

2021

Neutral Isn't Neutral: An Analysis of Misinformation and Sentiment in the Wake of the Capitol Riots

Daniel Robert Walsh
West Virginia University, dowalsh@mix.wvu.edu

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>



Part of the [Journalism Studies Commons](#)

Recommended Citation

Walsh, Daniel Robert, "Neutral Isn't Neutral: An Analysis of Misinformation and Sentiment in the Wake of the Capitol Riots" (2021). *Graduate Theses, Dissertations, and Problem Reports*. 8055.
<https://researchrepository.wvu.edu/etd/8055>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Neutral Isn't Neutral: An Analysis of Misinformation and Sentiment in the Wake of the Capitol Riots

Daniel Walsh

Thesis submitted

to the Reed College of Media

at West Virginia University

in partial fulfillment of the requirements for the degree of

Master's of Science in

Journalism

Bob Britten, Ph.D., Chair

Ashton Marra, M.S.

John Temple, M.F.A.

Rachael Woldoff, Ph.D.

Reed College of Media

Morgantown, West Virginia

2021

Keywords: Twitter, Tweets, Misinformation, Sentiment, Dissemination, News Gathering, Trump

Copyright 2021 Daniel Walsh

ABSTRACT

Misinformation and Sentiments: An Analysis of Twitter in the Week After the Capitol Riots

Daniel Walsh

January 6th, 2021 was a significant moment in the history of the United States of America. Protestors stormed the Capitol building over the results of the 2020 presidential election in which Joseph R. Biden defeated incumbent president Donald J. Trump. The Capitol riots were partially incited by the presence of misinformation on social media and was an example of the power misinformation has. This study presented two questions. Question one pertains to the sentiment analysis of verified Twitter users and their sentiment towards Trump. Question two pertains to analyzing tweets from verified accounts for misinformation between the dates of January 6th, 2021 and January 13th, 2021. To answer these questions, a machine learning sentiment analysis was conducted on 13 randomly selected Twitter accounts with noted liberal and conservative political leanings to assess their sentiment towards Trump. The accounts were analyzed and then categorized as being either anti-Trump or Trump-neutral. Once the accounts were appropriately categorized a collection of their tweets mentioning Trump were documented to create a consecutive day sample to examine their reporting and analyze how misinformation differed between the two. The results of this study show that one, sentiment analysis is a useful tool for examining and categorizing tweets and their overall accounts based on their sentiments and two, that there was a notable difference in the spread of misinformation between the two categories.

Table of Contents

| | |
|--|-----------|
| <i>List of Tables</i> | <i>iv</i> |
| <i>Introduction</i> | <i>1</i> |
| <i>Literature Review</i> | <i>2</i> |
| <i>Online News Gathering</i> | <i>3</i> |
| <i>Partisanship and Misinformation on Social Media</i> | <i>4</i> |
| <i>The 2021 Capitol Riots</i> | <i>7</i> |
| <i>Sentiment Analysis</i> | <i>9</i> |
| <i>Methodology</i> | <i>11</i> |
| <i>Identifying Groups Through Sentiment Analysis</i> | <i>11</i> |
| <i>Quantifying the Misinformation</i> | <i>13</i> |
| <i>Results</i> | <i>15</i> |
| <i>Sentiment Analysis Results</i> | <i>15</i> |
| <i>Misinformation Analysis of Posts</i> | <i>20</i> |
| <i>Discussion</i> | <i>25</i> |
| <i>Limitations and Future Research</i> | <i>28</i> |
| <i>Conclusion</i> | <i>29</i> |
| <i>References</i> | <i>31</i> |
| <i>Appendix A: Tables</i> | <i>37</i> |

List of Tables

Table 1..... 16

Table 2..... 17

Table 3..... 19

Table 4.....22

Table 5..... 23

Table 6..... 25

Introduction

Social media has continued to grow and gain significance as a source of information for society. Due to its immense growth and its explosion as a unique source for news gathering, it has become necessary to study its place in journalism. This use of social media for receiving news and information has given rise to both benefits and issues. One of the most serious issues social media poses to journalism is its ability to facilitate and spread articles or news with misinformation. Misinformation, especially as easily accessed as it is on social media, can present a serious challenge to the quality of democratic self-governance (Kuklinski et al., 2000). As a pillar of democracy and a tool for society to gain information on the world around them, upholding journalism is crucial.

The social media microblogging site Twitter has revolutionized journalism and is becoming the par excellence social media to catch up on recent news and events (Zubiaga et al., 2013). From newspapers, to radios, to television and now to social media, Twitter has become extremely beneficial and essential for journalism and according to Zubiaga et al. it is changing the way news is gathered. However, it is not without its faults. The brevity of tweets and the increasing usage of mobile devices are facilitating the ability to quickly share information from anywhere at any time, (Kwak et al., 2010), regardless of the truth of the news. The fact is that information fabrication is not new, misinformation has been featured in human communication throughout history, (Posetti & Matthews, 2018), and with new technology it has never been easier to spread it. Because of this, it is pertinent to continue the research of misinformation into the realm of social media news.

Social media has not only changed how people are able to communicate, but also people's information behavior entirely. Given the unique way in which information and news is shared, it can be spread and circulated rapidly and widely. Popular examples of this phenomenon on Twitter include the Hudson River plane crash photo in 2009, or the witness photo in the 2012 Empire State shooting, among many others (Zubiaga et al., 2013). This way of information sharing is unlike any method the world has seen before and is still relatively new. Used responsibly, social media can positively impact society by giving its people rapid, informative, and impactful news at the tips of their fingers. After all, from a

democratic viewpoint, the most essential function of the news media is to aid its citizens in being informed on the world, policies, and happenings around them (What is the Purpose of Journalism, 2017). Ideally, if the news fulfills this basic function, it will create a better society of free and knowledgeable self-governing people.

However, like anything, there is a major downside to social media being a massive player in the news media. Because the Internet and social media are like the wild west it is almost impossible to control what is posted and spread. Credible and truthful stories exist in the same place as purposely false stories and can even be found right next to each other which creates confusion amongst the users. In addition to this, ordinary people who have no education in journalism or its ethics can become breaking news reporters, which could result in unintentional (or intentional) bias on the subject they are covering.

Not only is the freedom of Twitter a double-edged sword in terms of the information received, but users can unknowingly create echo chambers for themselves. According to Sunstein (2007), much of the online public sphere is captured in the influential conceptualization of the internet and that it is mostly a sphere for social networking and individual profiling, which creates biased echo chambers of like-minded individuals instead of focal points for political news. Echo chambers can be instrumental in disseminating misinformation, especially echo chambers aligned to one political side.

Literature Review

Scholars have defined social media as a group of internet-based applications that allow for users to create and exchange content and information (Kaplan & Harenlein, 2010). Though social media is a relatively recent modality, it represents an ongoing internet-based evolution of news outlets, following after earlier channels of print news, radio, and television journalism. One of social media's fundamental features is that it allows users to discover and share a plethora of information to anyone using the application (Anttiroiko & Savolainen, 2011). Social media has transformed the landscape of news and information sharing, with the Pew Research Center reporting that 68 percent of American adults actively seek out news from social media (Matsa & Shearer, 2018). The introduction of social media to society has

been correlated with enhanced participation by Twitter users in self-created echo chambers which uphold and expand their existing beliefs (Guess et al., 2019).

Online News Gathering

Americans' use of social media microblogging sites, such as Twitter, to gather news has been especially impactful because it has altered people's newsgathering tendencies. According to a report from the 2011 State of the News Media (Pew, 2011), all forms of news media suffered audience decline with online news being the only exception. This trend in online news consumption has only continued, as pundits now claim that on social media sites, such as Twitter, users do not seek news anymore, and that news instead finds them (Gangadharbatla et al., 2014). However, according to Gangadharbatla et al., even though people now prefer newsgathering on social media, they still perceive traditional media sources as more credible. Furthermore, sharing the information is more likely to happen among Twitter users if the message was in the form of a newspaper article since traditional media are perceived as more credible (Utz et al., 2013). Additionally, just as television media turned readers of media content into watchers of media content, Gangadharbatla et al. suggests social media has created a nation of media content creators which is unlike anything the world has ever experienced.

