Clinton After She Calls Bernie Supporters a "Bucket of Losers" in Goldman Sachs Speech
	nti GOP	https://newromantimes.com/trump-deport-lin-manuel-miranda-doesnt-realize-puerto-rico-us-territory/
		Trump continued with his racially insensitive diatribe . . . "Is [Puerto Rico] one of our 50 states? Is it? No, I thought not. It's a foreign power, and like all other

		foreign powers, I plan to cut all ties with the place as soon as I enter office.”
egative tone		http://www.fury.news/2016/11/wikileaks-john-podesta-attended-satanic-orgies-called-spirit-cooking/
		A shocking new revelation found within the latest Wikileaks email releases shows a dark satanic cult has infiltrated the circle of people surrounding Hillary Clinton
ositive Tone		http://worldtruth.tv/hillary-clinton-in-2013-i-would-like-to-see-people-like-donald-trump-run-for-office-theyre-honest-and-cant-be-bought/
		The former secretary of state [Clinton] argued that she thinks that businessmen “can’t be bought” and that they’re “very honest”

Table 1 provides examples of fake news content collected, and their respective categorization. The table includes a mixture of headlines and text from these articles, and displays text syntax as it appears in the original article (e.g., maintaining capitalization, punctuation). As evidenced by the quotes selected, the syntax of fake news articles relies heavily on long headline titles, strong punctuation, and capitalization – all highly visible forms of syntax and sentence structure – to communicate content and tone.

Furthermore, these examples are congruent with emerging literature comparing mainstream news to fake news content. This research finds that fake news syntax and structure differs substantially from “real” headlines, by using longer titles and simpler, more emotional, language (Horne & Adali 2017). Moreover, although one quote was chosen to exemplify “negative tone,” almost all of the examples (and all the articles in both datasets) demonstrate strong negative language, using verbs like “lying,” “AWFUL!,” “infiltrated,” and equally extreme adjectives.

Table 1 provides a general representation of the language and syntax seen in the fake news content I analyzed. It also demonstrates a clear variance in article subject matter, along with the intensity and partisan direction of these traits. The variance among fake news articles, and their impact on article engagement, are the focus of my analysis.

In sum, I measured my independent variable of fake news content through four characteristics: article subject, policy discussion, partisanship, and tone. I operationalized article subject, policy discussion and the presence of partisanship as dummy variables. I operationalized the nature of partisanship (i.e., candidate support or opposition) and tone according to three-point scales. The operationalization of my content dimensions in the manners listed allowed me to generate variables at the greatest level of measurement, while still maintaining reliability in coding.

Intercoder Reliability

To assess the reliability of article content analysis and coding criteria, I performed an intercoder reliability check. This involved another coder conducting a separate content analysis, following the same rubric and guidelines in order to test the replicability of my results. The other coder was provided examples of the fake news articles and received training as to the method of

content analysis. After that, they coded 15 randomly selected articles based on my coding methods. I use these results in my intercoder reliability check.

Table 2: Intercoder Reliability Check Statistic

Statistic	Personal Content	Policy Content	Partisan Content	Partisan Support	Partisan Opposition	Tone
<i>Simple Agreement</i>	87%	87%	100%	80%	87%	80%
<i>Cronbach's Alpha</i>	0.82	n/a	1	0.52	0.95	n/a

Table 2 displays the simple agreement (how often the coders agreed in percentage terms), along with Cronbach's alpha, a commonly used statistic to measure intercoder reliability. The statistic ranges from 0 - 1, and increases as the agreement between raters increases. In general, values of $a \geq 0.7$ are preferred, and indicate strong agreement, while values $a \geq 0.6$ are acceptable for exploratory research (Goforth 2015, Tavakol & Dennick 2011).

All the test items indicate high levels of simple agreement. Alpha values range from 0.52 to 0.94. One falls below 0.7, and two cannot be calculated due to lack of variance in coding results. Although some variables do fall below the 0.7 threshold preferred in research, I chose to include all the variables collected in my analysis. First, my content analysis is exploratory in nature; an alpha value that falls slightly below an arbitrary threshold is not sufficient cause to discard a variable. Secondly, time and resource constraints limited the extent to which another coder could be trained and used – a longer training period and more test articles might well lead to higher alpha regardless of the validity of the content measurement.

Furthermore, the Cronbach's alpha for each variable does not represent the actual degree of agreement between coders. This is due to the fact that the articles checked did not take on the full range of values available for each variable; for instance, none of the articles checked demonstrated support for the Democratic party. This meant some my variables had artificially

low variance and covariance, which lowered their alpha values despite not influencing the reliability of the method. Additionally, for some of the articles tested there was a lack of variance in results coded, which prevented the calculation of alpha (i.e., all the randomly selected articles were coded as “negative” by the checker) for both policy content and article tone variables. Moreover, the coders did not radically disagree in any specific instance regarding article content (i.e., no article was classified as “positive” by one coder, and “negative” by another), which is also not measured by α .

The high levels of simple agreement across variables more fully demonstrates the extent of agreement between coders. For example, the lowest level of simple agreement across variables was 80%; that is, 80% of the time both coders coded an article for the same value, which is a high level of agreement. Given the constraints on Cronbach’s alpha’s and additional context regarding agreement between coders, I believe my content analysis was sufficiently valid and reliable to continue further analysis.

Dependent Variable

The dependent variable of interest is the level of engagement a fake news article received on Facebook. This was operationalized as two variables: the number of likes and the number of shares each article received. To collect these metrics, I searched within Facebook and located the original post associated with a given website and article. The keywords associated with the article were used as search terms for locating the associated post within Facebook when necessary.

In sum, I operationalized my independent variable according to four different characteristics of fake news content. I operationalized my dependent variable of article engagement as the number of likes and shares an article received on Facebook. In order to

measure my variables I had to conduct my own data collection to acquire the information necessary to address my research question. The next chapter provides a more detailed overview of my data collection process and resulting datasets.

Chapter IV: Data

The data used were derived from data collected by Allcott and Gentzkow for their paper “Social Media and Fake News in the 2016 Election.” This dataset is comprised of 898 fake news URLs, covering 156 fake news stories. These stories were collected via a web scraper that collected fake news stories carried by factchecking websites Snopes and Politifact. These fake news articles were categorized as “False” or “Unproven” on Snopes, and “False” or “Pants on Fire” from Politifact and published between August 1st to November 7th, 2016 (Allcott & Gentzkow 2017). Multiple URLs are attached to one story, as several sources cover the same subject – for example, 21 fake news article URLs were collected discussing if Clinton had sold weapons to ISIS, which is connected to one Snopes factchecked article.

Although a useful starting point for identifying fake news stories, Allcott and Gentzkow’s dataset does not contain the article content or detailed engagement metrics needed to answer my research questions. To address this issue, I conducted my own data collection between June and July of 2018, collecting fake news article texts for the active URLs present in Allcott and Gentzkow’s dataset. I then collected engagement metrics for those articles with active URLs, searching within Facebook and recording article engagement metrics (likes, shares, etc.).

Dead URLs, stories that were not “False,” “Unproven,” or “Pants on Fire,” satire articles, videos, and fact checked stories without associated fake news articles were excluded from my data collection. I also excluded fake news stories without active URLs – if the article did not possess a live link, I did not include it in my search. Without the active URL, I could not retrieve

the article's content to perform content analysis on, and could not measure my independent variables.

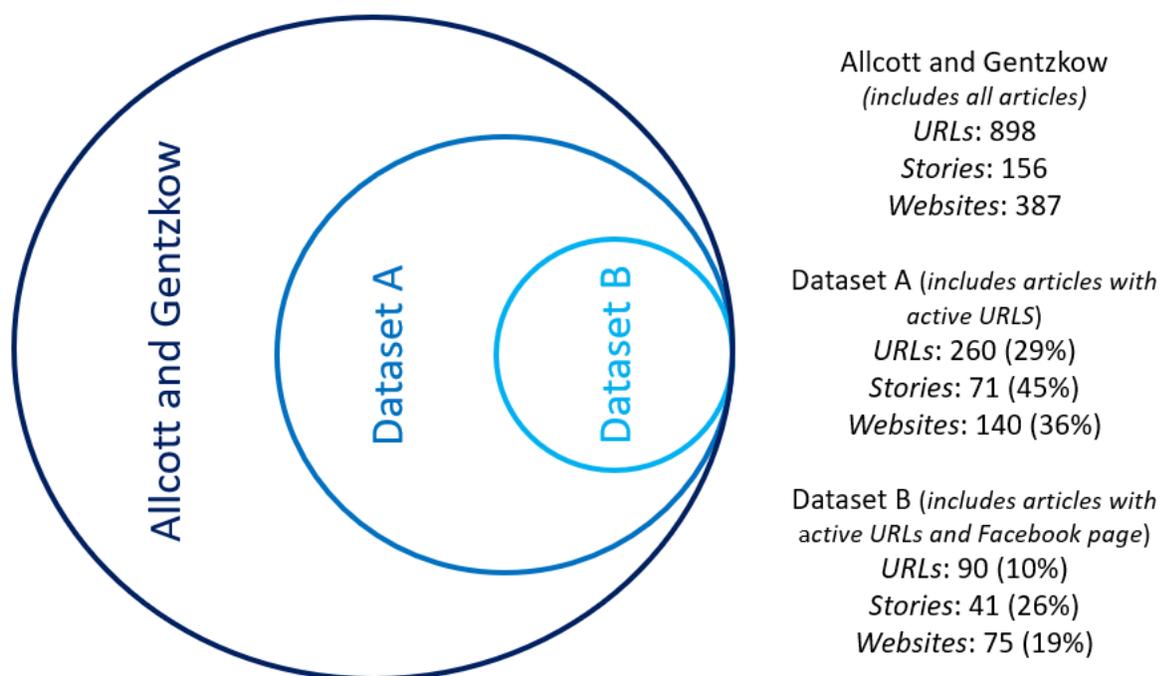
By conducting my own data collection, I was able to gather more measures of an article's engagement on Facebook and also produce a cleaner dataset (i.e., one that only contained truly "fake news" as defined by the authors) than that which Allcott and Gentzkow produced with their web scraper. Through the collection of more measures of engagement, along with my content analysis, my data were better able to analyze the impact of article content on engagement.

Discussion of Datasets

The result of this process is two related datasets, (or three, if the initial data collection performed by Allcott and Gentzkow is included). The first of these datasets is comprised of articles that have active URLs, but not necessarily a searchable Facebook post. It includes only shares on Facebook for engagement metrics, along with 263 active URLs/fake news articles covering 79 unique stories (Dataset A). The second dataset is a subset of Dataset A and is comprised of articles that have both an active URL and also a located Facebook post. This dataset contains 90 active URLs/fake news articles covering 41 unique stories (Dataset B). Dataset B contains likes, shares, angry, love, haha, wow, and sad reactions, and shares as measured by Allcott and Gentzkow. All of those metrics, excluding the share numbers as originally collected by Allcott and Gentzkow, were gathered during my data collection and are exclusive to Dataset B.¹⁵ Figure 2 visualizes the difference in datasets, and their connection to one another.

Figure 2: Dataset Characteristics

¹⁵ Although I did collect angry, love, haha, wow, and sad reactions, I do not conduct detailed analysis on them. This is due to small sample size for many of these reactions.



As Figure 2 displays, there is a substantial decrease in the number of fake news articles when restricting analysis to articles with active URLs. This decreases even further if restricting analysis to fake news articles that also have an active Facebook post at the time of data collection. Only approximately 10% of the URLs initially collected by Allcott and Gentzkow between September to November 2016 survived until June 2018.

This decay is not unusual or surprising, given characteristics of fake news articles and the websites that host them. As discussed in the literature review, one of the primary motivations for producing fake news is to collect advertising revenue, and several of the websites analyzed here are known to have used articles as clickbait – using fake news as a vehicle to generate advertising revenue. These fake news websites share characteristics of spam sites, including short lifespans (e.g., a few weeks), making it unlikely that those websites would continue to provide active URLs (Georgiou et al. 2008). Furthermore, Facebook has taken well publicized action to decrease the presence of fake news on its platform, in part due to the considerable scrutiny it incurred since the election. While the extent of these actions is unknown, Facebook

has removed billions of accounts and posts from the platform since the 2016 general election (Romm & Harwell 2018, Rosen 2018). This increases the likelihood that the Facebook post associated with a specific article was removed by the platform between Allcott and Gentzkow's initial data collection, and my data collection in June 2018. Given these events, the decline of live URLs and number of posts collected approximately a year and a half after the election is unsurprising.

While the decrease in live fake news URLs and Facebook posts is easily understood, it introduces concerns regarding selection bias. My data are not randomly selected: the web scraper used by Allcott and Gentzkow, the actions taken by Facebook, or the nature of the articles themselves all introduce different filters as to what articles were selected, and which ones remain analyzable. Furthermore, content that fake news users interact with is also not randomly selected; as I described previously, social media platforms specifically curate user content to increase user engagement.

However, I took a number of steps to minimize these issues. Use of the three datasets (Allcott and Gentzkow's, Dataset A, and Dataset B) provides ways to verify the representativeness of my sample by comparing the distribution of content across datasets. Additionally, I can compare the type and frequency of content I collected to descriptive details outlined in prior literature, and see if my results are substantially different. As I describe throughout this analysis, my findings are in line with current understandings of political fake news content. This similarity between my data and other research provides support for the idea that while issues of selection bias might be present in my research, the content present in those data is likely representative of the population of political fake news more broadly (Faris et al. 2017).

