SENTIMENT ANALYSIS OF TWEETS FROM THE 2020 PRESIDENTIAL ELECTION

by

Joseph M. Williams

A Dissertation Submitted to the Faculty of

The Charles E. Schmidt College of Science

in Partial Fulfillment of the Requirement for the Degree of

Doctor of Philosophy

Florida Atlantic University

Boca Raton, FL

May 2022

SENTIMENT ANALYSIS OF TWEETS FROM THE 2020 PRESIDENTIAL ELECTION

by

Joseph M. Williams

This dissertation was prepared under the direction of the candidate's dissertation co-advisors, Dr. Andrzej Nowak, Department of Psychology, and Dr. Robin Vallacher, Department of Psychology, and has been approved by all members of the supervisory committee. It was submitted to the faculty of the Charles E. Schmidt College of Science and was accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

SUPERVISORY COMMITTEE:

Andrzej Nowak
Andrzej Nowak (Apr 22, 2022 18:45 GMT+2)

Andrzej Nowak, Ph.D.
Dissertation Co-Advisor

Robin Vallacher

Robin Vallacher, Ph.D.
Dissertation Co-Advisor

Michael R. Maniaci
Michael R. Maniaci (Apr 22, 2022 16:38 EDT)

Michael Maniaci, Ph.D.

Elan Barenholtz
Elan Barenholtz (Apr 30, 2022 21:38 GMT+3)

Elan Barenholtz, Ph.D.

WH
William Hahn (May 3, 2022 09:24 EDT)

William Hahn, Ph.D.

Robin Vallacher

Robin Vallacher, Ph.D.
Interim Chair, Department of Psychology

Teresa Wilcox

Teresa Wilcox, Ph.D.
Interim Dean, Charles E. Schmidt College of Science

Robert W. Stackman

Robert W. Stackman Jr., Ph.D.
Dean, Graduate College

May 3, 2022

Date

ACKNOWLEDGEMENTS

I would like to thank my committee members, Dr. Elan Barenholtz, Dr. William Hahn, and Dr. Maniaci for their invaluable input. Special thanks to my advisors, Dr. Andrzej Nowak and Dr. Robin Vallacher for their advice and encouragement. Lastly, I thank my wife Jamie for her patience, love, and support. I could not have done this without her.

ABSTRACT

Author:	Joseph M. Williams
Title:	Sentiment Analysis of Tweets from the 2020 Presidential Election
Institution:	Florida Atlantic University
Dissertation Advisor:	Dr. Andrzej Nowak and Dr. Robin Vallacher
Degree:	Doctor of Philosophy
Year:	2022

We collected approximately 60 million tweets over a 6-month period during the 2020 presidential election, starting from the conventions through the inauguration. Each tweet referenced either the Republican incumbent Donald Trump or the Democrat challenger Joe Biden. The tweets were analyzed for sentiment and the frequency of the moral foundations using the standard LIWC2015 dictionary and the Moral Foundations Dictionary 2.0. We found that the tweets had an overall negative sentiment for both candidates, with tweets referencing Trump being more negative than tweets about Biden. Additional analyses showed that the Authority (Virtue) and Loyalty (Virtue) were the most frequently used moral foundations. This study provides an overview into social media discussions during a heated election cycle that ultimately culminated in the Jan. 6th Insurrection and the second impeachment of Donald Trump.

SENTIMENT ANALYSIS OF TWEETS FROM THE 2020 PRESIDENTIAL ELECTION

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This study will demonstrate the benefit of using social media (i.e., Twitter) data as a research tool. Our procedure for analyzing the tweets will utilize well-established and validated methods for computerized text analysis (the Linguistic Inquiry and Word Count and the Moral Foundations Dictionary 2.0) to provide a comparison to prior research. These methods will enable us to classify the sentiment and moral content for millions of tweets. By collecting the tweets daily over a seven-month period, we will be able to identify temporal patterns of change.

Social media is a widely popular method of modern communication, with nearly three-quarters of Americans having some form of social media (Auxier & Anderson, 2021). Twitter, one of the most popular platforms, is a networking and microblogging site where users post and interact with messages known as "tweets". In 2019, it was reported to have 134 million daily active users and 330 million monthly active users. By 2021, the number of daily active users increased to just under 200 million (Twitter, 2019; Twitter, 2021). Twitter is regularly used by politicians, athletes, and other celebrities. Notably, former President Donald Trump (with the username *@realDonaldTrump*) averaged dozens of tweets per day until the account's suspension in January 2021.

The massive popularity of social media sites inspired new approaches for conducting psychological research. Social media usage is indicative of personality characteristics (Kosinski et al., 2013; Youyou et al., 2015). Social media research, using Twitter in particular, can identify dynamical changes in attitudes (Serfass & Sherman, 2015; Wang et al., 2016), predict political ideology (Slywester & Purver, 2015), and measure the frequency of moral concepts (Kaur & Sasahara, 2016). The current study expands upon prior research by providing an exploratory overview of the sentiment of tweets about either candidate during the 2020 Presidential election. We also sought to identify how frequently each of the moral foundations (Graham et al., 2009) were used when discussing either candidate.

This study provides new insight into the attitudes expressed on social media throughout the 2020 presidential election, starting from the conventions until the inauguration. The 2020 presidential election cycle was filled with several unanticipated historically significant events, including but not limited to a global pandemic, an insurrection at the US Capitol building, the first time a president was impeached twice, and record turnout during the general election.

The current study built upon previous research on the dynamics of social media, the moral foundations, linguistic analysis, and political and social psychology. Particularly, this work built upon the moral foundations theory which identified differences between liberals and conservatives. We collected tweets over the course of a seven-month time frame, ranging from August 1st, 2020, to February 5th, 2021. There were an equivalent number of weeks before and after the general election on November 3rd, 2020. This timeframe covered the major events that normally occur during a U.S. presidential election cycle, including the Republican and Democratic conventions, the debates, and the inauguration. We also collected tweets during several unanticipated events such as the Jan. 6th Insurrection, the second impeachment of then President Donald Trump, and the continuing progression of the COVID-19 pandemic.

Linguistic Inquiry Word Count

Psychological factors can be identified in bodies of text using linguistic analysis techniques. One of the most common is the Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2015), which identifies the frequency of words of various categories indicating social, cognitive, and affective processes. Dictionary-driven text analysis is beneficial because the dictionaries can be reused across various datasets and studies (Kennedy et al., 2021). LIWC in particular provides a widely used and validated method for analyzing the usage of written and spoken words and has been used in studies across a variety of topics and cultures (Serfass et al, 2015; Yarkoni, 2010)

Computerized versions of LIWC and other dictionary-based programs facilitate the replication of the same analysis across different datasets. Tausczik & Pennebaker (2010) demonstrated that LIWC accurately identifies emotion in language use. For example, positive emotion words (e.g., love, nice, sweet) are used in writing about a positive event, and more negative emotion words (e.g., hurt, ugly, nasty) are used in writing about a negative event. This is further supported by Kahn et al. (2007). Automated LIWC ratings of positive and negative emotion words correspond with human ratings of writing excerpts (Alpers et al., 2005).