This shift into new territory has dramatically changed the techniques of newsgathering in the American people. Social media newsgathering has also become a more used source of news gathering than print media. A 2018 study by the Pew Research Center (Shearer, 2018) found that 20 percent of people get their news directly from social media while only 16 percent get it from print newspapers. Twitter is not only impactful for people as news gatherers but also for journalists. Twitter has become a gold mine for journalists to learn about breaking news and to gather information to broaden their coverage of news (Zubiaga, 2019). Furthermore, Zubiaga states that journalists now use Twitter to freely post stories in order to reach a wider audience, to analyze the preference and consumption of news gatherers, and finally, to keep up with breaking news and to retrieve additional context for news reports. Twitter for

news gathering can also be exhausting for people to use. Because Twitter users have a constant feed to news, the platform presents a unique challenge in that the stream of updates flows much faster than a person can process and makes it impossible for individuals to keep track of all the information.

Partisanship and Misinformation on Social Media

Twitter is a microblogging social media platform where information, or “tweets,” can be shared and spread regardless of whether or not its content is truthful. Due to Twitter’s unique structure compared to past news outlets, such as print or television, one can relay posts on the site with no significant third-party filtering, fact-checking, or editorial judgment. Twitter can be considered as a moving media system where users get information from both established media and each other, but also where not all information is an act of journalism (Hermida, 2010). Additionally, an individual user with no track record or reputation can reach as many readers as national news outlets (Allcott & Gentzkow, 2017). This is because Twitter provides users the ability to quickly spread the stories of community’s interest, which the site reflects as trending topics (Zubiaga et al., 2013). These trending topics often originate from breaking news, and then users can contribute, comment and spread (Zubiaga et al., 2011).

According to Weeks and Holbert (2013), dissemination is sharing news content with a defined set of others, either specific groups within one’s social network or the social media public. Dissemination is a proactive process, in which users take the initiative in distributing information or knowledge, giving it great potential for information collaboration among and between professionals and the public, though it is a dangerous force for those who want to use it to spread misinformation. The lack of guidelines and sheer amount of content that Twitter users encounter daily can be overwhelming, which is one reason that misinformation can be so successful on these platforms.

The mixture of increased partisanship, paired with the expanded use of social media by United States citizens, have created fears of widespread “echo chambers” and environments where people only encounter beliefs that coincide with their own (Sunstein, 2001; Pariser, 2011). Past research has shown

that people tend to prefer information that reaffirms their beliefs, especially in politics (Stroud, 2008; Hart et al., 2009; Iyengar & Hahn, 2009; Iyengar et al., 2008). Other research proposes that user trends towards reaffirming information could be intensified by sharing and consuming content online (Bakshy et al., 2015). This is politically notable because according to Jost et al (2017) conservative users have significantly more homogenous online networks than liberals. Because of hyperpartisanship and Twitter users' desire for reaffirming information, social media may be a breeding ground for misinformation.

Misinformation is information that is shown to be inaccurate. Due to social media, misinformation has rapidly taken on a life of its own (Karlova & Fisher, 2013). Some misinformation is unintentional in trying to deceive its viewers and is spread by regular users simply because of their unverified trust in the information by its sources. Instead of trying to deceive in spreading the misinformation, they believe they are spreading true information that is helpful to viewers (Wu et al., 2019). Under the umbrella of misinformation is disinformation. Disinformation differs from misinformation because it is information that is false, and it is deliberately created to harm a person, social group, organization, or country (United Nations Educational and Scientific Cultural Organization, 2020). Wu et al. suggests that there are several known and generally accepted terms for similar and related to misinformation. Those would be rumors, information not proven to be true and that has no source, spam, unverified information sent to many receipts, and the most common idea of misinformation held by the public, fake news, which is the intentional spread of misinformation in the format of a news program. Fake news is most similar and connected to disinformation because it is purposely deceitful, inaccurate or false.

There have been many examples of misinformation on social media especially since the 2016 presidential election. For instance, during the hearings of appointed Supreme Court Justice Brett Kavanaugh, misinformative posts on social media claimed that Christine Blasey Ford, a woman accusing Kavanaugh of sexual assault, had been photographed with George Soros, a frequent target of right-wing conspiracy theories. The alleged photograph of Blasey Ford and Soros was shared tens of thousands of times but was later proven to be Soros posing with a Ukrainian human rights activist (Roose, 2018). In

addition, Roose, detailed another instance of misinformation on social media is in the form of deceptive or misleading claims about political candidates. Roose states that a doctored photo of Stacey Abrams, the Democratic candidate for governor of Georgia, was disseminated thousands of times on social media. The photo of Abrams falsely showed her holding a sign reading “communist Stacy Adams” when the original sign read “Stacy Abrams Governor.” Though some of these examples seem harmless, according to Wu et al. (2019) they aim to undermine democracy and can cause various kinds of destructive effects through social media, especially when timely intervention is absent. One such example of the dangers of circulating misinformation on social media is Pizzagate. Pizzagate was a baseless conspiracy of a pizzeria allegedly being the location of a huge child sex trafficking ring. The misinformation led to a situation that motivated someone to fire a rifle inside the restaurant. These are all examples of misinformation on social media and illustrates the effects it can have on those viewing it. Some users could have been intentionally spreading these stories while others could have no desire or idea that they were spreading misinformation.

In relation to political partisanship and the consumption of misinformation, past research has shown some correlations. Grinberg et al. (2019), found that conservative users are more likely to engage with sources containing misinformation and Bovet and Maske (2019), found that misinformative users on Twitter and other social media sites are more likely to influence Trump supporters. Furthermore, a recent study conducted by Huang and Carley (2020), which looked at misinformation on Twitter during the Coronavirus outbreak, found that most of the people who tweet and retweet misinformation lean towards conservative users. It has also been established by Jost et al. (2018) that political conservatives are more likely than liberals to: (1) prioritize values of conformity and tradition; (2) possess a strong desire to share reality with like-minded others; (3) perceive within-group consensus when making political and non-political judgments; (4) be influenced by “watchful eyes” and sources who are perceived as similar to them; and (5) maintain homogenous social networks and favor an “echo chamber” environment that is conducive to the spread of misinformation. Continuing with conservatives' relationship with misinformation, conservatives endorse conspiracies more than liberals and are also more likely to spread

rumors and other types of misleading posts (Robbins, 2017). Due to the potential implications of these findings and the increasingly partisan nature of society it is important to continue to research how misinformation works on Twitter between users. That is what this research intends to accomplish.

The 2021 Capitol Riots

The Capitol riots on January 6th, 2021 was a signature moment in the history of the United States and a testament to the power of misinformation. A mob of people, predominantly Trump supporters stormed the Capitol building and the halls of Congress which resulted in the deaths of five people and numerous arrests. To make matters graver it happened as Congress was certifying the electoral votes for president elect Joe Biden. The Capitol riots are monumental for several reasons, but in the lens of this research it is critical because of the role misinformation played in inciting it.

Before the riots took place misinformation was swarming online and on social media sites which engulfed people and led them to storm the Capitol building. According to Munn (2021), the media worked to forge connections between disparate camps, to incite participants toward violent activity, and to legitimize this attack as moral or even spiritual. The media also took abstract ideologies and information and transformed them into the physical chaos that was seen during the riots. The media in question, as proposed by Munn, was highly misinformative. This historic event is an extreme example of what misinformation can do. Reuters (2021) reported that, “online misinformation that led to violent unrest at the Capitol last week has gone beyond false claims and has reached the point of ‘radicalization.’” The riot illustrated the power that misinformation on social media has and the lengths that it can push people.

In the months after the 2020 election, in which Joe Biden was elected the 46th president of the United States, and up to January 6th, misinformation, false claims and conspiracy theories spread through social media like wildfire. It was found that violent rhetoric and misinformation on social media ramped up in the weeks right before (Reuters, 2021). The misinformation about Joe Biden stealing the election from Donald Trump and Trump winning it flourished on social media despite mountains of evidence

against it. This is alarming because misinformation had affected people so intensely that they either refused to believe the evidence in front of them or didn't care. The misinformation was perpetuated by Trump himself, Republican officials, partisan media outlets and people on social media.

Due to the riot's social media sites like Twitter and Facebook suspended and then outright banned Trump's personal accounts on their applications. The social media sites also began to ban other perpetrators of the misinformation and posts spreading it. This came after Twitter initiated a feature which labeled Trump's misleading tweets on the website.

Because of the riots, and the misinformation which played a role in causing it, this research wanted to examine misinformation during, and a week immediately following it. It will accomplish this by analyzing tweets from partisan accounts mentioning Trump. To research the dissemination of misinformation from posted tweets requires appropriately coding what constitutes as a misinformative Twitter account. Based on the prior literature that has been presented, users who spread misinformation are defined as users who published two or more tweets that can be previously coded by Allcott and Gentzkow (2017) as false, misleading or misinformative. This research by Allcott and Gentzkow will be applied when evaluating tweets from the accounts categorized as anti-Trump or Trump-neutral to measure the spread of misinformation by each on Twitter.