Moreover, URLs that remained active, and the articles still have an active Facebook post are themselves data. What, if any, difference exist between the population of political fake news prior to the 2016 election and the sample remaining two years afterwards provides information as to what factors impact the longevity of fake news content, and also hints at what content might avoid detection by counter fake news efforts. These are important and valuable areas for future research.

I will primarily use the data I collected for Dataset B in my analysis, but use Dataset A and Allcott and Gentzkow's data to verify my results when possible. The use of Dataset B allows for the analysis of the detailed engagement metrics collected from Facebook. Dataset B also records what articles still contain active URLs and Facebook posts, which are not included in Dataset A or Allcott and Gentzkow's original dataset. While there are concerns regarding selection bias present in my sample, I am aware of the issue and impact of this bias, along with the shortcomings of my dataset, and keep that in mind throughout the analysis. The next chapter presents my data analysis and discusses the results and implications of that analysis.

Chapter V: Data Analysis

In this chapter, I describe my data analysis, and detail the impact of different types of fake news content on Facebook engagement. I first overview the basic features of the independent and dependent variables. I then examine each of my hypotheses, and test the effect of personal characteristic subject material, policy discussion, partisanship, and overall tone present in fake news content on Facebook engagement.

Data Overview: Fake News Content

Before testing my content variables, I first overview the subject matter discussed in each fake news article. I describe two main properties: which candidate a fake news article referred to, and whether a fake article was discussing a substantive or scandalous subject.¹⁸

Table 3: Frequency of Fake News Articles and Narratives per Candidate

Candidate	Articles A	Articles B	Narratives A	Narratives B
<i>Clinton</i>	73% (191)	60% (54)	66% (47)	66% (27)
<i>Trump</i>	23% (61)	36% (32)	30% (21)	27% (11)
<i>Other</i>	3% (8)	4% (4)	4% (3)	7% (3)
Total	100% (260)	100% (90)	100% (71)	100% (41)

Table 3 details the percentage of unique articles and stories found in Dataset A and B for each candidate. As illustrated, a majority of articles – 73% in Dataset A, and 60% in Dataset B – had subject matter which focused on Clinton. The concentration of fake news narratives displays a similar pattern, as approximately 66% of all the fake news narratives were focused on or related to Clinton. Thus, a majority of fake news articles and narratives in 2016 contained content related to Clinton.

Table 4: Percentage of Articles or Stories Discussing Scandal Content

Articles			
Candidate	Scandal	Substantive	Total
<i>Clinton</i>	93% (50)	7% (4)	100% (54)
<i>Trump</i>	84% (27)	16% (5)	100% (32)
<i>Other</i>	50% (2)	50% (2)	100% (4)
Total Articles	88% (79)	12% (11)	100% (90)
Narratives			
Candidate	Scandal	Substantive	Total

¹⁸ Going forward, my primary unit of analysis is fake news articles. As I mentioned in Chapter IV, multiple articles are attached to one major fake news narrative. I wish to measure the effect of fake news content and the variation between articles, which necessitates article level data - hence using individual articles as my unit of analysis. However, I occasionally include fake news stories/narratives in my analysis. I do this in order to show the concentration of articles around a narrative; it allows me to see if many articles were discussing only one or two major narratives, or if many articles were discussing many different narratives. In short, it measures the density of content, and lets me see what content attracted the most publication attention.

<i>Clinton</i>	85% (23)	15% (4)	100% (27)
<i>Trump</i>	55% (6)	45% (5)	100% (11)
<i>Other</i>	67% (2)	33% (1)	100% (3)
<i>Total Stories</i>	76% (31)	24% (10)	100% (41)

Table 4 displays the frequency of fake news articles discussing substantive or scandal content for each candidate. Of the articles related to Clinton, almost all (approximately 93% of articles in Dataset B), were discussing a “scandal” of sorts, rather than a substantive campaign topic. Scandals discussed included questions of Clinton’s health, issues related to her email server, Wikileaks document drops, and other similar topics. When keeping in mind Table 3, and how stories on Clinton comprised a majority of fake news articles collected, this suggests that most fake news articles discussed a scandal including or related to Clinton.

In contrast, fake news content regarding Trump contained slightly more substantive articles than Clinton, with 85% of the articles on Trump referring to a scandal of sorts. However, only half (55%) of the fake news narratives on Trump discussed a scandal. Thus, those narratives had a greater concentration of articles than scandal narratives on Clinton did – while approximately the same ratio of articles to stories focused on scandals for Clinton (i.e., 93% of the articles on Clinton are associated with 85% of the narratives on her), 77% of articles referring to scandals related to Trump are associated with only 57% of unique stories on him. In short, whereas fake news content on Clinton had many articles covering many topics, fake news content on Trump was more concentrated on a few specific topics.

The heavy focus of article content on Clinton seen in my data parallels that seen in traditional media. In mainstream news sources, Trump succeeded in shaping media coverage, whereby a majority of content focused on Clinton and scandals associated with her (Faris et al. 2017). Moreover, as I described in my literature review, fake news is a product of the far-right.

These far-right fake news websites are significantly more prolific and powerful than far-left equivalents (Faris et al. 2017, Guess et al. 2018, Howard et al. Junk News 2017, Lazer et al. 2017). Thus, the number and content of stories on Clinton is predictable given the political alignment of most fake news content producers.

In sum, Clinton received substantially more attention by fake news articles than Trump in both datasets. Fake news content for both candidates primarily discussed scandalous topics, though articles referencing Trump saw a greater mix of substantive and scandal content than ones referencing Clinton. Additionally, the composition of my data matches the descriptions of prior research, which provides support for the validity and representativeness of my data.

Having overviewed who and what was discussed in fake news content, I now describe the composition of my dependent variable; engagement as measured by the number of likes and shares received on Facebook. I provide measures of central tendency and dispersion for each measure of engagement, along with summary statistics on a per article basis.

Data Overview: Facebook Engagement

Table 5: Summary Statistics for Facebook Engagement Metrics

Engagement Measure	A&G Shares Dataset A	A&G Shares Dataset B	Shares	Likes
<i>Mean</i>	26,715	33,735	4,417	1,871
<i>Median</i>	6,050	5,700	416	354
<i>SD</i>	60,046	73,456	18,911	5,123
<i>Minimum</i>	1,000	1,000	2	1
<i>Maximum</i>	384,000	384,000	165,000	41,000
<i>N</i>	260	90	90	90

Table 5 presents summary statistics of shares received by articles in Dataset A. In addition, it also presents the summary statistics for number of likes and of shares received by articles in Dataset B as recorded in my data collection.

The most prominent result is the difference between the absolute number of shares an article received as recorded by Allcott and Gentzkow, and the absolute number of shares I recorded. For instance, the mean number of shares for articles in Dataset B using Allcott and Gentzkow's data is 33,735 shares per article. In contrast, the mean number of shares for articles in Dataset B using my data is much lower, at 4,417 shares per article. This is also the case if looking at the mean number of shares in Dataset A. Given that Allcott and Gentzkow's data collection was conducted a year and a half before mine, this difference is unexpected – one would expect for engagement to increase in the interim, rather than decrease so dramatically.

There are several possible explanations for why this result occurred. First, individuals might have deleted or removed certain posts from their Facebook accounts. Facebook measures the engagement of posts from a live dataset; therefore, deleting an account or post would lead to a subsequent loss in engagement for items that account engaged in. Furthermore, if an individual chose to unlike a page, or unshare it between data collection periods, those actions would also reduce the number of shares an article received (Facebook Business 2015, Whiting 2018). Given the size of the decrease, however, individual level variation is unlikely to account for the total decrease in the number of shares.

A more likely explanation which accounts for this difference is actions taken by Facebook to reduce the prevalence of fake news on its platform. The same purge of accounts that decreased the number of active Facebook posts I could collect also decreased the total shares each article received. If Facebook truly deleted billions of fake accounts and bots from its

platform, the activity of these fake accounts would also be deleted, leading to the subsequent loss in shares. Given the well documented and widespread influence of fake accounts, bots, and other inauthentic actors across platforms, this explanation is likely the primary reason for the difference in shares between datasets (Bessi & Ferrara 2017, Howard et al. Junk News 2017, O'Connor 2017).

More importantly, what remains of these shares is more likely to represent legitimate user engagement; that is, individuals with legitimate accounts (as opposed to bots or fake accounts) sharing a fake news article. Though analyzing the shares I collected, I am better able to measure the effect of fake news content on real user engagement, rather than engagement by bots or false accounts. Although there are limitations in my analysis (i.e., I could only make this comparison for the number of shares an article received, and cannot know what accounts are authentic), my results are still useful and increase our understanding of how many individuals likely engaged with fake news content.

The difference in absolute share numbers between the datasets, although interesting, represents a potential source of error in my analysis – some articles might have lost more shares than others, leading to inaccurate results. However, if the two sets of shares follow the approximately the same distribution, I can assume that the number of shares lost per article is roughly equal. That is, the relative number of shares each article received by both measures is the same, and thus does not impact the relative results.

Figure 3: Distribution of Normalized Shares Received

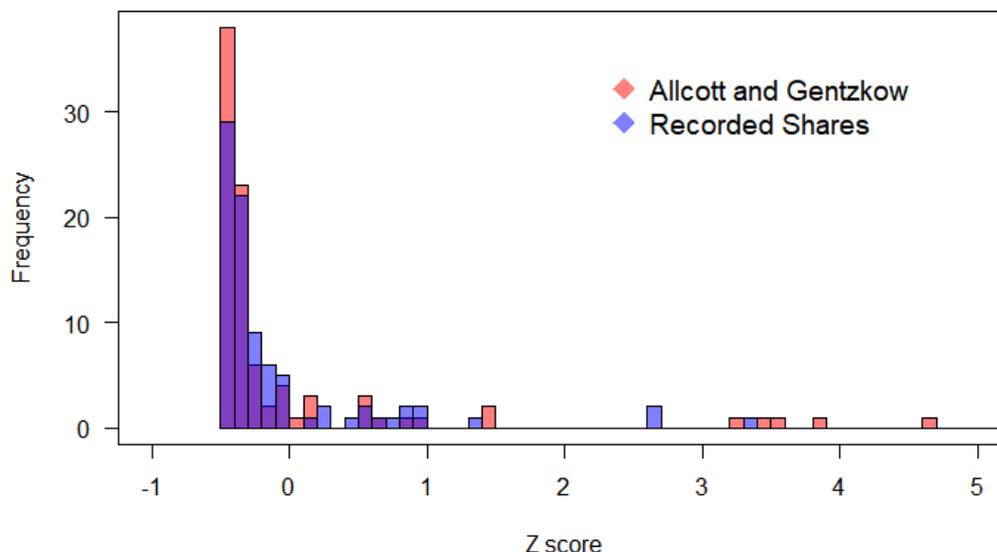
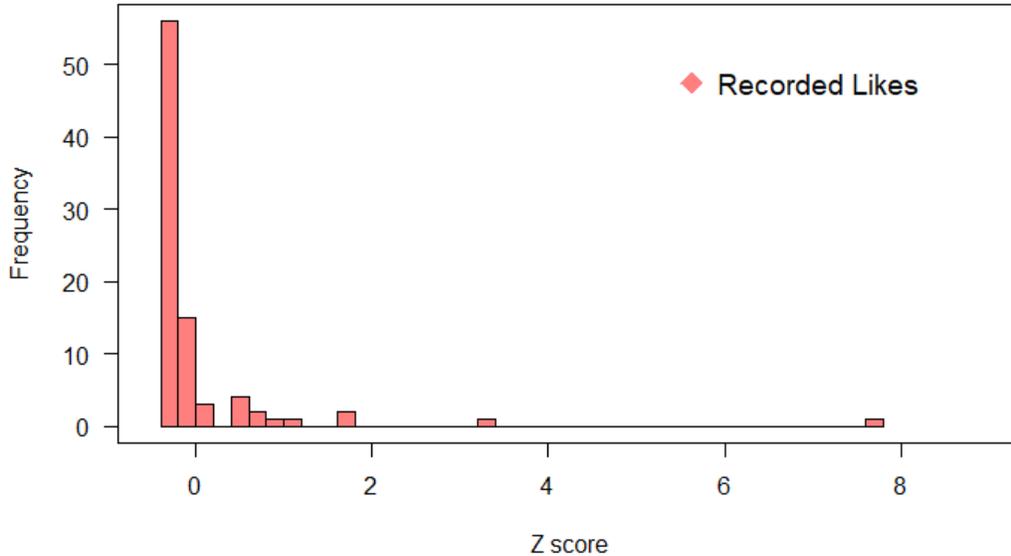


Figure 3 displays the distribution of shares received by articles in Dataset B, after normalizing the number of shares from both my data collection and Allcott and Gentzkow's collection. The results were normalized to reflect a Z-score, with the mean number of shares per article centered around zero to facilitate interpretation. Once corrected, the two shares measures generate approximately the same distribution. For each, the number of shares received is right skewed, with most articles possessing fewer than the mean number of shares, and a few outliers situated several standard deviations above the mean. A Mann-Whitney U test for difference in distributions between the two shares returned null results, meaning the two distributions are not statistically different.

Although the number of shares recorded by the two data collections differ in absolute terms, they follow a similar distribution once normalized, thus indicating the relative relationships between articles is the same. In short, both measures of shares per article are valid and provide useful information; shares data collected by Allcott and Gentzkow is a better measure of absolute engagement during the general election, while my shares data is a better measure of current real engagement.