In this study, we used the program LIWC2015 to analyze the emotional content of the tweets. LIWC has been validated across cultures (Igarashi et al., 2015) and with different forms of media (Meier et al., 2019). This method is also beneficial because it enabled us to analyze a large sample of data in a timely manner.

LIWC2015 is the latest in a series of text analysis programs developed by Pennebaker et al. (2015), which uses the Linguistic Inquiry and Word Count to determine the frequency of psychological and linguistic concepts in a body of text. LIWC2015 has two primary components: the dictionary and the processing module. The LIWC2015 dictionary contains a list of words related to various psychological and linguistic concepts (e.g., affective processes, parts of speech, etc.). Scores for each category are calculated as the percentage of words in a text that fall in a specific category. For this study, we focused on the categories of Positive Emotion, Negative Emotion, Anger, Anxiety, and Sadness. We also examined the summary variable Tone to identify overall changes in sentiment. LIWC2015 includes the ability to import additional dictionaries. To this end, we analyzed the tweets twice, once using the default LIWC2015 dictionary and again with the Moral Foundations Dictionary 2.0 (MFD 2.0; Frimer et al., 2019). Researchers can learn more about an individual or a group by understanding the moral concepts they prioritize or use.

Moral Foundations Theory

Linguistic analysis can also be extended to identify moral concepts. One of the leading theories is the Moral Foundations Theory (Haidt & Joseph, 2004; Haidt & Graham, 2007), which argues that people have five distinct, innate moral foundations:

- Care: caring for others and protecting them from harm.
- Fairness: showing equality in one’s actions and judgments towards others.
- Loyalty: having allegiance to one’s ingroup, family, and community.
- Authority: having respect for tradition and legitimate authority.
- Sanctity: avoiding disgusting things, foods, actions.

The Moral Foundations Questionnaire was developed by Graham et al., (2009) and later validated by Graham et al. (2011).

People emphasize and pull from different combinations of the five moral foundations and the moral foundations an individual draws on to make moral decisions are relatively accurate predictors of that person's political ideology. Liberals tend to emphasize the care and fairness foundations (referred to as "individualizing foundations") more than the loyalty, authority, and sanctity foundations (referred to as "binding foundations"). In contrast, conservatives tend to emphasize all of the moral foundations more equally (Haidt & Graham, 2007; Graham et al., 2009; Haidt, 2012; Neiman et al., 2016).

Liberals tend to score higher than conservatives on the care and fairness foundations, but lower on the ingroup, authority, and sanctity foundations. Liberals put more emphasis on caring for others and protecting them from care, as well as executing justice than on the other moral foundations, whereas conservatives are guided by all moral foundations to a similar extent (Slywester & Purver, 2015; Kraft, 2018). Endorsement of the moral foundations can predict support for political issues (e.g., abortion, gun control, etc.) more accurately than factors such as ideology, religiosity, etc. (Koleva et al., 2012).

The Moral Foundations Dictionary was first used to compare the moral concepts used in sermons delivered in liberal (e.g., Unitarian) and conservative (e.g., Southern Baptist) churches. Graham et al. (2009) developed several dictionaries reflecting the moral foundations and used a computerized version of LIWC to calculate word frequencies. Word frequency analysis yielded support for the direction of differences in care, fairness, authority, and sanctity, but not the loyalty foundation. Loyalty-related words were used more frequently in liberal than conservative sermons. However, they noted that liberal preachers were rejecting ingroup values instead of endorsing them.

The dominance of specific moral foundations can vary based on the topic of discussion. For example, Kaur and Sasahara (2016) found that Care was the most frequently used foundation in tweets about immorality. Clifford & Jerrit (2015) examined how opposing sides of the stem cell debate used different moral words to sway public opinion. They found that supporters of stem cell research focused almost exclusively on the Care foundation to bolster their position. In contrast, opponents relied mainly on the Care foundation, and the Sanctity foundation to a lesser extent. The Sanctity foundation was infrequently used overall despite similarities between it and stem cell attitudes. This was further supported by Sagi and Dehghani (2014). Brown and Silver (2020) found that Authority and Care were the most dominant foundations in discussions of crime control.

Lewis (2019) compared the differences in candidates' moral rhetoric during the 2015-16 United States presidential primary season. The Moral Foundations Dictionary was used to calculate how frequently candidates used words representing various moral foundations, distinguishing between positive and negative references to each. Democratic and Republican candidates used the moral foundations at similar rates, approximately 2% of the total words. The Republican candidates were more likely to use negative-valence moral terminology, describing violations of moral foundations.

The Moral Foundations Dictionary 2.0 (MFD 2.0) is a linguistic analysis dictionary that identifies the usage of five innate, moral foundations in text. The MFD 2.0 (Frimer et al., 2019) is derived from the Moral Foundation Theory (Haidt, 2003; Haidt & Graham, 2007) and the Moral Foundations Dictionary (Haidt et al., 2009), which defined the five moral foundations as Care (sympathy, compassion, and nurturance), Fairness/cheating (notions of equal rights and justice), Loyalty ("us vs. them" mentality), Authority (concerns about traditions and maintaining social order), and Sanctity (moral disgust). The MFD 2.0 contains a list of words related to one or more of the moral foundations such as "killing", "justice", and "loyal," which correspond respectively to the Care, Fairness, and Loyalty foundations. The dictionary contains an average of 210 words per category. The MFD 2.0 is compatible with the LIWC2015 software and follows the same scoring procedures as LIWC. However, the MFD uses a separate user-defined dictionary designed to calculate scores based on the five moral foundations.

Social Media Research

Research using social media has had massive implications for psychological science. Kosinski et al. (2013) studied the role of personality in website preferences by combining personality profiles and website choices from more than 160,000 users and then investigating whether different websites attracted audiences with different personalities. Personality traits of users can be predicted through regression models using Facebook Likes. Youyou et al. (2015) compared the accuracy of human and computer-based personality judgments, using a sample of 86,220 volunteers who completed a 100-item personality questionnaire. They showed that computer predictions based on a generic digital footprint (Facebook Likes) are more accurate than those made by the participants' Facebook friends using a personality questionnaire. Openness and Extraversion were the most predictable personality traits in both studies (Kosinski et al., 2013; Youyou et al., 2015).