The users that spread misinformation to be addressed in this research will be listed, and have little regard for adhering to journalistic norms, as ethical journalists should never mislead and always tell the truth. Though at times these users can publish accurate information, they frequently publish false claims, distortions, misinformation and copy news from other users. However, one cannot look at a user and definitively place them in a category as anti-Trump or Trump-neutral just based off them posting misinformation. Therefore, a sentiment analysis of the selected Twitter users will be implemented to place users in a binary category or anti-Trump or Trump-neutral.

Sentiment Analysis

Twitter is a platform where millions of users share their opinions, and because of this trying to collect a complete comprehension of these opinions can be difficult and time consuming. Sentiment analysis is an approach that uses Natural Language Processing (NLP) to extract, convert and interpret opinions from a text and classify them into positive, negative, or neutral sentiment (Drus & Khalid, 2020). Sentiment analysis in the digital age has become an increasingly valuable field of study that analyzes people's (or user's) opinions, sentiments, evaluations, attitudes, and emotions from written language (Liu, 2012).

Sentiment analysis is divided into three different levels which are sentence, document, and feature. Sentiment analysis on Twitter will be effective using a sentence level analysis due to Twitter's text length constrictions, as users only get 280 characters per tweet. Due to the abilities of sentiment analysis, a huge volume of opinionated data can be processed, analyzed, and then recorded digitally for further analysis and research.

Sentiment analysis is an active area of study in the field of natural language processing that analyzes people's opinions, sentiments, evaluations, attitudes, and emotions via the computational treatment of subjectivity in text (Hutto & Gilbert, 2014). Previous research shows that a high number of sentiment analysis uses rely on the sentiment lexicon-based application. Sentiment lexicons is a list of polar lexical features (e.g., words), that are then labeled according to their orientation. According to Kaity and Balakrishnan (2020), lexicon-based sentiment analysis relies on sentiment lexicons that are produced manually or semi automatically. Because sentiment lexicons need to be developed manually it can be very time consuming. The other approach to sentiment analysis and classification is the machine learning method. Machine learning-based classifiers depend on training datasets and employ an automated means to identify sentiment. This research paper will specifically use the machine learning approach of sentiment analysis, which utilizes algorithms to extract and detect sentiment in language (in this case tweets) to configure into data.

Uses of this method incorporate machine learning approaches to review the sentiment-relevant features of text. However, the machine learning approach is not without its flaws. Occasionally, machine learning can be difficult to interpret and can be expensive (Hutto & Gilbert, 2014).

In recent years there has been a lot of work done in the field of sentiment analysis, and specifically sentiment analysis on Twitter. Early on sentiment analysis was intended to be used as a system for binary classification of text data, meaning that it assigns opinions (tweets) to binary categories such as negative or positive (Gautam & Yadav, 2014). Turney (2002), was able to review sentiment by the average semantic orientation of a phrase that contains an adjective and adverb, making it possible to calculate whether the given phrase is positive or negative.

Pak and Paroubek (2010), worked on a sentiment analysis model to categorize tweets as either positive or negative or neutral. The duo constructed a Twitter corpus (collection of texts) by collecting tweets using Twitter API and then proceeded to automatically annotate those tweets using emoticons. Using the collection of texts, they developed a sentiment classifier using the machine learning method. However, this was ultimately an inefficient method of sentiment analysis because it relied on tweets containing emoticons. Kamps et al. (2004), used the lexical database WordNet to determine the emotional content of text. They then went on to develop a distance metric on WordNet and determine semantic polarity of adjectives which could then be used to interpret tweets.

Research Questions

To expand and continue the prior research on misinformation on Twitter regarding elections, two research questions arise:

RQ1: How does Twitter sentiment regarding Donald Trump differ between anti-Trump and Trump-neutral accounts?

RQ2: How did the spread of misinformation on Twitter differ between anti-Trump and Trump-neutral Twitter accounts in the week after January 6th, 2021?

Methodology

This research paper will work as a two-part exploratory study into the spread of misinformation from accounts on Twitter in the wake of the Capitol riots on January 6th to January 13th. The study will focus only on tweets mentioning the keyword “Trump.”

Identifying Groups Through Sentiment Analysis

The first component of this study will be a sentiment analysis of a pool of both conservative-leaning and liberal-leaning verified users on Twitter to categorize them into either pro-Trump users, Trump-neutral users or anti-Trump users. Due to the nature of Twitter, a sentiment analysis of 31 random tweets from users will be most effective in collecting data on people’s opinions and sentiments. For this, a simple random sampling method will be conducted. Using sentiment analysis tools to analyze these accounts' opinions will allow this paper to understand how the users ‘feel’ towards Donald Trump. Additionally, sentiment analysis will be useful in this research because it is an automated process of sorting text data.

This research paper will sort tweets (text data) into sentiments that are either positive, negative or neutral with regard to user’s sentiment towards Donald Trump. The more positive sentiment a Twitter user’s tweets have towards Donald Trump will result in the user being categorized as Pro-Trump. Users that share more negative sentiment towards Donald Trump will be categorized as anti-Trump. Users that are predominantly neutral on the subject will be categorized as Trump-neutral.

The Twitter users selected to be studied were chosen by way of simple random sampling. The only constraint was that the user needed to be verified on Twitter and there needed to be at least five accounts from each side of the political spectrum. Conservative-leaning Twitter users that will be looked at will be @OANN, @RedState, @BreitbartNews, @theblaze and @townhallcom. On the other side, liberal-leaning Twitter users will be @AlterNet, @newrepublic, @dailykos, @crooksandliars, @RawStory, @TheYoungTurks, democracynow and @Salon.

The previous users mentioned were randomly selected from a bias study conducted by Allsides.com. Allsides is a website that filters news sources by bias (left, right, center). Allsides uses a patented media bias rating system to classify news sources. The Allsides rating system is developed by crowdsourcing, surveys, internal research, and use of third party sources. A Pew Research Center Study by Gottfried et al. (2020) on bias will additionally be used as an identifier for news users along with the information taken from Allsides.com. This is to assess a complete and fair list of users.

To gauge the sentiment of Twitter users towards Donald Trump a month of random tweets will be constructed. The keyword “Trump” will be searched on Twitter along with one of the proposed accounts and at least 31 random tweets with the “Trump” keyword will be recorded to be analyzed by sentiment analysis. Once the tweets are examined by the sentiment analysis to show their detected polarity as either positive, neutral or negative, they will be appropriately categorized as either pro-Trump, Trump-neutral or anti-Trump.

A machine learning approach to sentiment analysis will be used in this research to gauge sentiment and categorize Twitter users into either a Pro-Trump, Trump-neutral, or anti-Trump group. It is an efficient tool for sentiment analysis because there are hundreds of millions of tweets a day, and the sheer volume of sentiment in the text data alone means that having an automated process for identifying sentiment will work in this research's favor.

Monkeylearn.com will be used to accurately find sentiment in users’ tweets based on their method of sorting text into either positive, neutral, or negative polarity, which is a similar approach used by Drus and Khalid (2020). MonkeyLearn is a machine learning sentiment analysis program that allows users to build their own sentiment analysis model and directly connect them to Twitter data by use of one-click integration. Furthermore, MonkeyLearn is an effective sentiment analysis tool because it provides a probability rating of the tweet’s sentiment right in its machine learning algorithm.

This research will obtain its collection of tweets to be analyzed by the sentiment analysis program through use of Allmytweets.net, a Twitter archive website. The tweets will be copied from Allmytweets.net and input verbatim to be analyzed by the sentiment analysis tool. Once the sentiment

analysis is produced it will be recorded along with the tweet and its percent of confidence for further study.

Polarity detection in sentiment analysis focuses on using binary classification outputs such as positive versus negative, thumbs up versus thumbs down or like versus dislike to categorize texts, (Cambria et al., 2018). Polarity detection will be used as the method of classifying tweets into either Trump-neutral or anti-Trump.

Finally, accounts will be labeled as either pro-Trump, Trump-neutral or anti-Trump based off its overall sentiment towards Trump. This is to say that at the end of the sentiment analysis study once every tweet has a labeled sentiment whatever sentiment dominates the others will be what the overall account will be categorized as.