Figure 4: Distribution of Normalized Likes Received, Dataset B



Given the results produced by normalizing shares, I also normalized likes per article. This allows me to examine the distribution of this measure of my dependent variable, and choose appropriate statistical test later in my analysis.

Figure 4 illustrates the number of likes received for articles in Dataset B, normalized using a Z-score with a mean centered on zero. Like the number of shares received, the distribution of likes per article is right skewed, with most articles receiving fewer than the mean number of likes per article, and a few outliers many standard deviations above from the mean.

Given the right skewed distribution for both likes and shares, and the small sample size in Dataset B (90 articles), I cannot assume that the measures for my dependent variable are normally distributed. To accommodate that fact, I report medians along with means (which allows for analysis that is more robust to outliers), and test data according to non-parametric methods when possible.

Having overviewed the key measures of central tendency and dispersion for my independent and dependent variables, and analyzed the composition of each, I now analyze each of my hypotheses. I test each hypothesis in the order they were presented, first testing whether

content describing the personal characteristic of a candidate receives more engagement on Facebook than other content.

Hypotheses Testing

H1: *Fake news articles discussing the private or personal characteristics of a candidate receive more engagement on Facebook than other types of fake news articles.*

My first hypothesis examines if content describing the personal characteristic of a candidate receives more engagement on Facebook than other forms of content. As described earlier, I define content referring to personal characteristics according to two criteria. First, content must reference the candidate themselves or an immediate family member. Secondly, content must discuss a candidate’s behavior or characteristics in their private life.

As shown in Section III, Table 1 provided an example of a fake news article displaying personal content, which discussed allegations that Clinton had Parkinson’s disease and attempted to “LIE HER WAY OUT” of it. The article is discussing Clinton herself, and is referencing a personal characteristic of hers (her health) in her private life – she was obviously not public about her alleged battle against Parkinson’s disease. Thus, this article contains discussion of personal characteristics, and was coded as a personal story.

Table 6: Summary Statistics; Impact of Personal Content on Engagement

Candidate	Category	Number of Stories	Average Likes	Median Likes	Average Shares	Median Shares	Average Shares, A&G
<i>Clinton</i>	<i>Non-Personal</i>	43	2,651	357	2,229	356	3,9477

	<i>Personal</i>	8	76	51	107	90	4,057
	Total	53	2,161	291	1,802	282	34,518
Trump	<i>Non-Personal</i>	31	709	350	596	296	5,242
	<i>Personal</i>	1	35	35	64	64	301,200
	Total	32	688	339	580	273	14,491
All	<i>Non-Personal</i>	78	1,822	376	1,572	356	29,763*
	<i>Personal</i>	12	2,036	95	1,792	99	66,609*
	Total	90	1,787	350	1,512	296	34,422

* Indicate statically significant difference between values at the $\alpha = 0.05$ level

Table 6 displays the mean and median engagement per article based on candidate and article subject. For both candidates, mean and median engagement is substantially lower for personal content. For example, personal stories on Clinton receive on average 76 likes and 107 shares per article, whereas all other stories on Clinton received 2,651 likes and 2,229 shares per article. Furthermore, the median engagement for personal stories on Clinton is also substantially less than the median engagement for non-personal stories. The data provide evidence against the hypothesis, and implies that content discussing the personal attributes or characteristics of a candidate actually receive less engagement than other types of content.

The difference between engagement received for personal stories and non-personal stories on Trump follows a similar pattern. Both mean and median engagement received by personal stories on Trump are substantially less than engagement for non-personal stories on Trump. While the small sample size for personal stories on Trump (only one story in Dataset B was coded as a personal story) is a limitation on the generalizability of this finding, the fact that only one story was coded as “personal” is notable itself.

The dearth of personal stories on Trump demonstrates that fake news content in Dataset B referencing Trump did not focus on his personal life or characteristics, despite Trump having a

tumultuous personal history. However, fake news is much more prevalent within conservative networks; the bias within these networks might select against unfavorable personal stories on the head of the Republican party. Fake news readers did not prefer content placing Trump in a negative light, and therefore fake news producers limited content which might do so.

It is important to note, however, that the articles in Dataset B represent content that has lasted approximately one and a half years after the election, and the issues of selection bias I discussed previously also apply to these results. Regardless, the absence of unfavorable personal content on Trump parallels prior descriptions of fake news content, and demonstrates intriguing relationships in fake news content production.

The pattern of lower engagement for personal stories reverses when looking at all articles in Dataset B (including the stories labeled as other, instead of being assigned to one candidate). For all articles, the mean number of like and shares for personal stories – 2,306 and 1,792 – exceeds the mean number for non-personal stories by approximately 200 to 300 engagements per article. This is likely due to the inclusion of three articles referring to Obama, coded as “other” in my content analysis.¹⁹ Those three articles were outliers among the personal stories, receiving an average of 7,272 likes as opposed to the average 1,787 likes for all articles. This generated an artificial increase in the average number of likes and shares. Furthermore, the median number of likes and shares for personal stories remains significantly lower than the median number for non-personal stories, providing further evidence that the inclusion of those stories artificially increased the average engagement received for all personal articles.

¹⁹ The three articles alleged Obama plans to move to Canada with his family if Trump won the election. While coded as personal (due to the fact that it references Obama’s private life, and refers directly to him and his family), the stories did not discuss a candidate as the main subject of the article, and were coded as “other.”

Both parametric and non-parametric test were carried out to see if engagement in personal articles differed from other articles in a statistically significant manner. Of those tests, the one which resulted in statically significance at the $\alpha = 0.05$ level was a t-test for difference in means between personal and non-personal articles, using Allcott and Gentzkow's share numbers as the dependent variable. However, the equivalent non-parametric test (Mann-Whitney's U) contradicted the results of the t-test, and indicated no difference in the number of shares received between those two groups. Thus, after eliminating outliers through non-parametric testing, the results are statically insignificant.²⁰

Overall, summary statistics display that personal articles receive less engagement than non-personal articles, providing evidence against my hypothesis. After removing outliers, my tests for statistical differences between those groups generate insignificant results. Thus, I find no evidence to support my hypothesis, and conclude that personal content does not receive more engagement than non-personal content.

H2: *Narratives that contain a lower level of policy discussion receive more engagement than stories that contain more policy discussion.*

As stated above, H2 tests if fake news articles that discuss policy related content receive less engagement than other types of fake news content. Given that individuals are generally apathetic and ignorant when it comes to political issues, I hypothesize that individuals will prefer content that reinforces those views, and is less technical in nature (Pew 2015, Somin 2004, Somin 2010).

²⁰ More importantly, the small N for Dataset B makes detection of a statically significant effect challenging. A larger N dataset would provide more opportunities for parsing out statistically significant results.

Of the 90 articles in Dataset B, only three articles and one narrative presented any policy discussion.²¹ Although there is possible variation in this result (i.e., changes in coding schema, different definitions of what counts as policy, etc.), the conclusion presented by my data is clear; there is minimal policy discussion in fake news articles analyzed in Dataset B.

This lack of policy content has the unfortunate effect of making H2 untestable through statistical comparison. However, it is an intriguing result with several possible explanations. First, fake news producers might struggle to produce “completely false” content discussing policy. Policy content tends to draw significant reporting by mainstream news sources, possibly making it challenging for a false narrative to spread in the relatively crowded space. Additionally, policy articles are frequently associated with information resources (i.e., government reports), making them easier to fact check, and therefore more challenging to fabricate a false narrative on.

Furthermore, the lack of policy content in fake news articles might suggest support for my hypothesis. As stated above, fake news consumers, and the American public more generally, dislike politics and possess low levels of political knowledge. Knowing that fake news consumers will engage with content that confirms their prior biases, and also fits their preferences, fake news producers choose to not manufacture false narratives which contain significant policy discussion. The immediate goal of fake news producers is to have users engage in their article, and find that other content facilitates that end. This idea in particular is consistent with the distribution of substantial and scandal content as described in Table 4 - the vast majority of fake news articles referenced a scandal (mostly related to Clinton) of sorts, rather than

²¹ The narrative that did reference policy was a fabricated story stating that former FBI director Comey received “millions” from the Clinton Foundation as a bribe to not place charges against her. <http://joeforamerica.com/2016/09/fbi-comey-clinton-foundation/>

fabricated a narrative around a substantial issue, of which policy related fake news would likely fall under.

Thus, given the lack of a testable sample size and little variation within my sample, I do not find conclusive support for my hypothesis. However, the lack of fake news articles and narratives referencing policy represents an interesting area for future research, and one which political scientists should explore further.

I have discussed my hypotheses related to the subject matter of a fake news article: the presence of personal content, and the presence of policy content. These variables measure the “what” of fake news content, and how it might affect engagement. I now shift to testing variables which measure the “how” of fake news content – that is, the partisan bias and tone used in an article, and how that might affect engagement. I test first the impact of partisanship on engagement, and then test article tone.

H3: *An increase in partisanship in an article’s content is positively correlated with the number of likes and shares an article receives – as partisanship increases, engagement will also increase, and as partisanship decreases, engagement will also decrease.*

To test my hypothesis on partisan fake news content, I coded partisanship as three variables. First, I coded fake news content for the presence or absence of partisanship. Then, I coded content according to the degree of partisan support. Finally, I coded content according to the degree of partisan opposition.

The presence of partisanship in the articles analyzed was categorized as either containing or not containing partisan content, and therefore coded as a dummy variable. Similar to the presence of policy content in my data, nearly all the articles in both Dataset A and B were coded

as containing partisan content; of the articles coded, there was no content that did not favor one party over another.²²

There are several potential reasons for this result. First, the initial population of articles sampled were in reference to political fake news – the sample excluded other forms of fake news such as urban legends and conspiracy theories, that are less likely to demonstrate partisan bias. Furthermore, fake news producers might choose to create partisan content, believing that partisan content matches user preferences better, and will collect more engagement than neutral articles. Moreover, partisanship is one of the key drivers of directionally motivated reasoning – when seeking out political news, partisanship can greatly bias a user’s preferences and choices. Given that, fake news producers cater to their audiences’ biases, and in political fake news, exploit a reader’s inherent partisanship to maximize article engagement.

Composition of Partisanship:

As stated above, partisanship was coded as a dummy variable. Although this coding enables me to test differences between articles based on the presence or absence of partisan content (or, as is the case here, indicate the overwhelming presence of it), this coding does not measure the direction of an article’s partisanship. To more accurately assess how partisanship manifested itself in fake news content, articles were coded according to two separate variables, one which measured partisan content in support of a party, and another which measured partisan content in opposition to a party.

Analyzing partisanship in this manner is particularly important for content related to 2016. Although many individuals did feel strongly in support or against one candidate, the two candidates had historically low favorability ratings, indicating widespread opposition to both

²² This particular variable received the highest agreement between coders, providing further support for this categorization.

parties (Saad 2016). If fake news content demonstrates strong opposition to one or both of the candidates, it implies that fake news producers were responding to consumer preferences for content and exploiting the hyper-partisan environment for their own ends. The same holds for content in strong support of one of the candidates.

Table 7: Frequency Table for Articles Coded For/Against Candidates

Favor Candidate	<i>Favor Clinton</i>	<i>Mixed/Neutral</i>	<i>Favor Trump</i>	<i>Total</i>
	9% (8)	84% (76)	7% (6)	100% (90)
Oppose Candidate	<i>Oppose Clinton</i>	<i>Mixed/Neutral</i>	<i>Oppose Trump</i>	<i>Total</i>
	63% (57)	5% (5)	31% (28)	100% (90)

Table 7 displays the number of articles coded as for or against a presidential candidate. Most articles (84%), did not explicitly favor one candidate over another, and approximately equal numbers of articles (8 for Clinton and 6 for Trump) demonstrated partisan lean in their favor. In contrast, articles opposing a candidate has an unequal distribution. Articles which oppose Clinton make up 63% of all the articles in Database B (57 out of 90 articles). Meanwhile, articles that oppose Trump only comprise 31% of articles analyzed (28 out of 90 articles).

These results suggest that partisanship manifests itself in fake news content through opposition, rather than support, towards a candidate and that Clinton specifically draws the most partisan opposition. Furthermore, since fake news is primarily created and circulated by the far-right, the composition of partisanship observed in the articles (i.e., anti-Clinton), is representative of bias found more generally in political fake news content.