Facebook is not the only social media platform that can be used to predict personality. Publicly shared information on Twitter can also predict the Big Five personality traits (Golbeck et al., 2015). They had participants complete a personality test and the researchers collected publicly accessible information from their profiles. They were able to train two machine learning algorithms to predict scores on each of the five personality traits to within 11% - 18% of the results of self-report personality questionnaires.

Lin et al. (2017) used machine learning to assess the predictability of moral traits from linguistic features using a set of manually coded texts from Twitter. They did not assess the orientation of each foundation (i.e., virtue vs vice), but only its presence in the tweet. Their machine learning model was successful at identifying the care, fairness, and authority foundations at similar rates as a human rater. For example, their model was able to successfully detect 76.3% of tweets that used the care foundation compared to the human performance of 76.0%. Their results also suggest that the moral foundations can be expressed in language outside of the Moral Foundations Dictionary.

Several studies have identified temporal patterns of behavior with tweets. Yang and Leskocev (2011) demonstrated how tweets change over a short period of time. The content of tweets is volatile, and topics can quickly become popular and fade away. They found that short quoted textual phrases ("memes") rose and decayed over the course of several days. However, named entities and general themes (i.e., "economy" or "Obama") changed over longer periods of time. Wang et al., (2016) identified dynamic changes in the content of tweets. They showed large increases in work stress and negative emotion expressed on Twitter occurred on Fridays. In contrast, positive emotion Tweets hit their lowest points on Tuesdays, Wednesday, and Thursdays and peaked from Friday through

Sunday, Tweets about work, money, etc. showed a small dipped to their lowest points on Fridays through Sundays.

Serfass at al. (2019) identified cyclical pattern of behavior of tweets using the DIAMONDS model, in which, actions were split into eight categories (duty, sociability, etc.). They collected nearly 20 million Tweets over the course of two weeks. Each tweet was scored on the eight situational DIAMONDS (Duty, Intellect, Adversity, Mating, Positivity, Negativity, Deception, and Sociality) (Rauthmann et al., 2014). They identified several daily and weekly trends. On weekdays, Duty peaked in the midmorning and declines steadily thereafter while Sociality peaked in the evening. Negativity was highest during the middle of the week and lowest on the weekends. In contrast, Positivity was highest on the weekends and lowest in the middle of the week.

Thelwall et al. (2011) examined a longer time scale by analyzing a month of English Twitter posts to determine whether popular events were associated with increases in sentiment strength. They showed that popular events are normally associated with increases in negative sentiment strength. There was also evidence to show that peaks of interest in events have stronger positive sentiment in the time before the peak. Many positive events (e.g., the Oscars) can generate increased negative sentiment in reaction to them.

Sano et al. (2019) extracted collective emotion from Japanese blog articles over a 10-year period. They found that collective emotion showed clear periodic cycles (i.e., weekly and seasonal behaviors) accompanied with spikes caused that coincided with several natural disasters. For example, there were increased levels of tension in April corresponding with the start of the Japanese school year.

Social Media and Politics

Since its launch on March 21st, 2006, Twitter has become increasingly intertwined with U.S. politics. Less than a year later, the Obama campaign created the @BarackObama Twitter account on March 5th, 2007. The @WhiteHouse account was created on April 27th, 2007. On May 18th, 2015, Obama sent his first tweet from the first Twitter account dedicated exclusively to the U.S. President (@POTUS). The U.S. presidential campaign of Barack Obama utilized Twitter, Facebook, MySpace as a key component to his victory (Williams & Gulati, 2008).

Tweets can convey real-world attitudes about political topics. They can be predictive of the outcome of elections, but reactive to specific events. For example, Tumasjan et al. (2010) analyzed over 100,000 Twitter messages mentioning parties or politicians prior to the German federal election 2009. They found that the quantity of tweets reflected voter preferences and closely resembled traditional election polls. The sentiment of tweets corresponded with political messaging and media coverage of the campaigns.

O'Connor et al. (2010) developed a sentiment detector based on tweets to predict consumer confidence and presidential job approval polls. They analyzed several surveys on consumer confidence and political opinion between 2008 and 2009 and found that they correlated to sentiment word frequencies in tweets. A different study by Bermingham and Smeaton (2011) modeled political sentiment through mining of social media from the 2010 Irish General Election. Their approach combined sentiment analysis using supervised learning and volume-based measures. They compared their model against the election polls and final election result and found that volume-based measures and sentiment analysis were predictive.

Murthy (2015) found that tweets are reactive rather than predictive in this election context. By examining the timeline of specific primary debates, and elections, they found that tweets were reacting to offline campaign results rather than predicting them.

Yaqub et al. (2017) investigated the sentiment of tweets by the two main presidential candidates, Hillary Clinton and Donald Trump, along with almost 2.9 million tweets by Twitter users during the 2016 US Presidential Elections. They found an overall negative public sentiment towards both candidates and noted that their timeline of Twitter trends and frequently used words corresponded with real world events.

Political Research

A large aspect of political research focuses on identifying differences between liberals and conservatives. Carney et al. (2008) demonstrated that liberals appeared to be more open, tolerant, creative, curious, expressive, enthusiastic, and drawn to novelty and diversity, in comparison with conservatives, who appeared to be more conventional, orderly, organized, neat, clean, withdrawn, reserved, and rigid. Most of these differences fall under the "Big Five" dimensions of Openness and Conscientiousness.

Ditto et al. (2019) examined differences in perceived bias between Democrats and Republicans. They asked participants how the term "biased" described the average Democrat and the average Republican. Respondents describing themselves as Democrats saw the average Republican as substantially more biased than the average Democrat. Republican respondents expressed the mirror image belief that the average Democrat was substantially more biased than the average Republican. Frimer et al. (2017) found that when participants thought about the election and relevant cultural issues, liberals and conservatives reported similar aversion toward learning about the views of their ideological opponents. They found that participants' disinterest was not due to already being informed about the other side or attributable to election fatigue. Participants on both sides indicated that they anticipated that hearing from the other side would induce cognitive dissonance (e.g., require effort, cause frustration). This would undermine a sense of shared reality with the person expressing disparate views, thereby damaging their relationship. This is further supported by a meta-analysis by Pronin (2007), who argued that partisan bias is a bipartisan problem, and that people may identify bias in others better than they see it in themselves.

Slywester & Purver (2015) identified differences between Democrat and Republican followers in discussed topics and personality characteristics. Republicans focused more on religion, nationalism, government and foundation in the Moral Foundations Theory. These results suggest that language used on Twitter reflects individual differences between liberals and conservatives.