Quantifying the Misinformation

Once the Twitter users are correctly classified as either pro-Trump, Trump-neutral or anti-Trump, the second component of the research will begin. Following in the footsteps of prior research by Riffe et al. (1993), the chosen users will be studied by way of a consecutive day sampling method. This method will be used because it will allow for the convenience in only needing to examine a week of articles to collect the data. It is also useful because it provides the ability to control the variables that go in the sample.

What will be analyzed over the seven-day period is tweets from the accounts which mentioned “Trump.” The tweets will be tracked from January 6th to January 13th, which was during and the week following the riots at the Capitol building. Only tweets from each of the chosen accounts mentioning “Trump” during the sample time frame will be analyzed. Likewise, the tweets to be selected will be posts directly from the chosen accounts. Retweets will not be reviewed in this study because it is looking at what the specific account is posting. This study wants to look directly at the posts from selected accounts and retweets will not count simply because they are not the selected accounts personally created

information. Each article linked by tweet that fits the prior criteria from an account will then be reviewed to determine whether it is truthful or misinformative.

The tweets selected based on the criteria will be fact checked by trusted sources such as PolitiFact, FactCheck.org and the Washington Post Fact Checker. If a selected tweet is not present or has not been fact checked by either PolitiFact, FactCheck.org or the Washington Post Fact Checker, this paper will use the credible fact checking rubric from PolitiFact. According to an article by Holan (2018), PolitiFact uses a detailed rubric for coding a statement to be true or false. The rubric will be used by this study and applied to the articles being shared from the selected accounts. The first mark of the rubric is to check if the statement is literally true and not an opinion. In short, can it be proved or disproved? The second mark is if the statement made is significant. The third mark is if the statement is likely to be spread, which is especially important in this work because all accounts being reviewed in this study have high followers and therefore high probability of spreading their posted tweets. Finally, the last mark for their decision in assigning a true or false label is if a typical person will read the tweet and wonder if it is true. Tweets that are obviously false will be coded accordingly.

Additionally, the Washington Post Fact Checker uses a reasonable person standard to analyze claims and classifies individual statements using up to four Pinocchios: one for “some omissions and exaggerations,” two for “significant omissions or exaggerations” leading to a false impression, three for “significant factual error,” and four for “whoppers” (Kessler, 2017a). This paper will code a tweet as being ‘misinformative’ if it gets even one ‘Pinocchio’ from the Washington Post Fact Checker.

In the case of Politifact, a tweet that is not ‘true’ according to the Truth-O-Meter, will be coded as misinformative. Additionally, Snopes and Factcheck.org will be used to evaluate whether a tweet will be coded as either ‘misinformative’ or not. Information in the selected tweets that these other fact checking sites deem to be untruthful will result in is being coded as misinformative. Only tweets that are 100% true from these Twitter accounts will be coded as such. This is because the paper is looking specifically at misinformation in the tweet, not the level of misinformation. Tweet articles that are even slightly misinformative will be coded as such. That means that as little as one false claim will result in the entire

article being labeled as misinformation. Once all the tweets from the accounts have been put into their proper categories, this research will quantitate the findings and compare the data to determine where the two variables differ from each other.

Results

This research set out to answer two research questions revolving around misinformation on Twitter. The questions stemmed from the fact that most people gather news from social media (Matsa & Shearer, 2018). Because of this style of news gathering on social media sites like Twitter, users can unknowingly place themselves self-created echo chambers which can advance and support their existing beliefs (Guess et al., 2019). Beliefs that may or may not be rooted in misinformation and provided by biased or unethical accounts.

This research began with wanting to look at how misinformation was spread during the month leading up to the 2020 president election. However, instead of examining misinformation regarding the 2020 presidential election, this research conducted a case study which examined posts from Twitter accounts in the week during and following the Capitol riots. This examination was conducted to analyze their role in the spread of misinformation.

Sentiment Analysis Results

RQ1 asked “how does Twitter sentiment regarding Donald Trump differ between anti-Trump and Trump-neutral accounts.” To answer this, 13 Twitter accounts which were identified as either liberal or conservative by Allsides.com were randomly selected to be studied because of them being verified on Twitter. According to help.twitter.com, to be given a verified status as a Twitter account, the account must be authentic, notable and active and it is of public interest. These accounts were then examined by a sentiment analysis machine learning algorithm to determine the overall sentiment of the accounts towards former president Donald Trump. The sentiment analysis was used to sort tweets from the selected

accounts in either positive, neutral, or negative categories. Once all the tweets were sorted, they would then be used to determine the dominant overall sentiment of the accounts which would be either positive, neutral or negative.

To determine the sentiment of a specific selected account, each account had 31 random tweets pulled from their accounts to create an artificial month for a sentiment analysis. This method led to the collection and analysis of 403 tweets spanning the 13 accounts. The keyword “Trump” was used to identify tweets specifically revolving around former president Donald Trump because the reporting of him is the basis of the sentiment analysis. Furthermore, all the tweets that were collected are tweets that were published directly from the accounts.

After collecting the simulated month of random tweets, each tweet was analyzed via a sentiment analysis machine learning program from MonkeyLearn.com. Tweets were copied and entered verbatim into the sentiment analysis program on MonkeyLearn.com. After putting the tweet into the sentiment analysis application, a sentiment of either negative, positive, or neutral was attributed to the tweet. Each sentiment was accompanied by a percentage to gauge the program's confidence in its sentiment categorization. Table 1 provides three examples of the process of tweets being labeled as positive, neutral, or negative below.

Table 1

Sentiment Analysis Examples

| Twitter User | Tweets | Overall Assessment |
|--------------|---|--------------------|
| @AlterNet | Complete incompetence': Biden officials fume after Trump leaves them 'nonexistent' vaccine plan | Negative 95% |

| | | |
|--------------|--|--------------|
| @townhallcom | Even if Trump disappears, his impact on the GOP is lasting. | Positive 58% |
| @dailykos | Voting Rights Roundup: Biden rescinds Trump order that tried to weaponize the census against Latinos | Neutral 72% |

The results determined by the sentiment analysis were then recorded and quantified (see Table 2). The sentiment analysis showed that between the 403 tweets and 13 Twitter accounts only 30 (7.44% of the total sample) had a positive sentiment towards Trump in their tweets. Comparatively, the 13 Twitter accounts had a combined 154 (38.21%) tweets with negative sentiment towards Trump. However, 219 (54.34%) tweets from the 13 Twitter accounts were neutral in their sentiment towards Trump. This means that just a little more than half of the 403 random tweets collected had neither positive nor negative sentiment towards Trump across both liberal-leaning and conservative-leaning Twitter accounts.

Table 2
Sentiment From Sampled Twitter Accounts.

| Twitter Account | Positive | Neutral | Negative |
|-----------------|------------|--------------------|------------|
| @OANN | 1 (3.23%) | 28 (90.32%) | 2 (6.45%) |
| @RedState | 4 (12.90%) | 23 (74.19%) | 4 (12.90%) |
| @democracynow | 3 (9.68%) | 23 (74.19%) | 5 (16.13%) |
| @BreitbartNews | 2 (6.45%) | 23 (74.19%) | 6 (19.35%) |

| | | | |
|-----------------|-------------------|---------------------|---------------------|
| @theblaze | 0 (0.00%) | 22 (70.97%) | 9 (29.03%) |
| @townhallcom | 4 (12.90%) | 22 (70.97%) | 5 (16.13%) |
| @TheYoungTurks | 2 (6.45%) | 21 (67.74%) | 8 (25.81%) |
| @RawStory | 0 (0.00%) | 6 (13.35%) | 25 (80.65%) |
| @newrepublic | 1 (3.23%) | 10 (32.26%) | 20 (64.52%) |
| @AlterNet | 5 (16.13%) | 7 (22.58%) | 19 (61.29%) |
| @Salon | 2 (6.45%) | 11 (35.48%) | 18 (58.06%) |
| @dailykos | 2 (6.45%) | 11 (35.48%) | 18 (58.06%) |
| @crooksandliars | 4 (12.90%) | 12 (38.71%) | 15 (48.39%) |
| Total | 30 (7.44%) | 219 (54.34%) | 154 (38.21%) |

To categorize a Twitter user into a specific category of anti-Trump, Trump-positive or Trump-neutral, the account had to have a quantified value in either negative, neutral, or positive higher than all the others. For example: @dailykos was categorized as anti-Trump because its dominant sentiment was negative, as 58.06% of its recorded tweets had negative sentiment towards Trump. In contrast, @BreitbartNews was categorized as Trump-neutral because it was 74.19% neutral in its sentiment towards Trump.