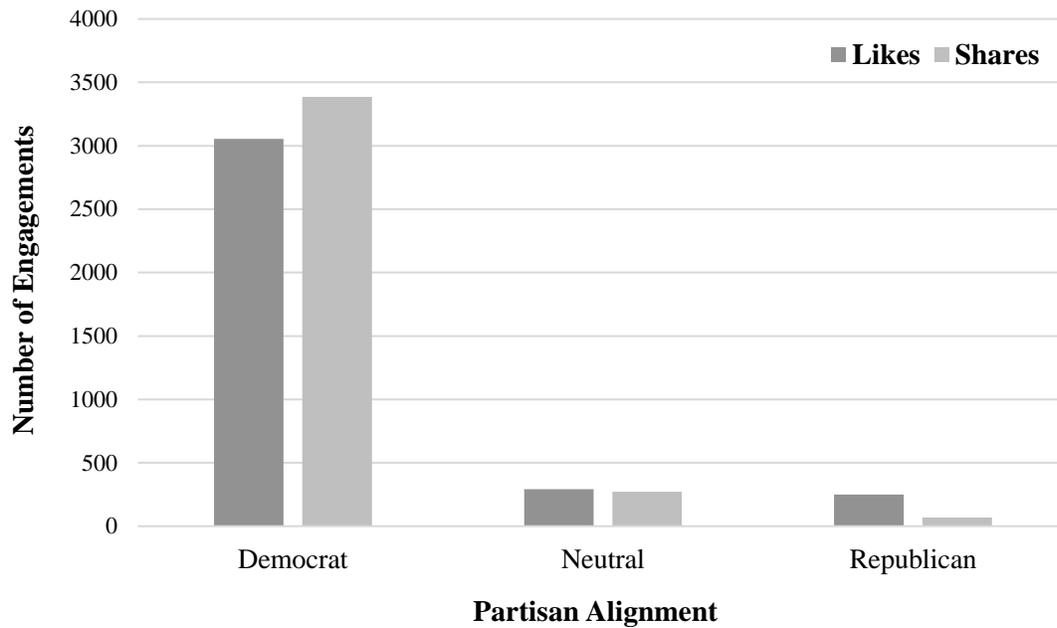
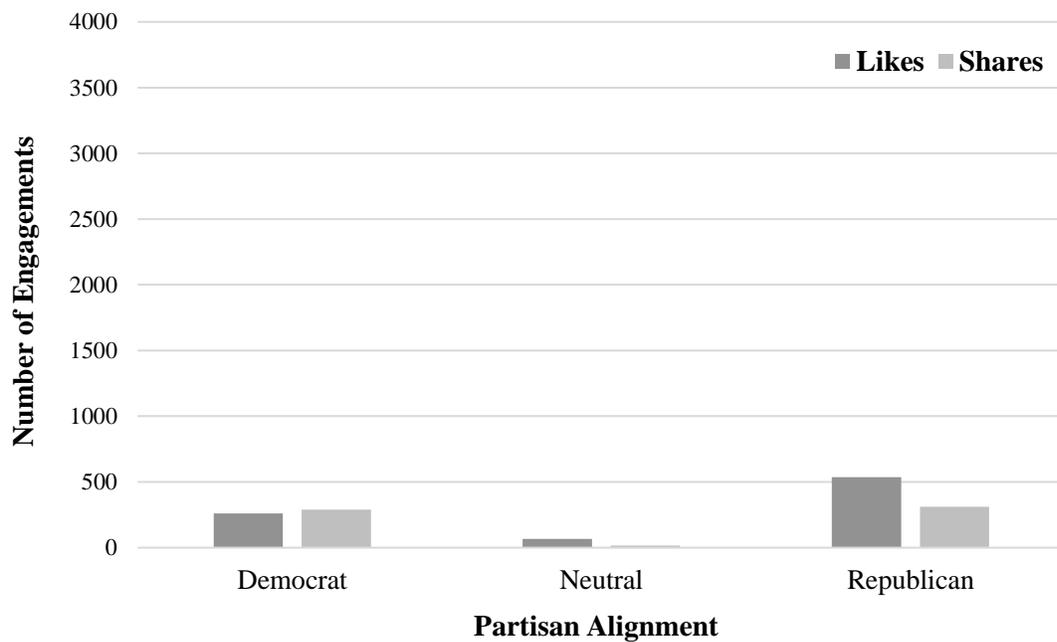
Figure 5: Effect of Partisan Support on Median Engagement per Article**Figure 6:** Effect of Partisan Opposition on Median Engagement per Article

Figure 5 illustrates the median likes and shares for content coded as supporting a candidate in Dataset B, while Figure 6 illustrates those same measures for content opposing a candidate. In examining both figures, articles that demonstrated positive support for Clinton had

a higher median engagement than all other categorizations of partisanship. Although this category contained only eight articles, the use of medians implies this result is not driven by outliers or the sample size of each category. This signals that content demonstrating partisan support for the Democratic Party actually receives more engagement on Facebook than other forms of content, despite comprising only a fraction of total stories.

Kruskal-Wallis tests were used to test for difference in medians between each group for all three engagement measures.²³ The results of these tests display a statistically significant difference between the number of likes and shares for pro-Democratic content, and pro-Republican/neutral content. There were no statistically significant differences between articles which opposed candidates. These results are surprising – whereas pro-Democratic articles comprise a small portion of Dataset B, they receive more engagement than articles which support the Republican party or remain neutral.

However, there are no statistically significant differences between pro-Democratic and Pro-Republican/neutral content when using Allcott and Gentzkow's share numbers as the engagement measure. Therefore, the statically significant difference seen in my share numbers was created sometime after the 2016 general election. This leads to three possibilities: fake news articles which supported Democrats received significantly more shares after the election, articles which supported Trump or remained neutral lost engagement (through the removal of posts and bots by social media platforms), or some combination of the first two options occurred.

Of those three possibilities, the second one – pro-Republican and neutral articles losing engagement – seems most probable. As I discussed earlier, social media platforms have taken

²³ A Kruskal Wallis test was used here as the nonparametric equivalent to an ANOVA test for difference between groups, as the data presented here has categories with an $N < 30$, and with an obviously not normally distributed data.

considerable efforts to curtail fake news dissemination, including deleting posts, pages, and accounts. This would cause the articles which the deleted account/page/etc. engaged with to lose likes and shares. If the targeting of these counter fake news actions were pointed towards pro-Trump fake news content (or anti-Democratic content, which most articles were), a disproportionate loss in engagement would occur in the pro-Republican or neutral articles. Though addressing this question is beyond the scope of my thesis, it presents an interesting challenge for future researchers to investigate.

In sum, almost all fake news content analyzed demonstrated partisan behavior. When partisanship is separated by support for and opposition to a candidate, most of the articles were anti-Democrat, while few articles supported either candidate. Statically significant differences in likes and shares received exist between pro-Democrat and pro-Republican/neutral articles, though not at the time of Allcott and Gentzkow's data collection. These findings provide evidence for the hypothesis proposed, but in relatively unusual manner; furthermore, these findings could be explained by changes which occurred after the study period.

H4: *Increased quantities of negativity in fake news content have a positive correlation with the number of shares an article received.*

Table 8: Engagement Statistics, by Article Tone

Article Tone	Average Likes	Median Likes	Average Shares	Median Shares	Number of Articles	Number of Stories
<i>Positive</i>	438	438	31	31	1	1
<i>Neutral</i>	1,176	537	1,123	588	21	11
<i>Negative</i>	2,088	261	1,721	263	67	32

Table 8 displays the frequency of fake news content with a positive, neutral, or negative tone, along with the mean and median engagement for each category. The mean engagement per article for negative content far exceeds that of articles with neutral/mixed or positive content. On

average, an article with a negative tone received 2,088 likes and 1,721 shares per article, while neutral articles received only 1,176 likes and 1,123 shares per article. Similarly, positive articles received 438 likes and 31 shares per article. This pattern shifts when looking at median engagement, with median likes and shares for neutral articles exceeding that of both positive and negative articles.

When examining mean engagement per article, it appears that negative articles collected the most engagement. However, when looking at median engagement, neutral articles have the greater engagement. This suggests that neutral articles have a more consistent engagement per article rate, as demonstrated by the higher median, and that negative articles contain a few outliers which increase the overall mean for that category.

Correlations run between article tone and shares all came back insignificant, with correlation values around zero. The same results held for Kruskal–Wallis test for difference in medians, meaning I find no difference between different article tones and shares received. Therefore, although there is strong theoretical evidence supporting a connection between article tone and engagement, my data did not exhibit this relationship in a statically significant manner.

Still, my results do not preclude such a relationship; limitations in my data (including a small sample size, amongst other issues) make it difficult to gather the statistical power necessary to demonstrate an effect. Additionally, within my small sample, article tone is unequally distributed. Most articles were classified as negative, while only a single article was categorized as positive.

To examine these results further, I compared the median number of shares as collected by Allcott and Gentzkow for each tone category between Dataset A and B – the only shared engagement metric for both datasets. Since the distribution of shares per article are the same

between datasets (as discussed in Section IV), the relative rank of each tone should also be the same – neutral articles should have the highest median shares. Alternatively, these shares were collected in the period before the general election, and before the extensive corrective action taken by Facebook. As such, this measure of article engagement might generate different results than my data does.

Table 9: Median Shares (A&G) for Negative, Neutral, and Positive Content

Article Tone	Dataset A	Dataset B
<i>Negative</i>	6,700	6,250
<i>Neutral</i>	3,000	4,440
<i>Positive</i>	4,900	4,400

Table 9 compares the median shares as collected by Allcott and Gentzkow for articles in Dataset A and B. In contrast to the results in Table 8, articles with a negative tone have a greater median for shares than articles with a positive or neutral tone. The median number of shares received by negative articles in Dataset B using Allcott and Gentzkow’s data is 6,250 shares per article, while the median for negative articles from my own data is 263 shares per article.

These results support my hypothesis that negative articles receive more engagement, but are not consistent with the results derived from my own data – different measures of engagement produce different results. This is not to say that either result is invalid; as described above, the difference reveals what content before the 2016 election, and what content currently, receives the most shares on Facebook. The share numbers as collected by Allcott and Gentzkow show what article content gathered more engagement during the election period. In contrast, my data is a more accurate representation of current engagement levels. The difference in these shares likely comes from the actions taken by Facebook and others to counter fake news, which were discussed in reference to the other hypotheses.

In addition to the above check, I compared the distribution of article tones in Dataset B (with a sample of 90 articles) with that from Dataset A (260 articles) to see if the issues of distribution in Dataset B drive the findings above.

Table 10: Distribution of Article Tone, by Dataset

Tone	Dataset A	Dataset B
<i>Negative</i>	82% (213)	75% (68)
<i>Neutral</i>	16% (41)	23% (21)
<i>Positive</i>	2% (6)	1% (1)

Table 10 looks at the distribution of articles from Dataset A and B by article tone. The distribution for article tones in Dataset A is comparable to that of Dataset B - most of the articles in Dataset A have a negative tone, similar to Dataset B. Additionally, Kruskal-Wallis tests demonstrate that the distribution of article tones between the two datasets are not significantly different. Thus, the articles sampled in Dataset B are not overly negative; they are approximately the same as those from Dataset A. While this does not remove the issue of bias present in both datasets (fake news articles with active URLs might differ in their content than fake news articles generally), it does reduce concerns that the articles from Dataset B differed from those in Dataset A, and political fake news more generally.

In sum, approximately two-thirds of the articles analyzed had a negative tone, while almost no articles had a positive tone. When comparing mean and median shares within Dataset B, negative articles received on average the highest shares per article, whereas neutral articles had the highest median shares per article. When assessing shares as collected by Allcott and Gentzkow, negative articles have the highest median shares per article.

None of these results above (including difference of means, correlation test, and Mann-Whitney tests) is statically significant, meaning I do not find support for my hypotheses that

negativity correlates with shares per article. Despite the absence of statistical significance, the frequency of negative articles and strong theoretical evidence necessitate further investigation into the relationship between article tone and engagement on Facebook.

I have conducted each of my hypothesis tests, and have examined the impact of each of my characteristics individually. Although my test produce results which are statistically insignificant or have other limitations, I still gathered interesting findings which increase our understanding of fake news content. These include data demonstrating the lack of policy content in fake news, along with the overwhelming presence of partisan and negative content. I now discuss my regression analysis, which provides a multivariate test for my content characteristics, and measures the effect of those characteristics in context.

Regression Analysis:

In addition to difference of means and Mann-Whitney tests, I also conducted OLS regression to test the effect of fake news content on each measure of engagement. Conducting OLS regression enables me to perform a multivariate analysis and examine my independent variables in relation to one another. A fake news article is not just “personal,” or “negative” – an article can demonstrate one, two, or any number of the characteristics I measured. Similar to my use of holistic grading for content analysis, multivariate regression provides a more comprehensive analysis of article content, allowing me to test the effect of a specific variable in context.

To do so, I transformed my dependent variables (likes, shares, and shares from Allcott and Gentzkow) according to a Box-Cox test to correct for non-normality. I then built my regression model, representing each of my content measures as a dummy variable. The intercept of my regression model is also the reference category, or the engagement an article receives

when all other variables equal zero. Here, the reference category is an article which does not discuss a personal characteristic or policy, demonstrates no partisan bias in favor or against a candidate, and has a neutral tone.

Table 11: Table of Coefficients²⁴ (* = $p < 0.05$)

²⁴ Results reported using transformed dependent variables, according to Box-Cox transformations.

<i>Coefficients</i>	Model 1: Shares	Model 2: Likes
	<i>B_i</i> (SD)	<i>B_i</i> (SD)
<i>Intercept</i>	1.073* (0.422)	1.2088* (0.449)
<i>Personal Characteristic</i>	-0.111 (0.166)	-0.044 (0.177)
<i>Policy</i>	-0.291 (0.305)	-0.314 (0.324)
<i>Partisan</i>	0.844* (0.393)	1.068* (0.417)
<i>Pro-Democrat</i>	0.763* (0.215)	0.693* (0.228)
<i>Pro-Republican</i>	0.011 (0.259)	-0.091 (0.276)
<i>Anti-Democrat</i>	0.382 (0.261)	-0.051 (0.278)
<i>Anti-Republican</i>	0.207 (0.283)	-0.088 (0.301)
<i>Positive Tone</i>	-0.774 (0.582)	0.006 (0.619)
<i>Negative Tone</i>	-0.199 (0.140)	-0.182 (0.147)
<i>R²</i>	0.2496	0.212
<i>R² Adjusted</i>	0.1619	0.1199
<i>F Statistic</i>	2.846	2.301
<i>P – Value</i>	0.006	0.02395
<i>N</i>	90	90

Table 11 displays the results of my regression, indicating the effect of article content on the number of likes and shares per article, along with the R^2 , R^2 Adjusted, F statistic, and P-Value for coefficients and regression models. Of all the coefficients (excluding the intercept), only two are statically significant at the $p < 0.05$ level: the presence of partisanship in an article, and partisanship in favor of the Democratic Party. These results hold for both likes and shares.