Colleoni et al. (2014) used machine learning and social network analysis to classify users as Democrats or as Republicans based on the political content shared. In general, Democrats exhibited higher levels of political homophily, such that they tended to seek out or be attracted to those who are similar to themselves. However, Republicans who followed official Republican accounts (i.e., @realDonaldTrump) exhibit higher levels of homophily than Democrats. A similar study by Halberstam & Knight (2016) split a sample of Twitter users into conservatives and liberals based upon the political party of candidates most followed by the user. They also found evidence of homophily, with conservatives interacting more with other conservatives and liberals interacting more with liberals. Another study by Conover et al. (2011), examined tweets in the six weeks leading up to the 2010 U.S. congressional midterm elections and found that the network of political retweets exhibited a highly segregated partisan structure. There was limited interaction between left- and right-leaning users. These findings are relevant with the spread of misinformation and fake news (e.g., election conspiracies, anti-vax propaganda, etc.). This, combined with a tendency to form echo chambers on social media, hinders the development of a shared sense of reality.

The purpose of this study was twofold: 1) identify the sentiment of tweets during the 2020 presidential election and 2) determine how frequently each of the moral foundations were used within those tweets. We calculated the sentiment of the tweets and the occurrence of the moral foundations using the Linguistic Inquiry and Word Count and the Moral Foundations Dictionary. Based on previous research, we expect to find differences between tweets referencing either candidate. Our analysis will examine the overall sentiment and frequency of the moral foundations, identify changes from before the election to afterward, and provide an overview highlighting any notable weekly trends.

The 2020 presidential election was one of the most contentious and polarizing in United States history, coinciding with a global pandemic and civil protests. The election was widely discussed on social media, (i.e., Twitter, Facebook, etc.). Through this lens, we were able to compare differences in the sentiment in tweets referencing the two major candidates, Donald Trump and Joe Biden, throughout the election cycle. We examined changes over time identified moments where substantial shifts in sentiment corresponded with real-world events. Lastly, we measured how frequently innate moral foundations were expressed on Twitter.

This study provides an exploratory overview of the sentiment of tweets during the 2020 presidential election. This study should not be viewed as the definitive authority on tweets during the election, but rather as one piece of a larger puzzle. There were several practical and technical limitations, but these concerns were outweighed by the benefits of this study. The size and scope of this study provide insight into the lead-up and aftermath of the election. We can see the sentiment of tweets during events that normally occur during an election cycle (i.e., Conventions, Debates, etc.). We also tracked the sentiment of tweets during several unexpected and unprecedented series of events (i.e., having an insurrection, impeachment, and inauguration occur in consecutive weeks).

Yaqub et al. (2017) reported negative sentiment towards both candidate in the 2016 presidential election between Hilary Clinton and Donald Trump. It's possible that we will find in our exploratory analysis that our sample had similar levels of negativity given the polarizing nature of the election. The rhetoric used by the candidates was polarizing, and in some instances, violent and inflammatory. We expect to see the tweets to be reflective of that language. There's also a plethora of research highlighting the differences between liberals and conservatives on the use of moral foundations in messages by politicians and about politicians. We will examine possible differences in the frequency of use of the moral foundations between tweets about Joe Biden and tweets about Donald Trump. Since we do not have demographic information on the Twitter users, we cannot fully predict the direction of the differences. However, we can predict the dominance of specific foundations due to the topics of interest. The current study will analyze tweets about two presidential candidates. Therefore, we will if there was high frequency of words concerning the Authority foundation. We will also determine if the Fairness foundation was frequently used, because the tweets concerned an election surrounded by allegations of cheating and fraud. Lastly, we will explore whether the Sanctity foundation was commonly used because of the COVID-19 pandemic that occurred during the election cycle. The frequency of each moral foundation should change over time to reflect current events in the news.

METHOD

Collection of Tweets

Twitter is a social media platform for networking and microblogging where users post and interact with messages known as "tweets". Users with accounts can post, like, and share (known as “retweeting”) tweets, while those without an account can only view the tweets. Users access Twitter through its website (https://twitter.com) or an application for mobile devices (e.g., iPhone, iPad, Android phones, etc.).

Approximately 60,000,000 publicly available tweets were collected from the Twitter Application Program Interface (API), using the R package *rtweet* (R Core Team, 2020; Kearney, 2019). Access to the Twitter API required a Twitter developer account.

Figure 1. Example Tweet by Joe Biden

Note. Example of a tweet by President Joe Biden on the official White House Twitter account (Biden, 2021).

The “search_tweets” function in the *rtweet* package was used to access the Twitter API to identify recent tweets that contained the search queries of “Donald Trump”,

"Trump", "Joe Biden" and "Biden". Batches of up to 18,000 tweets were continuously collected in alternating fifteen-minute intervals for each candidate. This amount was selected because it is the maximum number of tweets that a third-party individual can collect before the Twitter API has to reset. The information collected for each tweet included the text of the tweet, the date the tweet was posted, the number of retweets, etc. All data were deidentified before analysis. No personally identifiable information was reported.

Tweets were collected over the course of a seven-month time frame, ranging from August 1st, 2020, to February 5th, 2021. This timeframe was selected to allow for an equivalent number of weeks before and after the general election on November 3rd, 2020 and would also encompass events including the Republican and Democratic conventions, the debates, and the inauguration. A timeline of major events that occurred during data collection can found in the Appendix.

Preprocessing Tweets

The dataset was cleaned to remove references to other individuals that share a similar name to the candidates (e.g., Hunter Biden, Ivanka Trump, etc.). Identifying information (e.g., screen name and user ID) was removed from each tweet. The tweets were aggregated at a weekly level with a week defined as Saturday through Friday (e.g., Week 1 included Saturday, August 1st through Friday, August 6th).

Linguistic Analysis of Tweets

LIWC2015 is the latest in a series of text analysis programs developed by Pennebaker et al. (2015), which uses the Linguistic Inquiry Word Count to determine the frequency of psychological and linguistic concepts in a body of text. LIWC2015 has two primary components: the dictionary and the processing module. The LIWC2015 dictionary contains a list of words related to various psychological and linguistic concepts (e.g., affective processes, parts of speech, etc.). Scores for each category are calculated as the percentage of words in a text that fall in a specific category. A complete list of categories calculated by LIWC2015 can be found in the Appendix. We focused on the categories of Positive Emotion, Negative Emotion, Anger, Anxiety, and Sadness. The latter three categories are subsets of the Negative Emotion category and are the only emotion categories defined in LIWC2015. We also examined the summary variable Tone to identify overall changes in sentiment. LIWC2015 includes the ability to import additional dictionaries. To this end, we analyzed the tweets twice, once using the default LIWC2015 dictionary and again with the Moral Foundations Dictionary 2.0 (MFD 2.0; Frimer et al., 2019).