Of the 13 accounts analyzed in this study by the standards which were set, not one account fell into the Trump-positive sentiment category. The accounts were either anti-Trump or Trump-neutral. This does not mean there was no positive sentiment towards Trump, as there were 30 recorded tweets with positive sentiment coming from both liberal and conservative accounts. But it does mean that no account

had a dominant number of positive tweets. Analysis showed that six accounts were anti-Trump in their sentiment, and all of them leaned politically liberal. The remaining seven accounts were Trump-neutral in their sentiment, and the classifications of accounts and results are seen in Table 3. The results of the sentiment analysis show that all the conservative-leaning Twitter accounts were Trump-neutral.

Table 3

Categorization of Twitter Accounts and Their Follower Count

| Organization | Classification | Political Leaning | Twitter Follower Count |
|-----------------|----------------|-------------------|------------------------|
| @BreitbartNews | Trump-Neutral | Conservative | 1.4M |
| @democracynow | Trump-Neutral | Liberal | 788.6K |
| @theblaze | Trump-Neutral | Conservative | 690.8K |
| @TheYoungTurks | Trump-Neutral | Liberal | 444.7K |
| @Townhallcom | Trump-Neutral | Conservative | 168.7K |
| @OANN | Trump-Neutral | Conservative | 1.5M |
| @RedState | Trump-Neutral | Conservative | 233.7K |
| @AlterNet | Anti-Trump | Liberal | 137.9K |
| @crooksandliars | Anti-Trump | Liberal | 94.5K |
| @dailykos | Anti-Trump | Liberal | 292.4K |
| @newrepublic | Anti-Trump | Liberal | 176.4K |
| @RawStory | Anti-Trump | Liberal | 222.3K |

Misinformation Analysis of Posts

RQ2 asked “How did the spread of misinformation on Twitter differ between anti-Trump and Trump-neutral Twitter accounts in the week after January 6th, 2021?” Once the sentiment analysis was finalized and each of the accounts were correctly placed in their respective sentiment categories, this study used the Twitter archive website Allmytweets.net to access and record all the tweets to be analyzed in this study. The timeline of the samples was set using the consecutive day sampling method to track seven days of tweets starting at and immediately following the Capitol riots on January 6th.

Ultimately, four Twitter accounts from each category (anti-Trump and Trump-neutral) were chosen because of changes in the methods (see page 19). The number of accounts had to be altered because one of the anti-Trump accounts (@RawStory) did not have enough tweets to produce an accurate sample for the case study, and one of the Trump-neutral accounts (@OANN) only posted videos with no articles attached. This research wanted to look solely at articles attached to posts, the video-only account could not be analyzed. The final selection of anti-Trump accounts was @dailykos, @newrepublic, @Salon and @AlterNet. The accounts selected from the Trump-neutral categories were @BreitbartNews, @RedState, @theblaze and @townhallcom.

Similarly, to the analysis of tweets in the sentiment analysis section, this section analyzed only original Twitter posts from the accounts. No retweets, likes, or replies were considered. Additionally, the keyword “Trump” was used to filter out tweets not pertaining to former president Donald Trump to make the sample and research more focused. The documentation of tweets from the eight accounts led to the collection of 416 tweets. The Trump-neutral accounts had a combined 198 tweets that met the criteria while the anti-Trump accounts had 218.

Once the study had all the tweets recorded, a comprehensive analysis of the tweets began. Each of the 416 tweets were analyzed using the process described in the methodology section. The tweets were conclusively examined, using trusted fact checking sources such as PolitiFact, Snops.com, FactCheck.org and the Washington Post Fact Checker. Furthermore, when a tweet from an account was not fact checked by either PolitiFact, FactCheck.org or the Washington Post Fact Checker, the study used the credible fact checking rubric from PolitiFact. The PolitiFact rubric, according to Holan (2018), used a detailed rubric for coding a statement to be true or false.

This research only examined whether there was misinformation present and not the levels of misinformation. Because of that, only tweets that were 100% misinformation-free were coded as true while tweets that were even slightly misinformative were coded as misinformation. Finally, once all the 416 tweets from the accounts had been put into their proper categories as either truthful or misinformative, the research quantitated the findings and compared the data to determine where the two categorized accounts differ from each other.

Beginning with the anti-Trump accounts, 218 tweets with attached articles during the consecutive day sample were recorded and analyzed (see Table 4). The account @Salon had 71 tweets over the focus period and had only two articles that contained misinformation. Of the anti-Trump accounts, @Salon had the most instances of misinformation. In contrast, the account @AlterNet had 58 tweets over the focus period, and none contained evidence of misinformation. This was the only account examined that had no instances of misinformation during the sample period throughout all of the eight examined accounts. Overall, only four of the 218 articles (1.83% of articles analyzed) contained any trace of misinformation from the anti-Trump categorized Twitter accounts.

Table 4

Misinformation Analysis in Anti-Trump Accounts

| Twitter Users | Number of posts with misinformation | Percent of posts with misinformation | Number of posts |
|---------------|-------------------------------------|--------------------------------------|-----------------|
| @newrepublic | 1 | 3.70% | 27 |
| @Salon | 2 | 2.82% | 71 |
| @dailykos | 1 | 1.54% | 65 |
| @AlterNet | 0 | 0.00% | 58 |
| Total | 4 | 1.83% | 218 |

The Trump-neutral accounts had 198 tweets (see Table 5) with attached articles during the consecutive day sample, which were fewer posted tweets than the anti-Trump accounts. Though they had fewer posts, the Trump-neutral accounts had 20 instances of posts with misinformation, which made up 10.10% of their overall posts. The account @townahlcom had the smallest sample size of tweets with only eight during the week. However, it had the highest average misinformation per tweets because three of its eight tweets contained misinformation. The account @RedState had the highest number of tweet linked articles containing misinformation, with 12 articles having evidence of misinformation. The account @theblaze had the least amount of misinformation within the Trump-neutral accounts with only one instance. Overall, the Trump-neutral accounts had fewer tweets regarding former president Donald Trump yet posted more tweets with misinformation.

Table 5

Misinformation Analysis in Trump-Neutral Accounts

| Twitter Users | Number of posts with misinformation | Percentage of posts with misinformation | Number of posts |
|----------------|-------------------------------------|---|-----------------|
| @townhallcom | 3 | 37.50% | 8 |
| @RedState | 12 | 26.09% | 46 |
| @BreitbartNews | 4 | 5.48% | 73 |
| @TheBlaze | 1 | 1.41% | 71 |
| Total | 20 | 10.10% | 198 |

The quantitative data from the analyzed tweets showed that one side published a greater proportion of misinformation than the other. The Trump-neutral accounts published 16 more tweets containing misinformation even though they had posted 20 fewer articles in the week of the sample study. Proportionally, that means in one week 10.10% of the tweets posted by the Trump-neutral accounts mentioning Trump contained misinformation. In comparison, only 1.83% of Trump-specific tweets by the anti-Trump accounts contained misinformation.

One glaring difference between the two categories was the type of misinformation being posted. While misinformation from the anti-Trump accounts did not focus primarily on any subjects, the misinformation from the Trump-neutral accounts had a number of posts centering on misinformation about the 2020 presidential election, and misinformation about the Russian collusion (see Table 6). For example, an article posted by @BreitbartNews claimed outright that the Democrats cheated to win the 2020 presidential election and always cheat during elections though no such evidence has proven this

claim. Additional claims from an article posted by @RedState claimed that Trump did not lose the election and that it was stolen from him.

Though sixty-five percent of misinformation posts made by the Trump-neutral accounts were not like each other (see Table 6), there were still cases of similar posting. For example, 25% of all the misinformation posted from the Trump-neutral accounts was about the 2020 election. Furthermore, 10% of the misinformation posted by the Trump-neutral accounts was about the Russian collusion investigation. The other instances of misinformation posted by the Trump-neutral accounts was more varied.

This is in contrast with the anti-Trump accounts who's posts with misinformation was not similar. For example, an article posted by @Salon claimed that the American Psychiatric Association prohibited psychiatrists to speak about Donald Trump's mental health, which is untrue. Another article by @Salon claimed that one of the people who stormed that Capitol, Ashli Babbitt, was from Maryland when she was from California.