Articles that were partisan received on average 0.844 more likes and 1.068 more shares than neutral articles, all else held constant. Articles that demonstrated a partisan bent in favor of the Democratic party received on average 0.763 more likes and 0.693 more shares than articles with no partisan alignment in support of one party, all else held constant. I discuss the latter of these two results first, and then discuss the effect of partisanship.

The statistical significance of partisan support for the Democratic party is not surprising, given the statically significant results from the Mann-Whitney test discussed earlier. As mentioned there, this variable was heavily impacted by the presence of outliers, along with possessing a small sample size. Furthermore, this difference was not statistically significant when using Allcott and Gentzkow's shares as the measure of the dependent variable. These caveats limit the extent to which this finding is substantially significant and generalizable.

In contrast, the statistical significance of partisanship overall (i.e., the initial test for the presence of partisanship) is unforeseen, but not unexpected given theoretical evidence and prior descriptions of fake news content. Almost every article demonstrated some form of partisan lean, making that particular hypothesis untestable. However, my regression results display that partisan content receives slightly more likes and shares per article than nonpartisan content. This provides evidence in support of my hypothesis that partisan content receives more engagement on Facebook than nonpartisan content.

For the other variables tested, personal characteristic content, policy related content, and article tone all do not have statically significant effects on article engagement. These results do not eliminate a possible relationship between my content characteristics and engagement; rather, limitations to my research make parsing statically significant results difficult.

First, the relatively small N and non-random sample violates several of the assumptions necessary to performing regression; for instance, I cannot assume that my errors are independently and identically distributed. Additionally, my dependent variables are not normally distributed. Although my transformations mitigate the worst of this non-normality, it is still a limitation to my analysis. Furthermore, I did not include control variables external to my analysis, meaning that I cannot eliminate a possible confounding relationship.

For a more substantially significant regression, I would include controls for the article date published, along with the website the article was published on. I would also include a control for the actor behind the fake content – that is, if the fake content was published by individuals within the U.S., or by an external actor (e.g., Russia), which would control for the resources available to a fake news publisher. While I could not include these controls in my analysis due to the lack of reliable data, I would look to include these or similar measures in later work.

To that end, my regression points to areas of future research. A future study that collected a larger sample size could ameliorate the issues of non-normality and the small N in my variables. Furthermore, it is difficult to achieve a random sample when studying social media and internet content, as data are almost always filtered through algorithms and user bias. While my regression did not use extensive control variables, due to the already large number of parameters and the exploratory nature of this analysis, a more comprehensive regression with more controls might yield different results.

In sum, my regression allowed me to conduct a multivariate analysis of my content characteristics, examining the impact of each in the context of the others. My regression demonstrated that pro-Democratic content and partisan content received more likes and shares

per article than neutral content. Although there were limitations to my regression, these results indicate what differences in content receive more engagement on Facebook.

Discrete Limitations

Although I have disclosed many of the limitations of my research throughout my thesis, there are several limitations which merit more discussion. These include the presence of selection bias in my research, the likely underestimate of the impact of fake news, and the examining of only “false” articles.

First, my research was likely impacted by selection bias, which produced the small number of articles I analyzed throughout. While I addressed this issue earlier, the impact of selection bias is a significant point of concern; it is difficult to know if the fake news articles I analyzed represent the population of fake news produced during the 2016 election. Stories that received more engagements during the 2016 cycle (and therefore are closer matches to consumer preferences) were more likely to have been removed by Facebook between Allcott and Gentzkow’s data collection and my own. Furthermore, it is difficult to test if the removed stories differ substantively from the ones I analyze, hence generating concerns of selection bias. Given the importance of this issue, I take several measures discussed prior to address this, and kept it in mind throughout my analysis.

Furthermore, my analysis is likely an underestimate of the effect of fake news. My research, and political science literature more broadly, operationalizes engagement in content on Facebook through measuring active engagement (e.g., number of “likes”) – or, direct user interaction with Facebook content. Although this is an important measure (engagement leads to increased reach, so this is a valid measure for what content might be spread most broadly) it does not capture the full influence of a fake news article. Active engagement does not measure how

much fake news is in a user's NewsFeed; that is, the information diet of the individual engaging in it. Moreover, active engagement does not measure the indirect influence of fake news that comes from more passive encounters, such as scrolling past a fake news article on your NewsFeed. This is especially important given directionally motivated reasoning, which makes individuals vulnerable to subconscious biases. Though an individual might not engage with a fake news article, it does not mean the headline of an article or the graphic in a post did not have an impact on an individual's opinion. In fact, there is tentative evidence for this strategy. The headlines of fake news articles are longer and are closer to the summary of the article itself rather than a traditional headline, indicating that they are optimized to influence individuals who are "passively scrolling" (Horne & Adali 2017)

My research also underestimates the impact of fake news content by excluding "partly true" articles from my analysis. I included only "fake" stories to maintain the validity of my research, as it allowed for me to control for authenticity, and better measure the variation between content. Additionally, "partly false" news is conceptually vague, as it is difficult to define when a story becomes false, or if two party false stories are each equally false (e.g., is a sensationalized statistic from cable news as false as a misleading article from a far-left website)? However, partly true stories might be more likely to receive engagement, a question I do not address in my research. Additionally, including partly false stories might change the partisan content of fake news, possibly including more far-left, or policy related content.

These are the primary limitations of my thesis; there are others that I mention and discuss throughout (e.g., issues in my content analysis, limitations to the regression). Thus, there are several limitations regarding the representativeness of my analysis, given the selection bias that is likely present in my sample, and the operationalization of fake news as "false." Furthermore,

there are additional limitations as to how valid a measure engagement is as an indicator for the influence of fake news; it likely underestimates the impact of a fake news article. Although I take precautions to minimize the impact of these limitations, such as with the use and comparison of results between multiple datasets and the political science literature more generally, they are present and limit the scope of my analysis. Despite these limitations, my thesis increases our understanding of the impact of fake news content on Facebook engagement, and the nature of the impact during the 2016 U.S. presidential election.

Chapter VI: Conclusion

Social media is a central part of many American lives, and is increasingly relied upon as a source of political information. While social media can facilitate the open discussion of ideas, it can also spread of false and unfounded content. Fake news as “intentionally and verifiably false, with the intent to mislead,” combined with social media, contributed to the negative and confusing discourse during the 2016 presidential election. Knowing this, my thesis sought to address the question of what specific types of fake news content received more engagement on Facebook during the 2016 general election, and thereby increase our understanding of fake news content.

I theorized that the type of content which receives more engagement on Facebook best matches the goals of fake news consumers, fake news producers, and social media platforms. Successful content caters to the directionally motivated reasoning of consumers which drives their search for political information, and also enables producers to maximize ad revenue or influence. Social media platforms control the interaction between producers and consumers, and establishes what producer content a consumer interacts with.

To investigate this theory, I collected data on fake news articles published in the months immediately before the 2016 general election, gathering article content and Facebook engagement metrics. Using content analysis, I coded each article based on different characteristics to test four hypotheses. First, personal fake news content would receive more engagement on Facebook than non-personal content. Second, policy related content would receive less engagement than other fake news content. Third, partisan content would receive more engagement than neutral content. Finally, negative content would receive more shares than neutral or positive content.

Using parametric and non-parametric measures, I tested these hypotheses, utilizing multiple datasets to verify my results when possible. I find no statistically significant results in support of my hypotheses, with my primary limitation being the small sample size of my dataset, and the selection bias that occurred from only using active sites in my analysis. A larger dataset containing more webpages at the time of the 2016 election might find different results, that would support the current literature on this subject. Furthermore, my analysis could not measure passive engagement with fake news (e.g., scrolling past an article on a user's NewsFeed), or the impact of different levels of authenticity in fake news content. Regardless of these limitations, however, my research still produced a number of intriguing results which are useful in informing future research.

First, my data collection contained few, if not zero, articles which were discussing policy, nonpartisan, or positive in tone. Whereas this had the unfortunate effect of making some of my hypotheses untestable, it does support current literature regarding the characteristics of fake news content. If consumers are directionally motivated in consuming fake content, then it follows that content which contains little policy, is partisan, and is negative best appeals to their preference

set, and producers will therefore create content to match that demand. An absence of such content implies fake news producers (combined with social media algorithmic curation) are simply not producing alternative types of fake news content. Although I could not find statistical significance, my results provide a substantial finding that provides support for my hypothesis.

Second, my data collection provides evidence about what types of fake news content can last on social media and the web. The approximately 10% of the population of fake news articles which had both an active URL and Facebook page have managed to maintain both active websites, and active Facebook pages, despite considerable action taken to remove fake news from platforms and web-searches. Although I could not compare the content of active articles to the content of the inactive ones, my research increases our understanding of how many fake news articles survive a year and a half out from their initial publishing, and also what content remains on Facebook following “corrective” action. Future research, if they could collect some measure of the false content of the inactive articles, could compare the two groups and see what fake news content best avoids detection by social media sites.

Finally, my research provides an estimate regarding how much “real” engagement the fake news articles in my dataset received. The initial share numbers as collected in Allcott and Gentzkow’s dataset were orders of magnitude greater than my own, which were collected later. Though this did not pose a problem for my analysis, the absolute difference in shares and the likely difference in content is intriguing, and demonstrates some action was taken to reduce the engagement in these fake news articles over the past year and a half. I cannot state exactly what caused the difference; however, given the size of the reduction and the corresponding news reports on Facebook’s removal of fake accounts and bots, it is possible that most of this loss in engagement comes from the removal of fake and bot accounts. Thus, my share numbers provide

a better estimate of how much real engagement fake news content actually received, and how much of it was driven by automated or fake accounts. This provides an important insight as to how many individuals actually engaged in fake content, and how widely disseminated it actually was.

My research contributes to the literature by increasing our understanding of political fake news during the 2016 election, and what content received more engagement. It also increases our understanding of how fake news works on Facebook. Most importantly, it demonstrates areas which need additional research to better understand fake news, and how to better control its influence. As a democracy, the legitimacy of the U.S. government is derived from its citizens, which in turn must be responsive to their demands. The open debate of ideas in public discourse is essential part of democracy as well, as it is how individuals form opinions with which they base their political participation on. While this process can be contentious, American democracy is designed to work because of, rather than in spite, of these differences in opinions.

However, this public discourse is based on the assumption that the ideas individuals debate and form their opinions from are true. Fake news, by definition, violates that assumption and inserts misinformation and falsehoods into public discourse, enabling individuals to come to conclusions and form political opinions based on falsehoods.

The presence of fake news in political discussion leads to many negative outcomes. Fake news leads to ineffective or inefficient policy – as a democracy, government must be responsive to the wishes of citizens. If those wishes are formed from falsehood, this can cause government to enact policy counter to the optimal social outcome. For instance, the influence of climate change deniers has significantly hindered federal government action in response to phenomena such as global warming, and the consequences from it (Tesler 2018). Another example of

ineffective policy fueled by fake news was Trump's Advisory Commission on Election Integrity, created to investigate voter fraud (characterized specifically as in-person voter fraud). The commission was set up in response to citizen's belief that the 2016 election was influenced by rampant voter fraud, despite no meaningful evidence of the kind. These beliefs regarding the pervasiveness of voter fraud were the result of many different fake news campaigns, but most notably the repeated claim by then-candidate Trump that the election was "rigged" against him (Taylor 2018). Thus, government time and resources were spent to solve an issue with little basis in fact, and where the political benefit far outweighed any societal benefit.

More critically, fake news, and the ways in which it influences opinions, represents a real threat to the legitimacy of the U.S. government. As stated above, the legitimacy of democracies comes from its citizens; while dissent against government is a healthy part of democracy, dissent based on false information is not, and serves to undermine the authority and effectiveness of governance. It serves to weaken the U.S., and make it more difficult for the country to respond in times of real crisis. This concept was what motivated the Russian government in their misinformation campaign in 2016 – as stated in the intelligence community report, the goal of the Russian government was to “to undermine public faith in the US democratic process”(Intelligence Community 2018; 1) which they did through spreading false political information. Thus, fake news undermines the legitimacy of democratic institutions, posing a real threat to American democracy.

Given the real and pressing concerns fake news poses to U.S. democracy, it is crucial that researchers continue to investigate how fake news is spread, and how best to combat its influence. To that end, there are many possible areas for future researchers to examine.

First, more research of fake news on Facebook is needed, rather than on Twitter; while Twitter offers a wealth of (free and easily accessible) data, many more Americans use Facebook (69% as opposed to 23%) (Pew Factsheet 2018). Despite challenges, scholars need to place emphasis on researching the more widely used platform. Future research should also examine videos and pictures of fake news; while visual elements were not the focus of my analysis, the reliance of fake news on pictures, graphics, syntax, and other cues warrants further investigation by scholars. Additionally, an analysis into the content differences between “false” and “partly false,” or other classifications would also be welcomed; if the best lies always have a kernel of truth (as the saying goes), then these partly false stories might be a more effective vector for transmitting misinformation. Moreover, a study conducted using my methods, except with more articles, a more recent data collection and more resources, might yield different results. This list is not comprehensive; there is much we need to study and test to better understand political fake news, and social media communication more generally.