The processing module of LIWC2015 first reads written or transcribed verbal texts stored in digital formats (such as CSV files). Then, the program identifies all the words in the text and then calculates the percentage of words that match with each of the dictionary categories. To demonstrate, first refer to President Biden's tweet (See Figure 1). The text of the tweet reads "*There is no time to waste when it comes to tackling the crises we face. That's why today, I am heading to the Oval Office to get right to work delivering bold action and immediate relief for American families.*" LIWC2015 identifies a total of 39 words in the tweet. According to the program, there is one word in the First-person Singular Pronoun category (i.e., I) and two Positive Emotion category (i.e., bold, relief). In the example tweet, LIWC2015 outputs a score of 2.6 (equivalent to 2.6% of the words in the tweet) for the First-person Singular Pronoun category and 5.1 (equivalent to 5.1%) for the Positive Emotion category.

Traditional LIWC Categories

The default LIWC2015 dictionary is composed of approximately 6,400 words, word stems, and selected emoticons. LIWC2015 outputs over 90 categories for each tweet, including overall affective processes (positive or negative), parts of speech (pronouns, verbs, etc.) and emotions (anxiety, anger, sadness, etc.). Each word in the dictionary corresponds to one or more word categories. For example, the word "Cried" is part of five word categories: Overall Affect, Negative Emotion, Sadness, Verb, and Past Focus. If the word "Cried" is found in a tweet, the score would be counted for each of the categories. Some LIWC2015 categories overlap with others. For example, all Sadness words are also categorized as Negative Emotion and Overall Affect words. The Overall Affect category includes Negative Emotion and Positive Emotion words. The Negative Emotion category includes all the Anger, Anxiety, and Sadness words.

LIWC Summary Variables

LIWC2015 also includes four summary variables that have been converted to 100-point scales where 0 = very low along the dimension and 100 = very high. The summary variables are: Analytic (analytical or formal thinking), Clout (authoritative or confident writing). Authenticity (personability and honesty), and Emotional Tone (combination of the positive emotion and negative emotion dimensions). We focused on the Emotional Tone variable, which is scored such that higher numbers (> 50) are more positive and lower numbers are more negative.

Moral Foundations Dictionary

The Moral Foundations Dictionary 2.0 (MFD 2.0) is a linguistic analysis dictionary that identifies the usage of five innate, moral foundations in text. The MFD 2.0 (Frimer et al., 2009) is derived from the Moral Foundation Theory (Haidt, 2003; Haidt & Graham, 2007) and the Moral Foundations Dictionary (Graham et al., 2009), which defined the five moral foundations as Care (sympathy, compassion, and nurturance), Fairness/cheating (notions of equality and justice), Loyalty ("us vs. them" mentality), Authority (concerns about traditions and maintaining social order), and Sanctity (moral disgust). The MFD 2.0 contains a list of words related to one or more of the moral foundations such as "riots", "justice", and "virus," which correspond respectively to the Authority, Fairness, and Sanctity foundations. The dictionary contains an average of 210 words per category. According to Frimer et al. (2019), the MFD 2.0 has higher validity than Graham et al.'s (2009) dictionary.

Table 1. Moral Foundations Dictionary 2.0 Categories

Moral Foundation	**Description**	**Virtue Examples**	**Vice Examples**
Care	Caring for others and protecting them from harm.	Love, Help, Hospitalized	Bully, Killed, Suffered
Fairness	Showing equality in one's actions and judgments towards others.	Honest, Justice, Trust	Fraud, Lying, Racist
Loyalty	Having allegiance to one's ingroup, family, and community.	Family, Country, Patriot	Enemy, Traitor, Unpatriotic
Authority	Having respect for tradition and legitimate authority.	Dictator, Protect, President	Disrespect, Overthrow, Riot
Sanctity	Avoiding morally or physically disgusting things, actions, etc.	Faith, Prayer, Righteous	Corruption, Pandemic, Virus

To calculate the moral foundations expressed in tweets, the MFD 2.0 was first imported into LIWC2015 as a user-defined dictionary. Next, the processing module identified all the words in a tweet and calculated the percentage of words that matched each of the moral foundation categories. The MFD 2.0 categorizes the words in each tweet into one of the five foundations each with a positive (virtue) and negative (vice) component for a total of ten categories (Care-Virtue, Care-Vice, Fairness-Virtue, etc.). The outputs were saved as CSV files and later merged into a dataset with the LIWC outputs. Lastly, the merged dataset was imported into R for statistical analysis.

When we conducted our analyses, the tweets were separated into groups based on the candidate referenced within the tweet. In this paper, tweets that mentioned Joe Biden will be referred to as “B-Tweets”, while tweets that referenced Donald Trump will be called “T-tweets”. Any tweets that referenced both candidates were excluded from analysis.

Modified Moral Foundations

The Moral Foundations Dictionary is limited by the list of words used for each foundation. The dominance of a particular foundation may be due to the topic of focus within the text, rather than a natural use of the foundations. To test this, we modified two of the moral foundation categories: Authority (Virtue) and Sanctity (Vice). We reran the Moral Foundations Dictionary, this time excluding the words “President”, “Presidential”, and “Vice President” from the Authority (Virtue) foundation. We removed these words because within the context of the election, people may have been referring to the candidates by their titles (e.g., “President Trump”, “Vice President Biden”, etc.) which would then overinflate the use of the moral foundation. We also modified the Sanctity

(Vice) foundation by adding the words “covid”, “COVID-19”, and “coronavirus”. The default dictionary lacked the necessary vocabulary to fully represent the discussions about the COVID-19 pandemic.

RESULTS

We analyzed the sentiment and moral content of the tweets. First, we calculated the overall sentiment of the tweets and frequency of each moral foundation. Then we identified differences in tweets referencing either candidate. Next, we noted any changes from before the election to after the election. Lastly, we identified any temporal changes in tweets about either candidate.