Adding to the points made above, not only were the Trump-neutral accounts more likely to post misinformation that was focused on specific subjects, but a number of the topics they posted about were more favorable towards Trump. Though the accounts were categorized as having neutral sentiment towards Trump, the posts they made were typically friendly to him and even advanced views Trump had. This discrepancy illustrates a limitation with sentiment analysis because the approach used here focused on the tone of the text rather than the content. This was only made clearer by the fact that the anti-Trump accounts did not post about similar topics, nor was their misinformation related to that of the Trump-neutral accounts (see Table 6).

Table 6

Types of Misinformation Found

| Categorization | Misinformation about 2020 election | Misinformation about Russian Collusion | Other | Total posts containing misinformation |
|----------------|------------------------------------|--|--------|---------------------------------------|
| Trump-neutral | 25.00% | 10.00% | 65.00% | 20 |
| Anti-Trump | 0.00% | 0.00% | 100% | 4 |
| Total | 20.83% | 8.33% | 70.83% | 24 |

Discussion

This research examined the prevalence of misinformation within eight Twitter accounts across seven days using a consecutive day sampling method and applied the relatively new tool of machine learning to conducting of sentiment analysis. To accomplish this, the two research questions focused on the dissemination of misinformation during and immediately after the Capitol riots to see how it might differ between Trump-neutral and anti-Trump categorized accounts. The findings help to expand the research on misinformation by providing a case study and documented evidence of how disseminated misinformation appears from accounts with certain sentiments. It also provided an illustration of certain limitations of machine learning as a sentiment analysis tool. Specifically, how the classification of the text through sentiment analysis can contrast with the actual content of the text.

The first research question, “how does Twitter sentiment regarding Donald Trump differ between anti-Trump and Trump-neutral accounts,” relates to prior studies in sentiment analysis of Twitter and tweets. Gautam and Yadav (2014) proposed that sentiment analysis could be used as a system for the binary classification of text data, which involved assigning opinions such as negative or positive to the

data. In comparison to their work, the data assigned to sentiment classifications in this research came in the form of tweets. Elsewhere in the research on sentiment analysis, Pak and Paroubek (2010) advanced the binary classification model and were able to categorize text data as either positive or negative or neutral. That work was adapted by this research to analyze the sentiment of tweets from the selected accounts using the keyword “Trump.”

This research was built on existing applications of machine learning to sentiment analysis to determine the sentiment of tweets mentioning Donald Trump by specific Twitter users. The use of sentiment is valuable in the digital age, according to Liu (2012), because it can analyze a user's sentiments and emotions from written language (tweets). The sentiment analysis of tweets this in study also expanded on an approach from Drus and Khalid (2020) to analyze and classify the tweets' sentiment as either positive, negative, or neutral (see Table 2). That research extracted sentiment from text, and this study expanded on that to analyze the extraction of sentiment from tweets to then use to determine the overall sentiment of the account. Though the parameters of this research could classify tweets as either positive, negative, or neutral, the outcome of the dominant sentiment was only as neutral or negative.

The second research question, “how did the spread of misinformation on Twitter differ between anti-Trump and Trump-neutral Twitter accounts in the week after January 6th, 2021,” examined the dissemination of misinformation on the day of, and immediately following, the Capitol riots. This question was proposed to see how the spread of misinformation differed between the Trump-neutral and anti-Trump categorized accounts that were classified by the sentiment analysis. These findings expanded upon prior research into misinformation on Twitter by providing evidence of how disseminated misinformation appears from accounts with certain dominant sentiments.

Huang and Carley (2020) concluded that conservative Twitter users are more likely to engage and disseminate misinformation, and Grinberg et al. (2019) and Bovet and Maske (2019) established that Trump supporters and conservatives are more likely to engage with, spread, and be influenced by misinformation on Twitter. These past findings by Grinberg et al. and Bovet and Maske are consistent with the findings of this research, which showed that the Trump-neutral accounts posted more

misinformation than accounts that were anti-Trump (see Table 6). This research, through sentiment analysis, was able to examine misinformation between Twitter users who were neutral in their sentiment towards Trump and those who were negative toward him. The results show the accounts who are Trump-neutral (exclusively conservative in their political leaning) disseminated more misinformation than the liberal-leaning anti-Trump accounts, as seen in Table 4 and Table 5.

It is not only noteworthy that in this research the Trump-neutral accounts posted more tweets containing misinformation than the anti-Trump accounts, but that some of the misinformation was connected to right-wing narratives. For example, some posts contained posts about the 2020 election being rigged and the Russian collusion investigation being a hoax. Additionally, these accounts seemingly echoed some views promoted by Trump himself. However, as seen in Table 6, many of the posts from the Trump-neutral accounts did not contain misinformation about the 2020 presidential election or Russian collusion.

Continuing with the statements made above, this research illustrated the idea that neutral is not neutral, and because of these instances there was a noted shortcoming in sentiment analysis. It is evident from the results of the sentiment analysis of Trump-neutral accounts that it is possible that the news outlets may be using tone to position themselves as a more credible, neutral source of journalism. According to Ekström et al. (2020), this is done by normalizing and legitimizing political views of the radical right through distinctive practices. Additionally, Liebhart (2019) suggests that through changes in political discourses by the right, a normalization of their ideologies occurs. These findings are in line with this research in that the normalization in the tone of these accounts transforms the sentiment into a more neutral appearing manor.

The fact that in this study anti-Trump accounts posted less misinformation than the Trump-neutral accounts is meaningful due to its connection with established research on the ideological differences between conservatives and liberals. The conservative Trump-neutral accounts posted five times the amount of misinformation during the consecutive day sampling than the anti-Trump accounts. Conservative ideological online practices relate to this research because like the results of this research,

conservatives are more likely than liberals to post misinformation (Robbins, 2017). This can be dangerous because not only are conservatives more apt to engage with misinformation, but also because conservative users have a significantly more uniform echo chamber online (Jost et al., 2017). So, the misinformation is circulating around their online networks where it is likely that more and more users will engage with the misinformation. According to Jost et al., not only do conservatives engage in the dissemination of misinformation but they are also more likely to endorse and act on the misinformation. This ideological asymmetry and relationship between conservatives and online misinformation can have consequences when it comes to users' political judgement and actions (Jost et al., 2017).

Limitations and Future Research

Because of its use of the simple random sampling method, this research had limitations. First, this research cannot be used to assess the overall sentiment of these accounts towards Trump. This research analyzed a small sample size due to time constraints and because it is a manageable method to extract research data from a large population. Each account had numerous tweets, across more than four years mentioning and about Donald Trump. Therefore, this research decided to use such a limited sample. Rather than examining a representative sample, this research provided a descriptive case study of accounts on Twitter and their misinformative posts in the wake of the Capitol riots. However, even in the small sample, this research still provides a solid illustration of the sentiment.

Because only 31 tweets were analyzed by the sentiment analysis program, their sentiment towards Trump cannot be 100% certain. It is entirely possible that a more expansive sentiment analysis of tweets from these accounts could produce different results and classifications than the ones present in this study. There has been over four years of tweets from these accounts pertaining to Trump during his presidency that could easily shift the sentiment towards him which could have resulted in a different study. It is possible, for example, that a Trump-neutral account could have ended up being dominantly positive towards him if a complete record of tweets were documented. Though these possibilities exist, it was

advantageous for this research to use the simple random sampling method to provide a detailed and intense case study into misinformation on Twitter that can be used to provide an exploratory look at the accounts during a short, targeted period.

Another limitation is the number of accounts studied. This study was only able to analyze eight accounts because of the sampling adjustments described above. The findings from this study still provide a useful snapshot of misinformation on Twitter, but they cannot be generalized broadly. Further studies could expand on this research to gain better insight and a sounder conclusion. With more accounts studied, the results could be different.

A final limitation is the sample of days that were chosen. Because the original, more extensive sampling plan was not possible, the research was instead focused on a purposive sample of January tweets to focus on pre-inauguration issues. To only analyze one week out of the entire year for misinformation provides a useful case study but not a generalizable sample.

Conclusion

Combating and correcting misinformation is a crucial function of the press in a democratic society (Kovach & Rosenstiel, 2010), and misinformation is constantly evolving. The need for curbing and understanding misinformation on social media is evident, especially now because all forms of news media have seen a decrease in use by people except for online news (Pew, 2011). This study can serve as a base for future research into misinformation in journalism as well as the application of machine learning and sentiment analysis.