Understanding how and what fake news content provokes reactions on social media is an important question to answer, for it will better enable social media companies and others to stop or limit its spread. Although there are many areas left to explore, my research study addresses this question, and provides insight into the connection between fake news content and social media during the 2016 election cycle.

Appendix A: Content Analysis Rubric

H1	<i>Personal Story/Not Personal Story</i>	
<i>Scale</i>	0 = not a personal story	1 = personal story

H2:	<i>Policy Article/Not policy oriented</i>	
<i>Scale</i>	0 = no policy	1 = some policy

H3: A	<i>Partisan/Not Partisan</i>				
<i>Scale</i>	0 = non-Partisan		1 = Partisan		
B	<i>Pro Candidate</i>				
<i>Scale</i>	1 = Pro Democrat	2 = lean Democrat	3 = neutral/mixed	4 = lean GOP	5 = Pro GOP
C	<i>Anti Candidate</i>				
<i>Scale</i>	1 = Anti Democrat	2 = lean anti Democrat	3 = neutral/mixed	4 = lean anti GOP	5 = Anti GOP

H4:	<i>Article Tone</i>				
<i>Scale</i>	1 = positive	2 = lean positive	3 = neutral/mixed	4 = lean negative	5 = negative

Candidate	<i>Clinton</i>	<i>Trump</i>	<i>Other</i>
<i>Scale</i>	0	1	2

Subject	<i>Substantive</i>	<i>Scandal</i>
<i>Scale</i>	0	1

Substantive	<i>Jobs</i>	<i>Immigration</i>	<i>Trade</i>	<i>Taxes</i>	<i>Muslims</i>	<i>Policing</i>
(13 = "other")	1	2	3	4	5	6
	<i>Race</i>	<i>Guns</i>	<i>Healthcare</i>	<i>Women's Rights</i>	<i>Climate Change</i>	<i>Education</i>
	7	8	9	10	11	12

Scandal		<i>Foundation</i>	<i>Emails</i>	<i>Benghazi</i>	<i>Russia</i>	<i>Other</i>	<i>Wikileaks</i>	<i>Health</i>
<i>Scale</i>	<i>HRC</i>	1	2	3	4	5	6	7
	<i>DJT</i>	<i>Foundation</i>	<i>University</i>	<i>Russia</i>	<i>Women</i>	<i>Other</i>		
		1	2	3	4	5		

Bibliography

- Allcott, Hunt and Gentzkow, Matthew. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31, no. 2 (Spring 2017): 211-36. Accessed October 31, 2017. doi=10.1257/jep.31.2.211.
- 100PercentFedUp.com. "CAN HILLARY LIE HER WAY OUT OF THIS ONE? PHYSICIAN Says Hillary Has Parkinson's Disease...Hillary Admits She Couldn't Even "Get Up" After Convention." 100PercentFedUp. Last modified 2016. Accessed November 16, 2018. <https://100percentfedup.com/can-hillary-lie-way-one-physician-says-hillary-parkinsons-disease-hillary-admits-couldnt-even-get-convention/>.
- Adair, Bill, and Angie Drobnic Holan. "The Principles of PolitiFact, PunditFact and the Truth-O-Meter." PolitiFact. Last modified November 1, 2013. Accessed November 28, 2017. <http://www.politifact.com/truth-o-meter/article/2013/nov/01/principles-politifact-punditfact-and-truth-o-meter/>.
- Agwin, Julia, Surya Mattu, and Terry Parris, Jr. "Facebook Doesn't Tell Users Everything It Really Knows About Them." ProPublica. Last modified December 22, 2016. Accessed November 28, 2017. <https://www.propublica.org/article/facebook-doesnt-tell-users-everything-it-really-knows-about-them>.
- Applebaum, Anne. "President Trump is now a troll." *The Washington Post* (Washington, DC), November 30, 2016. Accessed December 11, 2017. https://www.washingtonpost.com/opinions/global-opinions/our-president-is-an-online-troll/2017/11/30/42345a86-d5f3-11e7-b62d-d9345ced896d_story.html?utm_term=.7e03a14d19a3.
- Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. "Exposure to ideologically diverse news and opinion on Facebook." *Science* 348, no. 6239: 1130-32. Accessed May 3, 2018. DOI:10.1126/science.aaa1160.
- Balbinus, Cornelius. "Trump to Deport Lin-Manuel Miranda, Doesn't Realize Puerto Rico is US TerritoryNew Roman Times." New Roman Times. Last modified November 1, 2016. Accessed November 17, 2018. <https://www.newromantimes.com/trump-deport-lin-manuel-miranda-doesnt-realize-puerto-rico-us-territory/>.
- Barthel, Michael, Amy Mitchell, and Jesse Holcomb. *Many Americans Believe Fake News Is Sowing Confusion*. December 16, 2016. Accessed May 2, 2018. <http://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/>.
- Bessi, Alessandro, and Emilio Ferrara. "Social Bots Distort the 2016 US Presidential Election Online Discussion." *First Monday* 21, no. 11 (November 7, 2016). Accessed October 31, 2017. <https://ssrn.com/abstract=2982233>.

- Blake, Aaron. "Welcome to the Next, Most Negative Presidential Election of Our Lives." *Washington Post* (Washington, DC), July 29, 2016, The Fix. Accessed November 28, 2017. https://www.washingtonpost.com/news/the-fix/wp/2016/07/29/clinton-and-trump-accept-their-nominations-by-telling-you-what-you-should-vote-against/?utm_term=.c914a3443b69.
- Boyd, Danah M., and Nicole B. Ellison. "Social Network Sites: Definition, History, and Scholarship." *Journal of Computer-Mediated Communication* 13, no. 1 (October 1, 2007): 210-30. Accessed November 2, 2018. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>.
- Brady, William J., Julian A. Wills, John T. Jost, Joshua A. Tucker, and JJ Van Bavel. "Moral contagion in social networks." *Proceedings of the National Academy of Sciences*, June 2017. Accessed November 2, 2018. <https://doi.org/10.1073/pnas.1618923114>.
- Bullock, John. "The Enduring Importance of False Political Beliefs." *Conference Papers -- Western Political Science Association*, 2006, 1-49. Accessed May 3, 2018.
- Bump, Philip. "How Hillary Clinton and Donald Trump became historically unpopular presidential nominees." *The Washington Post* (Washington, DC), June 16, 2016, The Fix. Accessed November 2, 2018. https://www.washingtonpost.com/news/the-fix/wp/2016/06/16/how-hillary-clinton-and-donald-trump-became-historically-unpopular-presidential-nominees/?utm_term=.7999ac4067da.
- Burroughs, Benjamin. "Obama Trolling: Memes, Salutes and an Agonistic Politics in the 2012 Presidential Election." *The Fibreculture Journal*, no. 22 (2013): 258-77. Accessed October 7, 2017. <http://fibreculturejournal.org/wp-content/pdfs/FCJ-165Benjamin%20Burroughs.pdf>.
- Centre for Applied Research, Estonian National Defense College. *Russian Information Campaign Against Ukrainian State And Defence Forces*. Accessed May 2, 2018. <https://www.stratcomcoe.org/russian-information-campaign-against-ukrainian-state-and-defence-forces>.
- Chakraborty, Abhijnan, Paranjape Bhargavi, Sourya Kakarla, and Niloy Ganguly. "Stop Clickbait: Detecting and Preventing Clickbaits in Online News Media." *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, October 31, 2016. Accessed November 2, 2018. <https://arxiv.org/abs/1610.09786>.
- Chaykowski, Kathleen. "Mark Zuckerberg: 2 Billion Users Means Facebook's 'Responsibility Is Expanding.'" *Forbes Tech*. Last modified June 27, 2017. Accessed May 3, 2018. <https://www.forbes.com/sites/kathleenchaykowski/2017/06/27/facebook-officially-hits-2-billion-users/#4795b9643708>.

- Coldewey, Devin. "Here are the Russian Ads that Deceived Users on Facebook and Instagram." Tech Crunch. Last modified November 1, 2017. Accessed May 2, 2018. <https://techcrunch.com/gallery/here-are-15-of-the-russian-bought-ads-aimed-at-influencing-the-election/slide/11/>.
- Cook, John, Stephan Lewandowsky, and Ulrich K.H. Ecker. "Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence." *PLoS ONE* 12, no. 5 (May 5, 2017). Accessed October 31, 2017. . <https://doi.org/10.1371/journal.pone.0175799>.
- Crider, Vanessa. "BREAKING: Hillary Clinton VINDICATED After WikiLeaks Caught FABRICATING Fake Emails (DETAILS)." Bipartisan Report. Last modified October 12, 2016. Accessed November 16, 2018. <http://bipartisanreport.com/2016/10/12/breaking-hillary-clinton-vindicated-after-wikileaks-caught-fabricating-fake-emails-details/>.
- Davis, Wynne. "Fake Or Real? How To Self-Check The News And Get The Facts." npr All Tech Considered. Last modified December 5, 2016. Accessed May 2, 2018. <https://www.npr.org/sections/alltechconsidered/2016/12/05/503581220/fake-or-real-how-to-self-check-the-news-and-get-the-facts>.
- Difonzo, Nicholas, and Prashant Bordia. *Rumor psychology: Social and organizational approaches*. Vol. 1. Washington, DC: American Psychological Association, 2007. Accessed May 2, 2018. DOI:10.1037/11503-000.
- DiGrazia, J., K. McKelvey, J. Bollen, and F. Rojas. "More Tweets, More Votes: Social Media as a Quantitative Indicator of Political Behavior." *PLoS ONE* 8, no. 11 (November 2013). Accessed October 31, 2017. doi:10.1371/journal.pone.0079449.
- Domonoske, Camila. "Students Have 'Dismaying' Inability To Tell Fake News From Real, Study Finds." npr The Two Way. Last modified November 23, 2016. Accessed May 2, 2018. <https://www.npr.org/sections/thetwo-way/2016/11/23/503129818/study-finds-students-have-dismaying-inability-to-tell-fake-news-from-real>.
- Donath, Judith S. "Identity and Deception in the Virtual Community." In *Communities in Cyberspace*, by Marc A. Smith and Peter Kollock, 29-55. London: Routledge, 2006. Accessed December 11, 2017. https://wiki.cc.gatech.edu/scqualifier/images/d/db/Donath-Identity_and_deception.pdf.
- Einstein, Katherine, and David Glick. "Do I Think BLS Data are BS? The Consequences of Conspiracy Theories." *Political Behavior* 37, no. 3 (September 2015): 679-701. Accessed May 3, 2018. doi:10.1007/s11109-014-9287-z.
- Eslami, Motahhare, Aimee Rickman, Kristen Vaccaro, Aleyasen Amirhossein, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. "I always assumed that I wasn't really that close to [her]": Reasoning about Invisible Algorithms in News Feeds." *CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in*

- Computing Systems*, April 2015, 153-62. Accessed November 2, 2018. doi>10.1145/2702123.2702556.
- Evans, David S. "The Online Advertising Industry: Economics, Evolution, and Privacy." Abstract. *Journal of Economic Perspectives* 23, no. 3 (Summer 2009): 37-60. Accessed May 3, 2018. <https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.23.3.37>.
- Facebook. "How News Feed Works." Facebook Help Center. Last modified 2018. Accessed November 2, 2018. <https://www.facebook.com/help/1155510281178725>.
- . "Success Stories." Facebook Business. Last modified 2018. Accessed May 3, 2018. <https://www.facebook.com/business/success/categories/custom-audiences>.
- Facebook Business. "Making Page Likes More Meaningful." Facebook Business. Last modified March 5, 2015. Accessed November 2, 2018. <https://www.facebook.com/business/news/page-likes-update>.
- Faris, Rob, Hal Roberts, Bruce Etling, Nikki Bourassa, Ethan Zuckerman, and Yochai Benkler. *Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election*. August 16, 2016. Accessed November 2, 2018. <https://cyber.harvard.edu/publications/2017/08/mediacloud>.
- FBI Department of Justice. *Indictment*. By Robert S. Mueller, III. Publication no. 1:18-cr-00032-DLF. Washington, DC: Government Printing Office, 2018. Accessed May 3, 2018. <https://www.justice.gov/file/1035477/download>.
- Fitzgerald, Mary Ann. "Misinformation on the Internet: Applying evaluation skills to online information." *Emergency Librarian* 24, no. 3 (January/February 1997): 9-14. Accessed May 2, 2018. <https://search.proquest.com/openview/7e79afbc6d89c33838cc4e9eff647625/1?pq-origsite=gscholar&cbl=38018>.
- Fourney, Adam, Miklos Z. Racz, Gireeja Ranade, Markus Mobius, and Eric Horvitz. "Geographic and Temporal Trends in Fake News Consumption During the 2016 US Presidential Election." In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, compiled by Ee-Peng Lim, 2071-74. New York, NY: ACM, 2017. Accessed October 14, 2018. <http://doi>10.1145/3132847.3133147>.
- Frank, Russell. "Caveat Lector: Fake News as Folklore." *The Journal of American Folklore* 128, no. 509 (Summer 2015): 315-32. Accessed October 31, 2017. DOI:10.5406/jamerfolk.128.509.0315.
- Freedom's Final Stand. "SHOKING: IT'S OVER: Hillary's ISIS Email Just Leaked & It's Worse Than Anyone Could Have Imagined...." Freedom's Final Stand. Accessed November 16, 2018. <http://freedomfinalstand.com/shoking-its-over-hillarys-isis-email-just-leaked-its-worse-than-anyone-could-have-imagined/>.