Comparison to Prior Research

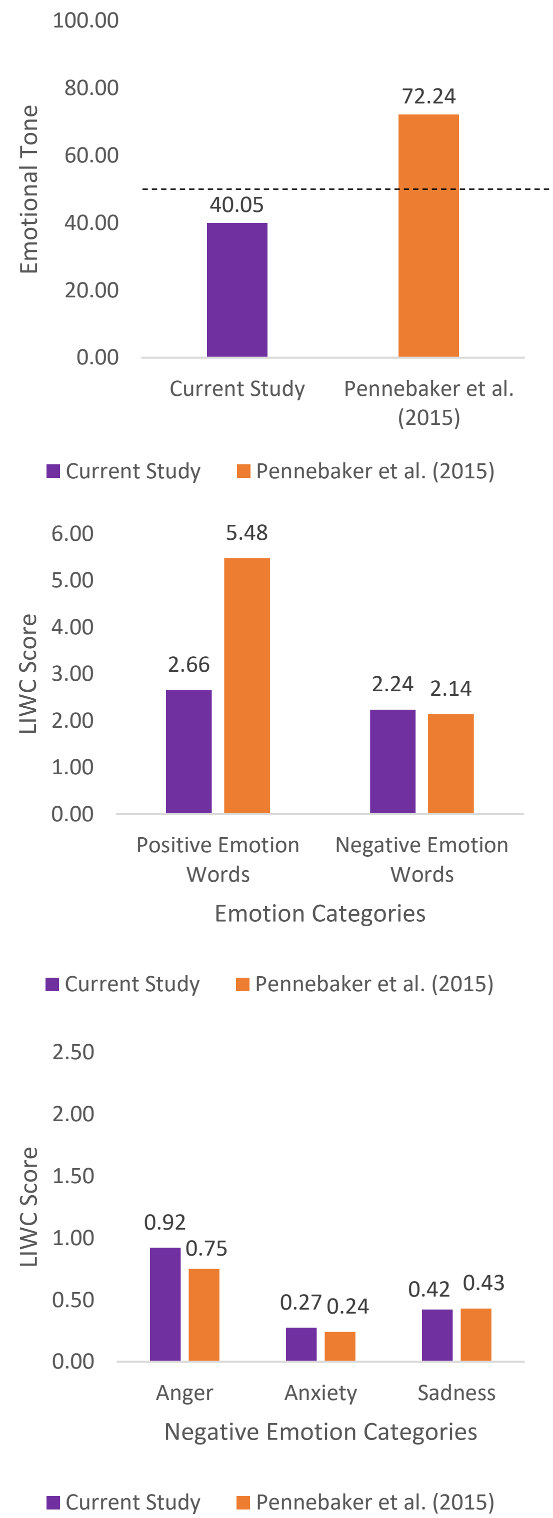
First, we compared the mean LIWC scores from our sample to those reported by Pennebaker et al. (2015), who analyze 35,269 tweets collected from publicly accessible profiles. This comparison demonstrates how our data compares to a general sample of tweets. The average LIWC scores (i.e., the proportion of words in the tweet that fell into specific categories) for positive and negative emotion words were similar between the two samples (See Figure 2). However, we observed several key differences. Our sample had a more negative emotional tone overall. Our sample and Pennebaker et al.'s sample used negative emotion words at a similar rate, but our sample used positive emotion words half as frequently. Anger words were used more frequently in our sample, while the frequency of Anxiety and Sadness words were similar between samples (Figure 2).

In our sample, 3.30% of words used in B-Tweets (tweets that referenced Joe Biden) and 3.65% of words used in T-Tweets (tweets that referenced Donald Trump) matched words found in the Moral Foundations Dictionary 2.0. This is higher than the frequency of 2.2% reported by Lewis (2019), which examined the use of moral foundations in candidate's statements during the 2016 presidential primary debates. However, this difference may be the result of analyzing tweets compared to other forms of media. For example, Pennebaker et al. (2015) reported that positive emotion words and negative emotion words were used more frequently in tweets than natural speech, blogs, and New York Times articles.

Differences Between Candidates – LIWC Scores

We conducted a series of Welch t-tests to evaluate the differences between B-Tweets and T-Tweets. We followed a similar analysis procedure as prior studies that examined LIWC scores and the moral foundations in tweets (Grover et al., 2019; Doğruyola et al., 2019). All the p-values

Figure 1. Comparisons to Pennebaker et al. (2015)



Note. Comparisons of the mean LIWC scores in the current study to Pennebaker et al. (2015).

in our analyses are statistically significant due to the size of sample. To maintain scientific transparency, the results of the t-tests are provided in the Appendix.

We will describe our results with different nomenclature based on the effect size. For example, we will note a difference between conditions if Cohen's d is greater 0.1. If Cohen's d is between 0.05 and 0.1, we will classify the relationship as a small difference. If Cohen's d is less than 0.05, then we will describe the relationship as having no notable differences.

We first examined Emotional Tone, a summary variable from LIWC2015. Emotional Tone converts the frequency of positive and negative emotion words into a composite score on a 100-point scale. Scores above 50 have a positive tone and scores below 50 have a negative tone. The mean score for Emotional Tone for all the tweets in our sample was (M =40.05, SD = 35.96), which indicates an overall negative tone throughout the election cycle. We a difference in Emotional Tone, with T-Tweets (M = 37.78, SD = 35.72) having a more negative average Emotional Tone than B-Tweets (M = 42.43, SD = 36.05).

Next, we examined the LIWC categories of Positive Emotions and Negative Emotions, which are calculated as the proportion of words in a tweet that use the specific emotion. Overall, Negative Emotions were used more frequently than Positive Emotions. There was no notable difference in the frequency of Positive Emotion words between B-Tweets (M = 2.71, SD = 3.65 and T-Tweets (M = 2.61, SD = 3.58). In contrast, T-Tweets used more Negative Emotions words (M = 2.55, SD = 3.60) than B-Tweets (M = 1.91, SD = 3.09).

We further examined the negative emotion words by breaking them into three categories: Anger, Anxiety, and Sadness. Anger was the most frequently used emotion (M = 0.92, SD = 2.23), followed by Sadness (M = 0.42, SD = 1.46), and Anxiety (M = 0.27, SD = 1.08). All three negative emotion categories were used more frequently in T-Tweets than B-Tweets (See Table 2). T-Tweets used Anger words most frequently (M = 1.06, SD = 2.38), followed by Sadness (M = 0.50, SD = 1.65), and Anxiety (M = 0.30, SD = 1.14). Similarly, B-Tweets used Anger words most frequently (M = 0.78, SD = 2.05), followed by Sadness (M = 34, SD = 1.23), and Anxiety (M = 0.24, SD = 1.02).