First, the method used to categorize the sentiment of the selected accounts and collect the data will be useful to future researchers and the study of misinformation, especially on Twitter. This research expanded on the study of Drus and Khalid (2020), which extracted sentiment from text and classified its sentiment to then extracting sentiment from tweets. The sentiment classification of tweets was then applied to characterizing the users' overall sentiment to create a sentiment profile of the user. According to Liu (2012), this can be used to analyze Twitter users' sentiments from their own tweets and to create a

classification of them on a number of subjects. However, it is noted in this research that although the sentiment analysis can analyze the tone of the tweet, it is limited because the content could contrast with the sentiment (as seen in Table 6).

This research also serves the study of how journalism and journalists on Twitter curate their accounts. Future studies could examine accounts and tweets from journalists and media outlets using sentiment analysis to determine conclusions on their objectivity and biases. This identification would be able to give the public and readers valuable context into what they're engaging with. According to Gangadharbatla et al. (2014), people now prefer to gather news and information on social media and knowing the sentiment of the sources they're getting their information and news from would provide better context and overall credibility to their news gathering. This research's focus on Twitter could be used across social media sites such as Facebook or Instagram, which is useful because most American adults get their news from social media (Matsa & Shearer, 2018).

Finally, future research could examine the effects of exposure to these misinformative accounts. While it's been suggested that misinformative news was unlikely to have truly impacted the 2016 U.S. presidential election (Allcott & Gentzkow, 2017), the continued exposure of misinformation can still erode the quality of public debate, promote misperceptions, foster greater hostility toward political opponents, and corrode trust in government and journalism (Guess et al., 2019). For instance, the Capitol riots were perpetuated, at least in part, by misinformation that circulated on social media up to that point (Reuters, 2021). This showed a real-world effect misinformation can produce. Because of such a recent example of misinformation's power, the results of this study will contribute to the overall study of the spread of misinformation and how it functions on social media.

References

- Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31 (2): 1–28.
- Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. 2015. "Exposure to ideologically diverse news and opinion on Facebook." *Science* 348 (6239): 1130–1132.
- Bovet, A., Makse, H. A., *Nature communications* 10, 1 (2019).
- Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2018). *PRACTICAL GUIDE TO SENTIMENT ANALYSIS* (Vol. 5). Place of publication not identified, UK: SPRINGER INTERNATIONAL PU.
- Coddington, M., Molyneux, L., & Lawrence, R. G. (2014). Fact Checking the Campaign: How Political Reporters Use Twitter to Set the Record Straight (or Not). *The International Journal of Press/Politics*, 19(4), 391–409. <https://doi.org/10.1177/1940161214540942>
- Drus, Z., & Khalid, H. (2020). Sentiment Analysis in Social Media and Its Application: Systematic Literature Review. Retrieved November 13, 2020, from <https://www.sciencedirect.com/science/article/pii/S187705091931885X>
- Ekström, M., Patrona, M., & Thornborrow, J. (2020, May 26). *The normalization of the populist radical right in news interviews: a study of journalistic reporting on the Swedish democrats*. Taylor & Francis. <https://www.tandfonline.com/doi/full/10.1080/10350330.2020.1762984>.
- Gangadharbatla, H., Bright, L., & Logan, K. (2014). Social Media and News Gathering: Tapping into the Millennial Mindset. Retrieved November 13, 2020, from <https://thejsms.org/tsmri/index.php/TSMRI/article/view/63/33>
- Gottfried, J., Mitchell, A., Kiley, J., & Matsa, K. (2020, August 28). Political Polarization & Media Habits. Retrieved November 13, 2020, from <https://www.journalism.org/2014/10/21/political-polarization-media-habits/>

- Gautam, G., & Yadav, D. (2014). Sentiment analysis of twitter data using machine learning approaches and semantic analysis. *2014 Seventh International Conference on Contemporary Computing (IC3)*, 437-442.
- Guess, A., Nyhan, B., & Reifler, J. (2019). Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign. Retrieved November 13, 2020, from <http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf>
- Grinberg, N., Joseph, L. Friedland, B. Swire-Thompson, D. Lazer, *Science* 363, 374 (2019).
- Hart, William, Dolores Albarracín, Alice H. Eagly, Inge Brechan, Matthew J. Lindberg, and Lisa Merrill. 2009. "Feeling validated versus being correct: a meta-analysis of selective exposure to information." *Psychological Bulletin* 135 (4): 555.
- Hermida, A. From tv to twitter: How ambient news became ambient journalism. *M/C Journal* 13, 2 (2010)
- Holan, A. (2018). PolitiFact - The Principles of the Truth-O-Meter: PolitiFact's methodology for independent fact-checking. Retrieved December 10, 2020, from <https://www.politifact.com/article/2018/feb/12/principles-truth-o-meter-politifact-methodology-i/>
- Huang, B., & Carley, K. (2020, June 07). Disinformation and misinformation on Twitter during the novel Coronavirus outbreak. Retrieved April 13, 2021, from <https://arxiv.org/abs/2006.04278>
- Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Retrieved November 13, 2020, from <http://eegilbert.org/papers/icwsm14.vader.hutto.pdf>
- Iyengar, Shanto, and Kyu S. Hahn. 2009. "Red media, blue media: Evidence of ideological selectivity in media use." *Journal of Communication* 59 (1): 19–39.

- Iyengar, Shanto, Kyu S. Hahn, Jon A. Krosnick, and John Walker. 2008. "Selective exposure to campaign communication: The role of anticipated agreement and issue public membership." *Journal of Politics* 70 (01): 186–200
- Jost, Linden, Panagopoulos, & Hardin, (2018, January 12). Ideological asymmetries in conformity, desire for shared reality, and the spread of misinformation. Retrieved April 19, 2021, from <https://www.sciencedirect.com/science/article/pii/S2352250X17302828>
- Journalism, 'Fake News' and Disinformation: A Handbook for Journalism Education and Training. (2020, October 15). Retrieved November 13, 2020, from <https://en.unesco.org/fightfakenews>
- Kaity, M., Balakrishnan, V. Sentiment lexicons and non-English languages: a survey. *Knowl Inf Syst* 62, 4445–4480 (2020). <https://doi.org/10.1007/s10115-020-01497-6>
- Kamps, J., Marx, M., Mokken, R. J., & De Rijke, M. (2004). Using WordNet to Measure Semantic Orientations of Adjectives. Retrieved November 15, 2020, from <http://www.lrec-conf.org/proceedings/lrec2004/pdf/734.pdf>
- Karlova, N., & Fisher, K. (2013). A social diffusion model of misinformation and disinformation for understanding human information behaviour. Retrieved November 13, 2020, from <http://www.informationr.net/ir/18-1/paper573.html>
- Kovach, B., and T. Rosenstiel. 2010. *Blur: How to Know What's True in the Age of Information Overload*. New York: Bloomsbury Books.
- Liebhart, K. (2019, July 9). *The Normalization of Right-Wing Populist Discourses and Politics in Austria*. Springer Link. https://link.springer.com/chapter/10.1007/978-3-658-25419-3_4.
- Liu, B. (2012, May). Sentiment analysis and opinion mining. Retrieved April 18, 2021, from <https://www.morganclaypool.com/doi/abs/10.2200/S00416ED1V01Y201204HLT016>

- Kessler, G. (2017a, January 1). About the Fact Checker. The Washington Post. Retrieved from <https://www.washingtonpost.com/politics/2019/01/07/about-fact-checker/>
- Kuklinski, J. H., P. J. Quirk, J. Jerit, D. Schwieder, and R. F. Rich. 2000. "Misinformation and the Currency of Democratic Citizenship." *Journal of Politics* 62 (3): 790–816.
- Kwak, H., Lee, C., Park, H., and Moon, S. What is twitter, a social network or a news media? In Proceedings of the 19th international conference on World wide web, WWW '10, ACM (New York, NY, USA, 2010), 591–600.
- Mishra, N., & Jha, C. K. (2012). Classification of opinion mining techniques. *International Journal of Computer Applications*, 56(13).
- Munn, L. (2021). More than a mob: Parler as preparatory media for the U.S. Capitol storming. *First Monday*, 26(3). <https://doi.org/10.5210/fm.v26i3.11574>
- N/A., 1.9. Naive Bayes¶. (n.d.). Retrieved November 15, 2020, from https://scikit-learn.org/stable/modules/naive_bayes.html
- Pak, A., & Paroubek, P. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. Retrieved November 15, 2020, from http://www.lrec-conf.org/proceedings/lrec2010/pdf/385_Paper.pdf
- Pariser, Eli. 2011. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). The development and psychometric properties of LIWC2007. Austin, TX: LIWC.net. Pennebaker, J. W., Francis, M., & Booth, R. (2001). *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NJ: Erlbaum.
- Posetti, J., & Matthews, A. (2018). A short guide to the history of 'fake news' and disinformation. Retrieved April 06, 2021, from <https://www.icfj.org/sites/default/files/2018->