- Fridkin, Kim, Patrick Kinney, and Amanda Wintersieck. "Liar, Liar, Pants on Fire: How Fact-Checking Influences Citizens' Reactions to Negative Advertising." *Journal of Political Communication* 32, no. 1 (February 2015): 127-51. Accessed October 31, 2017. <http://dx.doi.org/10.1080/10584609.2014.914613>.
- Fridkin, Kim L., and Patrick J. Kenney. "The Role of Candidate Traits in Campaigns." *The Journal of Politics* 73, no. 1 (January 14, 2011): 61-73. <http://www.jstor.org/stable/10.1017/s0022381610000861>.
- Geiger, Abigail. *For Many Voters, It's Not Which Presidential Candidate They're For But Which They're Against*. September 2, 2016. Accessed November 28, 2017. <http://www.pewresearch.org/fact-tank/2016/09/02/for-many-voters-its-not-which-presidential-candidate-theyre-for-but-which-theyre-against/>.
- Gentzkow, Matthew. "Polarization in 2016." *TAP* (blog). Entry posted March 18, 2016. Accessed November 28, 2017. http://www.techpolicy.com/Gentzkow_PolarizationIn2016.aspx.
- Georgiou, Eleni, Marios D. Dikaiakos, and Athena Stassopoulou. "On the properties of spam-advertised URL addresses." *Journal of Network and Computer Applications* 31, no. 4 (November 2008): 966-85.
- Gillespie, Tarleton. "The Relevance of Algorithms." In *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot. The MIT Press, 2014. MIT Press Scholarship Online, 2014. doi: 10.7551/mitpress/9780262525374.003.0009.
- Goforth, Chelsea. "Using and Interpreting Cronbach's Alpha." University of Virginia Library. Last modified November 16, 2015. Accessed November 2, 2018. <https://data.library.virginia.edu/using-and-interpreting-cronbachs-alpha/>.
- Hannak, A., D. Margolin, B. Keegan, and I. Weber. "Get Back! You Don't Know Me Like That: The Social Mediation of Fact Checking Interventions in Twitter Conversations." *International Conference on Weblogs and Social Media*, 8th ser., 2014, 187-96. Accessed October 31, 2017. file:///C:/Users/Grace/Downloads/8115-37746-1-PB.pdf.
- Hart, Kim. "Here are the election Facebook ads Russia bought." Axios. Last modified November 1, 2017. Accessed May 3, 2018. <https://www.axios.com/here-are-the-election-facebook-ads-russia-bought-1513306606-f8d07e39-085f-44e7-9958-975d35d3b4f0.html>.
- Hawkins, "Is Chavez Populist? Measuring Populist Discourse in Comparative Perspective," *Comparative Political Studies* 42 (August 2009): 1040-67

- Heikkilä, Niko. "Online Antagonism of the Alt-Right in the 2016 Election." *European Journal of American Studies* 12, no. 2 (2017). Accessed October 7, 2017. <http://ejas.revues.org/12140>.
- Hochschild, Jennifer L. "Do Facts Matter? Information and Misinformation in American Politics." *Political Science Quarterly* 130, no. 4 (2015): 585-624. Accessed November 2, 2018. <https://doi.org/10.1002/polq.12398>.
- Hoft, Jim. "AWFUL! AP Shills for Hillary Clinton After She Calls Bernie Supporters a “Bucket of Losers” in Goldman Sachs Speech." The Gateway Pundit. Last modified October 9, 2016. Accessed November 17, 2018. <https://www.thegatewaypundit.com/2016/10/awful-ap-shills-hillary-clinton-calls-bernie-supporters-bucket-losers-goldman-sachs-speech/>.
- Horne, Benjamin D., and Sibel Adali. "This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News." *The 2nd International Workshop on News and Public Opinion at ICWSM*, March 28, 2017. Accessed November 2, 2018. <http://arxiv.org/abs/1703.09398>.
- Horrigan, John B. *Information Overload*. December 7, 2016. Accessed May 3, 2018. http://assets.pewresearch.org/wp-content/uploads/sites/14/2016/12/05162316/PI_2016.12.07_Information-Overload_FINAL.pdf.
- Howard, Philip, Gillian Bolsover, Bence Kollanyi, Lisa-Maria Neudert, and Samantha Bradshaw. *Junk News and Bots during the U.S. Election: What Were Michigan Voters Sharing Over Twitter*. Research report no. 2017.1. March 26, 2017. Accessed October 31, 2017. <http://275rzy1ul4252pt1hv2dqyuf.wpengine.netdna-cdn.com/wp-content/uploads/2017/07/2206.pdf>.
- Howard, Philip N., Bence Kollanyi, Samantha Bradshaw, and Lisa-Maria Neudert. *Social Media, News and Political Information during the US Election: Was Polarizing Content Concentrated in Swing States*. August 2017. Accessed October 31, 2017. <http://comprop.oii.ox.ac.uk>.
- Hughes, Adam, and Onyi Lam. "Highly ideological members of Congress have more Facebook followers than moderates do." Pew Research Center. Last modified August 21, 2017. Accessed November 2, 2018. <http://www.pewresearch.org/fact-tank/2017/08/21/highly-ideological-members-of-congress-have-more-facebook-followers-than-moderates-do/>.
- Intelligence Community. *Assessing Russian Activities and Intentions in Recent US Elections*. Report no. ICA 2017-01D. January 6, 2017. Accessed October 31, 2017. https://www.dni.gov/files/documents/ICA_2017_01.pdf.
- Ipsos, and BuzzFeed. *IPSOS PUBLIC AFFAIRS: BuzzFeed Fake News*. December 6, 2016. Accessed May 2, 2018. <https://www.ipsos.com/en-us/news-polls/ipsosbuzzfeed-poll-fake-news>.

- Kahne, Joseph, and Benjamin Bowyer. "Educating for Democracy in a Partisan Age Confronting the Challenges of Motivated Reasoning and Misinformation." *American Educational Research Journal* 54, no. 1 (February 2017). Accessed May 3, 2018. <http://journals.sagepub.com/doi/full/10.3102/0002831216679817>.
- Kardas, Tuncay. "Trump and the Rise of the Media-Industrial Complex in American Politics." *Insight Turkey* 19, no. 3 (Summer 2017): 93-120. Accessed October 8, 2017. DOI:10.25253/99.2017193.08.
- Keating, Dan, Kevin Schaul, and Leslie Shapiro. "The Facebook ads Russians targeted at different groups." *The Washington Post* (Washington, DC), November 1, 2017, Business. Accessed May 3, 2018. https://www.washingtonpost.com/graphics/2017/business/russian-ads-facebook-targeting/?utm_term=.b7fa0b3d30c5.
- Krippendorff, Klaus. "Validity in Content Analysis." In *Computerstrategien für die Kommunikationsanalyse*, 69-112. Frankfurt, Germany, 1980. Accessed November 2, 2018. https://repository.upenn.edu/asc_papers/291/.
- Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F. Rich. "Misinformation and the Currency of Democratic Citizenship." *The Journal of Politics* 62, no. 3 (August 2000): 790-816. Accessed May 2, 2018. <http://www.jstor.org/stable/2647960>.
- LaCapria, Kim. "Chuck E. Sleaze." Snopes. Last modified November 21, 2016. Accessed November 28, 2017. <https://www.snopes.com/pizzagate-conspiracy/>.
- Lacy, Stephen, Brendan R. Watson, Daniel Riffe, and Jennette Lovejoy. "Issues and Best Practices in Content Analysis." *Journalism & Mass Communication Quarterly* 92, no. 4 (September 28, 2015): 791-811. Accessed November 2, 2018. <https://doi.org/10.1177/1077699015607338>.
- Lazer, David, Matthew Baum, Nir Grinberg, Lisa Friedland, Kenneth Joseph, Will Hobbs, and Carolina Mattsson. "Combating Fake News: An Agenda for Research and Action." Paper presented at Combating Fake News: An Agenda for Research and Action; Sponsored by Harvard Kennedy School and Northeastern University, Harvard University, Cambridge, MA, February 2017. Accessed October 14, 2018. <https://shorensteincenter.org/combating-fake-news-agenda-for-research/>.
- Lewandowsky, Stephan, Ullrich K. H. Ecker, Colleen M. Seifert, Norbert Schwarz, and John Cook. "Misinformation and Its Correction Continued Influence and Successful Debiasing." Abstract. *Psychological Science in the Public Interest* 13, no. 3 (September 17, 2012): 106-31. Accessed May 3, 2018. <https://doi.org/10.1177/1529100612451018>.

- Lewis, Kevin, Marco Gonzalez, and Jason Kaufman. "Social selection and peer influence in an online social network." *PNAS* 109, no. 1 (January 12, 2012): 68-72. Accessed May 3, 2018. <https://doi.org/10.1073/pnas.1109739109>.
- Lodge, Milton, Kathleen M. McGraw, and Patrick Stroh. "An Impression-Driven Model of Candidate Evaluation." *The American Political Science Review* 83, no. 2 (June 1989): 399-419. Accessed November 28, 2017. <http://links.jstor.org/sici?sici=0003-0554%28198906%2983%3A2%3C399%3AAIMOCE%3E2.0.CO%3B2-5>.
- Malmgren, Evan. "Don't Feed the Trolls." *Dissent*, Spring 2017. Accessed October 7, 2017. <https://www.dissentmagazine.org/article/dont-feed-the-trolls-alt-right-culture-4chan>.
- McBeth, Mark K., and Randy S. Clemons. "Is Fake News the Real News? The Significance of Stewart and Colbert for Democratic Discourse, Politics, and Policy." In *The Stewart/Colbert Effect: Essays on the Real Impacts of Fake News*, edited by Amarnath Amarasingam, 79-93. N.p.: McFarland & Company Inc, 2011.
- McCombs, Maxwell E., and Donald L. Shaw. "The Agenda-Setting Function of Mass Media." Abstract. *The Public Opinion Quarterly* 36, no. 2 (Summer 1972). Accessed May 2, 2018. <https://www.jstor.org/stable/2747787>.
- Mitchelstein, Eugenia, and Pablo J. Boczkowski. "New Media, News Production and Consumption." In *International Encyclopedia of the Social & Behavioral Sciences*, 793-97. Second ed. 2015. Last modified March 12, 2015. Accessed May 3, 2018. <https://doi-org.ezproxy.lib.vt.edu/10.1016/B978-0-08-097086-8.95083-9>.
- Narayanan, Vidya, Vlad Barash, John Kelly, Bence Kollanyi, Lisa-Maria Neudert, and Philip N. Howard. Memorandum, "Polarization, Partisanship and Junk News Consumption over Social Media in the US," February 6, 2018. Accessed May 3, 2018. <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/02/Polarization-Partisanship-JunkNews.pdf>.
- Nyhan, B., J. Reifler, and PA Ubel. "The hazards of correcting myths about health care reform." *Med Care* 51, no. 2 (February 2013): 127-32. Accessed October 31, 2017. doi:10.1097/MLR.0b013e318279486b.
- . "The hazards of correcting myths about health care reform." *Med Care* 51, no. 2 (February 2013): 127-32. Accessed October 31, 2017. doi:10.1097/MLR.0b013e318279486b.
- Nyhan, Brendan, and Jason Reifler. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32, no. 2 (June 2010): 303-30. Accessed May 3, 2018. doi:10.1007/s11109-010-9112-2.

- . "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32, no. 2 (June 2010): 303-30. Accessed May 3, 2018. doi:10.1007/s11109-010-9112-2.
- Office of the Deputy Attorney General. "Appointment Of Special Counsel To Investigate Russian Interference With The 2016 Presidential Election And Related Matters." News release. May 17, 2017. Accessed May 3, 2018. <https://www.justice.gov/opa/press-release/file/967231/download>.
- Ohlheiser, Abby. "'We actually elected a meme as president': How 4chan celebrated Trump's victory." *The Washington Post* (Washington, DC), November 9, 2016, The Intersect. Accessed December 11, 2017. https://www.washingtonpost.com/news/the-intersect/wp/2016/11/09/we-actually-elected-a-meme-as-president-how-4chan-celebrated-trumps-victory/?utm_term=.dee7bb4388c5.
- O'Sullivan, Donie, and Dylan Byers. "Exclusive: Fake black activist accounts linked to Russian government." CNN Money. Last modified September 28, 2017. Accessed May 2, 2018. <http://money.cnn.com/2017/09/28/media/blacktivist-russia-facebook-twitter/index.html>.
- Palma, Bethania. "Hillary Clinton 'Accidentally' Gave ISIS \$400 Million." Snopes. Last modified November 3, 2016. Accessed November 2, 2018. <https://www.snopes.com/fact-check/hillary-clinton-accidentally-gave-isis-400-million/>.
- Pew Research Center. "Social Media Fact Sheet." Pew Research Center. Last modified February 5, 2018. Accessed November 2, 2018. <http://www.pewinternet.org/fact-sheet/social-media/>.
- Pew Research Center. *What the Public Knows — In Pictures, Words, Maps and Graphs*. April 28, 2015. Accessed November 2, 2018. <http://www.people-press.org/2015/04/28/what-the-public-knows-in-pictures-words-maps-and-graphs/>.
- Rader, Emilee, and Rebecca Gray. "Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, edited by Bo Begole, 173-82. New York, NY: ACM, 2015. Accessed November 2, 2018. <https://doi.org/10.1145/2702123.2702174>.
- Ratkiewicz, J., M. D. Conover, M. Meiss, B. Goncalves, A. Flammini, and F. Menczer. "Detecting and Tracking Political Abuse in Social Media." *International AAAI Conference on Weblogs and Social Media*, 5th ser., 2011, 297-304. Accessed October 31, 2017. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewFile/2850/3274>.
- Rosen, Guy. "Facebook Publishes Enforcement Numbers for the First Time." Facebook Newsroom. Last modified May 15, 2018. Accessed November 2, 2018. <https://newsroom.fb.com/news/2018/05/enforcement-numbers/>.