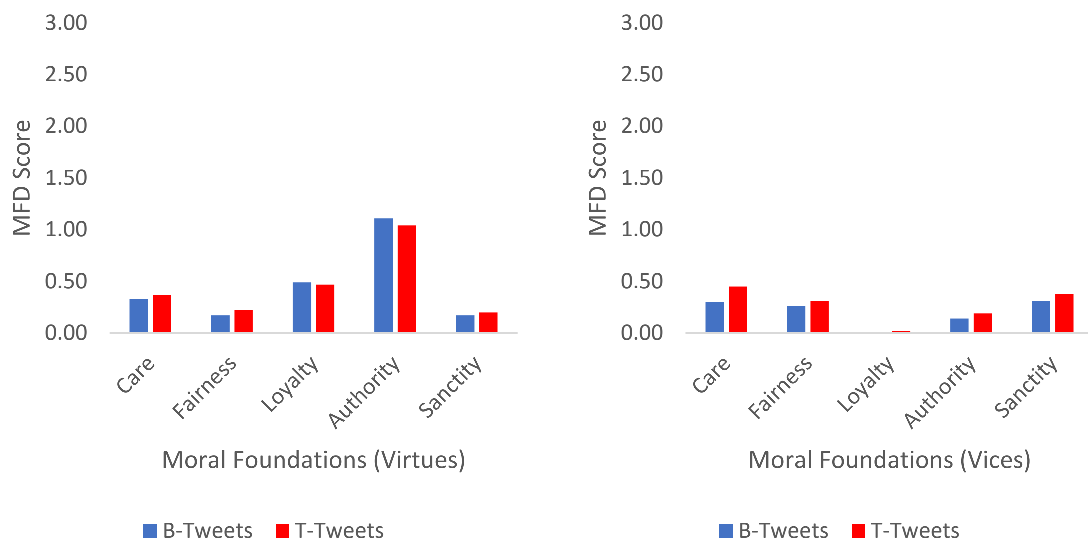
Table 2. Means LIWC Scores for the Negative Emotion Categories in B-Tweets and T-Tweets

	Anger		**Anxiety**		**Sadness**	
	M	*SD*	*M*	*SD*	*M*	*SD*
Biden	0.78	2.05	0.24	1.02	0.34	1.23
Trump	1.06	2.38	0.30	1.14	0.50	1.65
Overall	0.92	2.23	0.27	1.08	0.42	1.46

Moral Foundations

We ran the tweets through LIWC2015 using the Moral Foundations Dictionary 2.0 to calculate scores for the five foundations (each with a virtue and vice component). Refer to Figure 2 for details on each of the moral foundation categories. Figure 3 shows the mean scores for each moral foundation for B-Tweets and T-Tweets. The most frequently used foundation overall was Authority (Virtue) (M = 1.08, SD = 2.28), followed by Loyalty (Virtue) (M = 0.48, SD = 1.41). The top two foundations and bottom two foundations were the same for B-Tweets and T-Tweets. B-Tweets used Authority (Virtue) (M = 1.11, SD = 2.33) and Loyalty (Virtue) (M = .49, SD = 1.43) most frequently. Comparatively, T-Tweets used Authority (Virtue) (M = 1.04, SD = 2.24) and

Figure 3. Means Scores for the Moral Foundations in B-Tweets and T-Tweets



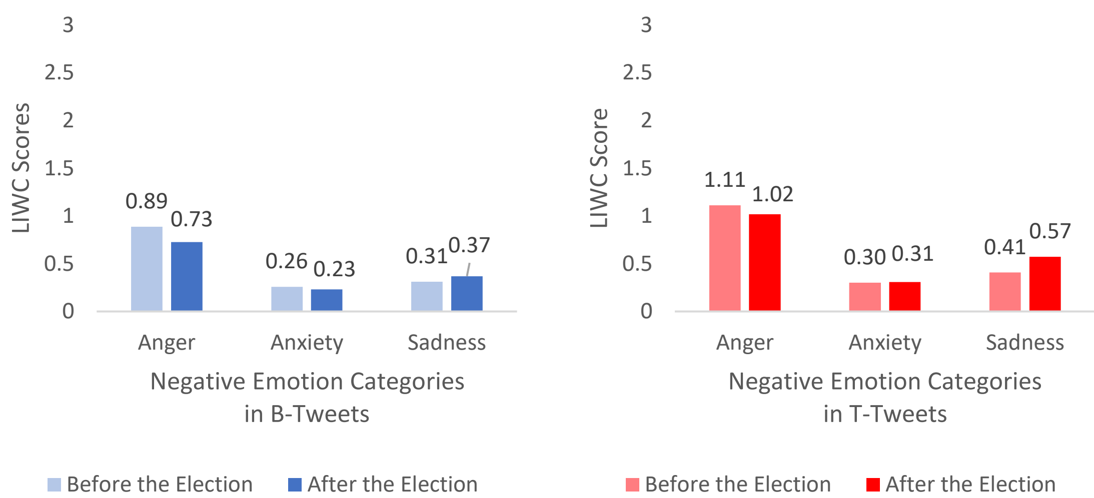
Loyalty (Virtue) (M = .47, SD = 1.40). There were no notable differences between B-Tweets and T-Tweets in the frequency of any foundation.

Both groups had Care as the third most used foundation. However, it was Care as a virtue (M = .33, SD = 1.20) for B-Tweets and Care as a vice for T-Tweets (M = .45, SD = 1.43). Lastly, Loyalty (Vice) was hardly used in either B-Tweets (M = .01, SD = .23) or T-Tweets (M = .02, SD = .29). T-Tweets used all of the vices more than B-Tweets. A full breakdown of the mean frequencies for each of the moral foundation categories is presented in Table 5 (See Appendix).

Pre-Election Versus Post-Election

We examined the sentiment before and after the election by splitting the tweets into two equally sized groups. Tweets from Weeks 1-13 were aggregated into one group and tweets from Weeks 15-27 were in the second group. In this section, we excluded tweets from Week 14, the week of the election.

Figure 4. Mean LIWC Scores Before and After the Election in B-Tweets and T-Tweets