07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ
%20Final.pdf

Riffe, D., Aust, C. F., & Lacy, S. R. (1993). The Effectiveness of Random, Consecutive Day and Constructed Week Sampling in Newspaper Content Analysis. *Journalism Quarterly*, 70(1), 133–139.
<https://doi.org/10.1177/107769909307000115>

Robbins, J. (2017, December 1). Brexit voters and Britain First supporters are more likely to believe conspiracy theories. *International Business Times*. <https://amp.ibtimes.co.uk/brexit-voters-britain-first-supporters-are-more-likely-believe-conspiracy-theories-1649106>

Roose, K. (2018, November 04). We Asked for Examples of Election Misinformation. You Delivered. Retrieved December 10, 2020, from <https://www.nytimes.com/2018/11/04/us/politics/election-misinformation-facebook.html>

Shearer, E. (2020, August 27). Social media outpaces print newspapers in the U.S. as a news source. Retrieved November 13, 2020, from <https://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/>

Stroud, Natalie Jomini. 2008. “Media use and political predispositions: Revisiting the concept of selective exposure.” *Political Behavior* 30 (3): 341–366.

Sunstein, C. 2001. *Republic.com*. Princeton University Press.

Sunstein, C. 2007. *Republic.com 2.0*. Princeton, NJ: Princeton University Press.

Tandoc, E., Wei Lim, Z., & Lang, R. (2017). Defining "Fake News." A Typology of Scholarly Definitions. Retrieved November 13, 2020, from <https://www.rcmediafreedom.eu/Publications/Academic-sources/Defining-Fake-News.-A-Typology-of-Scholarly-Definitions>

- Turney, P. D. (2002). Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. Retrieved November 15, 2020, from <https://www.aclweb.org/anthology/P02-1053.pdf>
- Utz, S., Schultz, F. and Glocka, S. (2013). Crisis communication online: How medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster. *Public Relations Review*, 39(1), pp. 40–46.
- Weeks, B. E., & Holbert, R. L. (2013). Predicting Dissemination of News Content in Social Media: A Focus on Reception, Friending, and Partisanship. *Journalism & Mass Communication Quarterly*, 90(2), 212–232. <https://doi.org/10.1177/1077699013482906>
- What is the purpose of journalism? (2017, July 18). Retrieved from <https://www.americanpressinstitute.org/journalism-essentials/what-is-journalism/purpose-journalism/>
- Wu, et al., (2019). Misinformation in Social Media: Definition, Manipulation, and Detection. Retrieved 2020, from https://www.kdd.org/exploration_files/8._CR.10.Misinformation_in_social_media_-_Final.pdf
- Zubiaga, A., Ji, H., & Knight, K. (2013, March 01). Curating and contextualizing twitter stories to assist with social newsgathering. Retrieved April 05, 2021, from https://dl.acm.org/doi/abs/10.1145/2449396.2449424?casa_token=8P7fA2IIULUAAAAA%3A604a5vgaYVKPn7-4K5t_2GsIHGDuxDZ04HjphcRoOWGwTS-htxY4pyfSAM3RGVK2JULm0q6ZNJc8Yg
- Zubiaga, A., Spina, D., Fresno, V., and Martínez, R. Classifying trending topics: a typology of conversation triggers on twitter. In *Proceedings of the 20th ACM international conference on Information and knowledge management, CIKM '11*, ACM (New York, NY, USA, 2011), 2461–2464.
- Zubiaga, A. (2019). Mining Social Media for Newsgathering: A Review. Retrieved November 13, 2020, from <https://arxiv.org/pdf/1804.03540.pdf>

Appendix A: Tables

Table 1

Sentiment Analysis Examples

| Twitter User | Tweets | Overall Assessment |
|--------------|--|--------------------|
| @AlterNet | Complete incompetence': Biden officials fume after Trump leaves them 'nonexistent' vaccine plan | Negative 95% |
| @townhallcom | Even if Trump disappears, his impact on the GOP is lasting. | Positive 58% |
| @dailykos | Voting Rights Roundup: Biden rescinds Trump order that tried to weaponize the census against Latinos | Neutral 72% |

Table 2

Sentiment From Sampled Twitter Accounts.

| Twitter Account | Positive | Neutral | Negative |
|-----------------|------------|-------------|------------|
| @OANN | 1 (3.23%) | 28 (90.32%) | 2 (6.45%) |
| @RedState | 4 (12.90%) | 23 (74.19%) | 4 (12.90%) |
| @democracynow | 3 (9.68%) | 23 (74.19%) | 5 (16.13%) |
| @BreitbartNews | 2 (6.45%) | 23 (74.19%) | 6 (19.35%) |

| | | | |
|-----------------|------------|--------------|--------------|
| @theblaze | 0 (0.00%) | 22 (70.97%) | 9 (29.03%) |
| @townhallcom | 4 (12.90%) | 22 (70.97%) | 5 (16.13%) |
| @TheYoungTurks | 2 (6.45%) | 21 (67.74%) | 8 (25.81%) |
| @RawStory | 0 (0.00%) | 6 (13.35%) | 25 (80.65%) |
| @newrepublic | 1 (3.23%) | 10 (32.26%) | 20 (64.52%) |
| @AlterNet | 5 (16.13%) | 7 (22.58%) | 19 (61.29%) |
| @Salon | 2 (6.45%) | 11 (35.48%) | 18 (58.06%) |
| @dailykos | 2 (6.45%) | 11 (35.48%) | 18 (58.06%) |
| @crooksandliars | 4 (12.90%) | 12 (38.71%) | 15 (48.39%) |
| Total | 30 (7.44%) | 219 (54.34%) | 154 (38.21%) |

Table 3

Categorization of Twitter Accounts and Their Follower Count

| Organization | Classification | Political Leaning | Twitter Follower Count |
|-----------------|----------------|-------------------|------------------------|
| @dailykos | Anti-Trump | Liberal | 292.4K |
| @crooksandliars | Anti-Trump | Liberal | 94.5K |
| @TheYoungTurks | Trump-Neutral | Liberal | 444.7K |
| @newrepublic | Anti-Trump | Liberal | 176.4K |

| | | | |
|----------------|---------------|--------------|--------|
| @RawStory | Anti-Trump | Liberal | 222.3K |
| @Salon | Anti-Trump | Liberal | 977.2K |
| @democracynow | Trump-Neutral | Liberal | 788.6K |
| @AlterNet | Anti-Trump | Liberal | 137.9K |
| @RedState | Trump-Neutral | Conservative | 233.7K |
| @BreitbartNews | Trump-Neutral | Conservative | 1.4M |
| @OANN | Trump-Neutral | Conservative | 1.5M |
| @theblaze | Trump-Neutral | Conservative | 690.8K |
| @Townhallcom | Trump-Neutral | Conservative | 168.7K |

Table 4

Misinformation Analysis in Anti-Trump Accounts

| Twitter Users | Number of posts with misinformation | Percent of posts with misinformation | Number of posts |
|---------------|-------------------------------------|--------------------------------------|-----------------|
| @newrepublic | 1 | 3.70% | 27 |
| @Salon | 2 | 2.82% | 71 |
| @dailykos | 1 | 1.54% | 65 |
| @AlterNet | 0 | 0.00% | 58 |
| Total | 4 | 1.83% | 218 |

Table 5

Misinformation Analysis in Trump-Neutral Accounts

| Twitter Users | Number of posts with misinformation | Percentage of posts with misinformation | Number of posts |
|----------------|-------------------------------------|---|-----------------|
| @townhallcom | 3 | 37.50% | 8 |
| @RedState | 12 | 26.09% | 46 |
| @BreitbartNews | 4 | 5.48% | 73 |
| @TheBlaze | 1 | 1.41% | 71 |
| Total | 20 | 10.10% | 198 |

Table 6

Types of Misinformation Found

| Categorization | Misinformation about 2021 election | Misinformation about Russian Collusion | Other | Total posts containing misinformation |
|----------------|------------------------------------|--|--------|---------------------------------------|
| Trump-neutral | 25.00% | 10.00% | 65.00% | 20 |
| Anti-Trump | 0.00% | 0.00% | 100% | 4 |
| Total | 20.83% | 8.33% | 70.83% | 24 |