- Ryan, Timothy. "What Makes Us Click? Demonstrating Incentives for Angry Discourse with Digital-Age Field Experiments." *The Journal of Politics* 74, no. 4 (October 2012): 1138-52. Accessed November 2, 2018. <https://doi.org/10.1017/S0022381612000540>.
- Saad, Lydia. "Trump and Clinton Finish With Historically Poor Images." Gallup. Last modified November 8, 2016. Accessed November 2, 2018. <https://news.gallup.com/poll/197231/trump-clinton-finish-historically-poor-images.aspx>.
- Seaver, Nick. "Knowing Algorithms." Unpublished manuscript, MIT, February 2014. Accessed November 28, 2017. <https://static1.squarespace.com/static/55eb004ee4b0518639d59d9b/t/55ece1bfe4b030b2e8302e1e/1441587647177/seaverMIT8.pdf>.
- Selyukh, Alina, and Tim Mak. "Facebook Shuts 652 Iran-Backed Accounts Linked In Global Disinformation Campaign." National Public Radio. Last modified August 21, 2018. Accessed November 2, 2018. <https://www.npr.org/2018/08/21/640696857/facebook-shuts-down-652-iranian-backed-and-some-russian-backed-accounts>.
- Shane, Scott, and Mike Issac. "Facebook's Russia-Linked Ads Came in Many Disguises." *The New York Times* (New York, NY), October 2, 2017, Technology. Accessed May 3, 2018. <https://www.nytimes.com/2017/10/02/technology/facebook-russia-ads-.html>.
- Shearer, Elisa, and Jeffery Gottfried. *News Use Across Social Media Platforms 2017*. September 7, 2017. Accessed November 2, 2018. <http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/>.
- Silver, Nate. "Donald Trump Is The World's Greatest Troll." FiveThirtyEight. Last modified July 20, 2015. Accessed December 11, 2017. <https://fivethirtyeight.com/features/donald-trump-is-the-worlds-greatest-troll/>.
- Silverman, Craig, ed. "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook." BuzzFeedNews. Last modified November 16, 2016. Accessed October 28, 2018. <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook>.
- Silverman, Craig, and Lawrence Alexander. "How Teens In The Balkans Are Duping Trump Supporters With Fake News." BuzzFeed. Last modified November 3, 2016. Accessed October 31, 2017. https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.abV11bGBE#.fo1lz43P8.
- Somin, Ilya. "Deliberative Democracy and Political Ignorance." *Critical Review* 22, nos. 2-3 (2010): 253-79. Accessed November 28, 2017. <https://doi.org/10.1080/08913811.2010.508635>.

- . *When Ignorance Isn't Bliss How Political Ignorance Threatens Democracy*. Research report no. 525. Cato Institute, 2004. Accessed November 28, 2017. <https://object.cato.org/sites/cato.org/files/pubs/pdf/pa525.pdf>.
- Starbird, Kate. "Starbird, Kate. "Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter." ICWSM (2017)." *ICWSM*, 9th ser., 2017, 1-10. Accessed October 31, 2017. http://faculty.washington.edu/kstarbi/Alt_Narratives_ICWSM17-CameraReady.pdf.
- Steward, Emily. "Trump buys into the conspiracy that social media censors conservatives." *Vox*. Last modified August 18, 2018. Accessed November 2, 2018. <https://www.vox.com/policy-and-politics/2018/8/18/17749450/trump-twitter-bias-alex-jones-infowars>.
- Stieglitz, Stefan, and Linh Dang-Xuan. "Political Communication and Influence through Microblogging--An Empirical Analysis of Sentiment in Twitter Messages and Retweet Behavior." In *2012 45th Hawaii International Conference on System Science: (HICSS) ; USA, 4-7 Jan. 2012*, edited by Ralph H. Sprague. Piscataway, NJ: IEEE, 2012. Accessed November 2, 2018. <https://ieeexplore.ieee.org/abstract/document/6149247>.
- Stinson, Elizabeth. "Facebook Doesn't Tell Users Everything It Really Knows About Them." *Wired*, February 24, 2016. Accessed November 28, 2017. <https://www.wired.com/2016/02/facebook-reactions-totally-redesigned-like-button/>.
- Subramanian, Samanth. "Inside the Macedonian Fake News Complex." *Wired*. Last modified February 15, 2017. Accessed May 2, 2018. <https://www.wired.com/2017/02/veles-macedonia-fake-news/>.
- Sunstein, Cass R., and Adrian Vermeule. "Symposium on Conspiracy Theories Conspiracy Theories: Causes and Cures." *The Journal of Political Philosophy* 17, no. 2 (2009): 202-27. Accessed October 31, 2017. doi:10.1111/j.1467-9760.2008.00325.x.
- . "Symposium on Conspiracy Theories Conspiracy Theories: Causes and Cures." *The Journal of Political Philosophy* 17, no. 2 (2009): 202-27. Accessed October 31, 2017. doi:10.1111/j.1467-9760.2008.00325.x.
- Sydell, Laura. "We Tracked Down A Fake-News Creator In The Suburbs. Here's What We Learned." *NPR All Tech Considered*. Last modified November 23, 2016. Accessed May 2, 2018. <https://www.npr.org/sections/alltechconsidered/2016/11/23/503146770/npr-finds-the-head-of-a-covert-fake-news-operation-in-the-suburbs>.
- Taber, Charles S., and Milton Lodge. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50, no. 3 (2006): 755-69. Accessed October 14, 2018. <https://www.jstor.org/stable/3694247>.

- Tandoc, Edson C., Jr, Zheng Wei Lim, and Richard Ling. "Defining "Fake News" A typology of scholarly definition." *Digital Journalism* 6, no. 2 (2018): 137-53. Accessed November 2, 2018. <https://doi.org/10.1080/21670811.2017.1360143>.
- Tavakol, Mohsen, and Reg Dennick. "Making sense of Cronbach's alpha." *International Journal of Medical Education* 2 (June 27, 2011): 53-55. Accessed November 2, 2018. [TavakolM,DennickR.MakingsenseofCronbach'salpha.IntJMedEduc.2011;2:53-55.Published2011Jun27.doi:10.5116/ijme.4dfb.8dfd](https://doi.org/10.5116/ijme.4dfb.8dfd).
- Taylor, Jessica. "Trump Dissolves Controversial Election Commission." NPR. Last modified January 3, 2018. Accessed December 12, 2018. <https://www.npr.org/2018/01/03/575524512/trump-dissolves-controversial-election-commission>.
- Tesler, Michael. "Elite Domination of Public Doubts About Climate Change (Not Evolution)." Abstract. *Political Communication* 35, no. 2 (2018): 306-26. Accessed May 3, 2018. <https://doi.org/10.1080/10584609.2017.1380092>.
- Tewksbury, David, and Scott Althaus. "Patterns of Internet and Traditional News Media Use in a Networked Community." Abstract. *Political Communication* 17 (March 2000): 21-45. Accessed May 3, 2018. DOI:10.1080/105846000198495.
- Thorson, Emily. "Belief Echoes: The Persistent Effects of Corrected Misinformation." *Political Communication* 33, no. 3 (2016): 460-80. Accessed October 14, 2018. <https://doi.org/10.1080/10584609.2015.1102187>.
- Tien, Shannon. "How the Facebook Algorithm Works and How to Make it Work for You." *Hootsuite Blog*. Entry posted April 25, 2018. Accessed November 2, 2018. <https://blog.hootsuite.com/facebook-algorithm/>.
- Tucker, Catherine E. "Social Networks, Personalized Advertising, and Privacy Controls." Abstract. *Journal of Marketing Research* 51 (October 2014): 546-62. Accessed May 3, 2018. <https://doi.org/10.1509/jmr.10.0355>.
- Twitter. "How can we help?" Twitter Business. Last modified 2018. Accessed May 3, 2018. <https://business.twitter.com/en/help/campaign-setup.html>.
- Uberti, David. "The Real History of Fake News." *Columbia Journalism Review*, December 15, 2016. Accessed October 31, 2017. https://www.cjr.org/special_report/fake_news_history.php.
- Van Bavel, JJ, and A. Pereira. "The Partisan Brain: An Identity-Based Model of Political Belief." *Trends in Cognitive Sciences* 22, no. 3 (March 2018): 213-24. Accessed November 2, 2018. <https://doi.org/10.1016/j.tics.2018.01.004>.
- Vicario, Micheala Del, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, Stanley H. Stanley, and Walter Quattrociocchi. "The Spreading of

- Misinformation Online." *PNAS* 113, no. 3 (January 4, 2016): 554-59. Accessed October 31, 2017. doi:10.1073/pnas.1517441113.
- Vorst, Eric C. "Too Legit to Quit (Tweeting): Leadership, Followership, and Candidate Legitimacy." Paper presented at American Political Science Association, San Francisco, CA, September 1, 2017. Accessed October 7, 2017. https://convention2.allacademic.com/one/apsa/apsa17/index.php?program_focus=view_paper&selected_paper_id=1249850&cmd=online_program_direct_link&sub_action=online_program.
- Brittany Vojak, "Fake News: The Commoditization of Internet Speech," *California Western International Law Journal* 48, no. 1 (Fall 2017): 123-158
- Vosoughi, Soroush, Deb Roy, and Sinan Aral. "The spread of true and false news online." *Science* 395, no. 6380 (March 9, 2018): 1146-51. Accessed November 2, 2018. <https://doi.org/10.1126/science.aap9559>.
- Watson, Paul Joseph. "'SPIRIT COOKING': CLINTON CAMPAIGN CHAIRMAN PRACTICES BIZARRE OCCULT RITUAL Menstrual blood, semen and breast milk: Most bizarre Wikileaks revelation yet." InfoWars. Last modified November 4, 2016. Accessed November 17, 2018. <https://www.infowars.com/spirit-cooking-clinton-campaign-chairman-invited-to-bizarre-satanic-performance/>.
- Weisburd, Andrew, Clint Watts, and JM Burger. "Trolling For Trump: How Russia Is Trying To Destroy Our Democracy." *War on the Rocks*. Last modified November 6, 2016. Accessed October 31, 2017. <https://warontherocks.com/2016/11/trolling-for-trump-how-russia-is-trying-to-destroy-our-democracy/>.
- White, David Manning. "The 'Gate Keeper': A Case Study in the Selection of News." *Journalism & Mass Communication Quarterly* 27, no. 4 (September 1950): 383-90. Accessed May 2, 2018. <https://doi.org/10.1177/107769905002700403>.
- Whiting, Geoff. "If I Delete My Facebook Account, Will All My Past Comments Be Erased?" Classroom. Last modified 2018. Accessed November 2, 2018. <https://classroom.synonym.com/delete-facebook-account-past-comments-erased-7872.html>.
- Wilder, Juan. "DONALD TRUMP ANNOUNCES HIS PLAN TO ELIMINATE 490,000 TEACHER JOBS – THE MAN LOVES TO FIRE PEOPLE." *Greenville Gazette*. Last modified October 11, 2016. Accessed November 16, 2018. <http://www.greenvillegazette.com/p/143404/>.
- Wojcieszak, Magdalena E., and Diana C. Mutz. "Online Groups and Political Discourse: Do Online Discussion Spaces Facilitate Exposure to Political Disagreement?" *Journal of Communication* 59 (2009): 40-56. Accessed May 3, 2018. doi:10.1111/j.1460-2466.2008.01403.x.

- WorldTruth.TV. "Hillary Clinton in 2013: 'I Would Like to See People Like Donald Trump Run for Office'; They're Honest & Can't be Bought." WorldTruth.TV. Accessed November 17, 2018. <https://worldtruth.tv/hillary-clinton-in-2013-i-would-like-to-see-people-like-donald-trump-run-for-office-theyre-honest-and-cant-be-bought/>.
- Yglesias, Matthew. "Donald Trump's Phony War with the Press, Explained." Vox. Last modified January 1, 2018. Accessed November 2, 2018. <https://www.vox.com/policy-and-politics/2018/1/9/16854368/trump-fake-news>.
- Zamith, Rodrigo, and Seth C. Lewis. "Content Analysis and the Algorithmic Coder What Computational Social Science Means for Traditional Modes of Media Analysis." *The ANNALS of the American Academy of Political and Social Science* 659, no. 1 (April 1, 2015): 307-18. Accessed November 2, 2018. <https://doi.org/10.1177/0002716215570576>.
- Zannettou, Savvas, Michael Sirivianos, Jeremy Blackburn, and Nicolas Kourtellis. "The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans." *CoRR*, 2018. Accessed November 2, 2018. <https://arxiv.org/abs/1804.03461>.