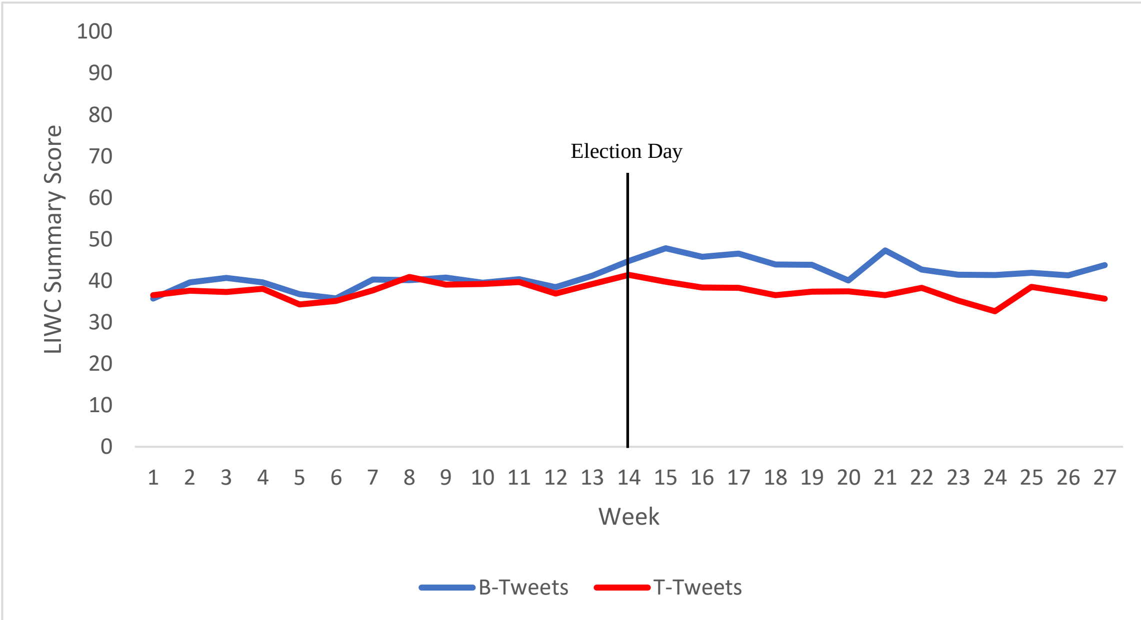


We then compared the LIWC scores for each group from before to after the election. B-Tweets saw a difference in Emotional Tone, with it increasing from M = 39.26 (SD = 35.27) to M = 44.04 (SD = 36.40). In contrast, T-Tweets did not have a noticeable change from before the election (M = 38.14, SD = 35.87) to after the election (M = 37.19, SD = 35.49). Anger decreased more in B-Tweets than T-Tweets Anxiety increased slightly for both groups, and Sadness increased more in T-Tweets than B-Tweets (See Figure 4),.

Weekly Trends

We examined weekly changes for each variable to identify possible connections with real-world events. Due to the exploratory nature of this study, we provided a descriptive overview of these findings. This will establish a guide for future research to further explore possible relationships between real-world events and changes in the sentiment or moral content of tweets. The LIWC categories highlighted the differences between the groups of tweets.

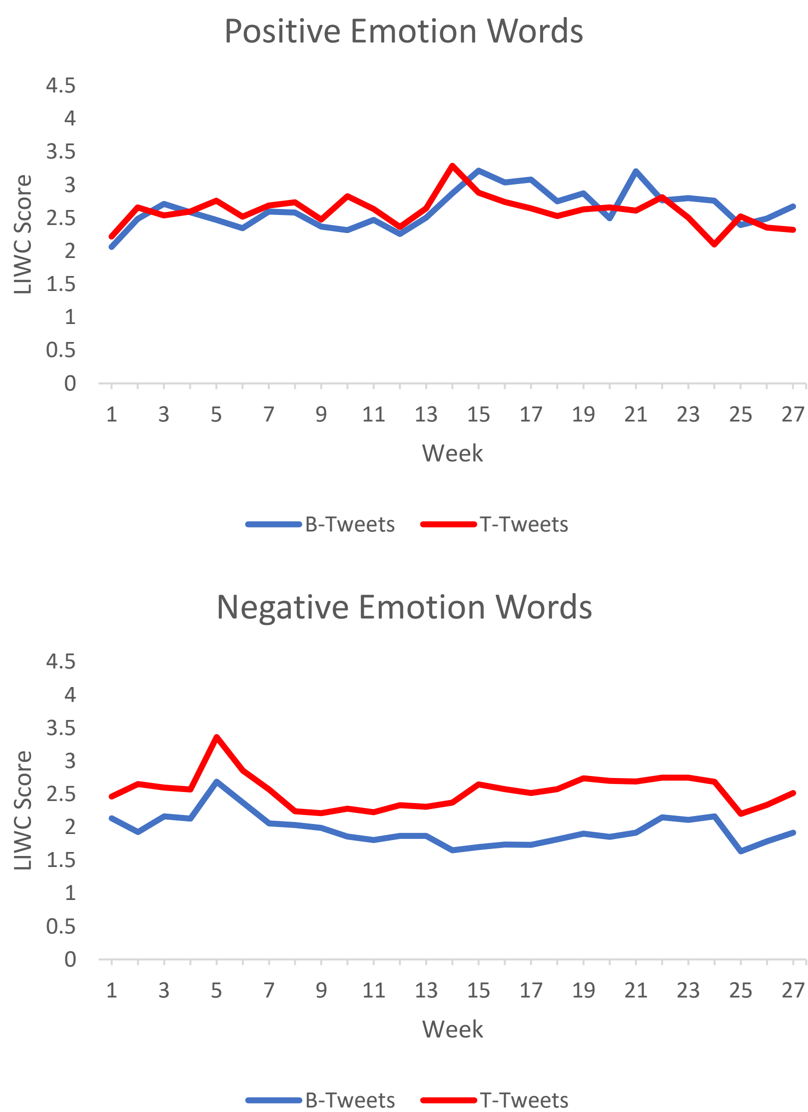
Figure 5. Weekly Changes in Emotional Tone



Note. Emotional tone uses a different scale than the normal LIWC scores. The scores range from 0 to 100. Scores above 50 have a positive sentiment and scores less than 50 are considered negative.

By examining the weekly means for emotional tone, we identified several points where shifts in the emotional tone that coincided with real-world events (See Figure 5). One of the major shifts occurred around the election (Week 14). There are distinct patterns of behavior before and after the election. Before the election, tweets about either candidate averaged similar levels of emotional tone. However, in the weeks immediately prior to the election (Weeks 12 and 13), the emotional tone scores begin to diverge. After that point, B-Tweets averaged a less negative emotional tone than T-Tweets. Despite this shift, neither B-Tweets nor T-Tweets averaged a positive emotional tone in any week during data collection. We can further analyze the shift in emotional tone by breaking down the frequencies of positive and negative emotion words.

Figure 6. Weekly Trends for Positive and Negative Emotion Words



The frequency of positive emotion words showed no notable differences between B-Tweets and T-Tweets. While the overall trend followed a similar pattern for both groups, we can see some small differences when comparing before and after the election (See Figure 6). In the beginning, T-Tweets used positive words slightly more frequently than B-Tweets. However, following a spike in the frequency for both groups that occurred within the weeks surrounding the election, B-Tweets used more positive words on average.

This is contrasted by the negative emotion word category. Throughout the entire period of data collection, T-Tweets used more negative words than B-Tweets. There was a large increase in the use of negative words following the Republican National Convention (Week 4) and a decrease during the week of Joe Biden's inauguration (Week 25). Notably, the frequencies diverged around the time of the election, with T-Tweets using more negative words and B-Tweets using less.

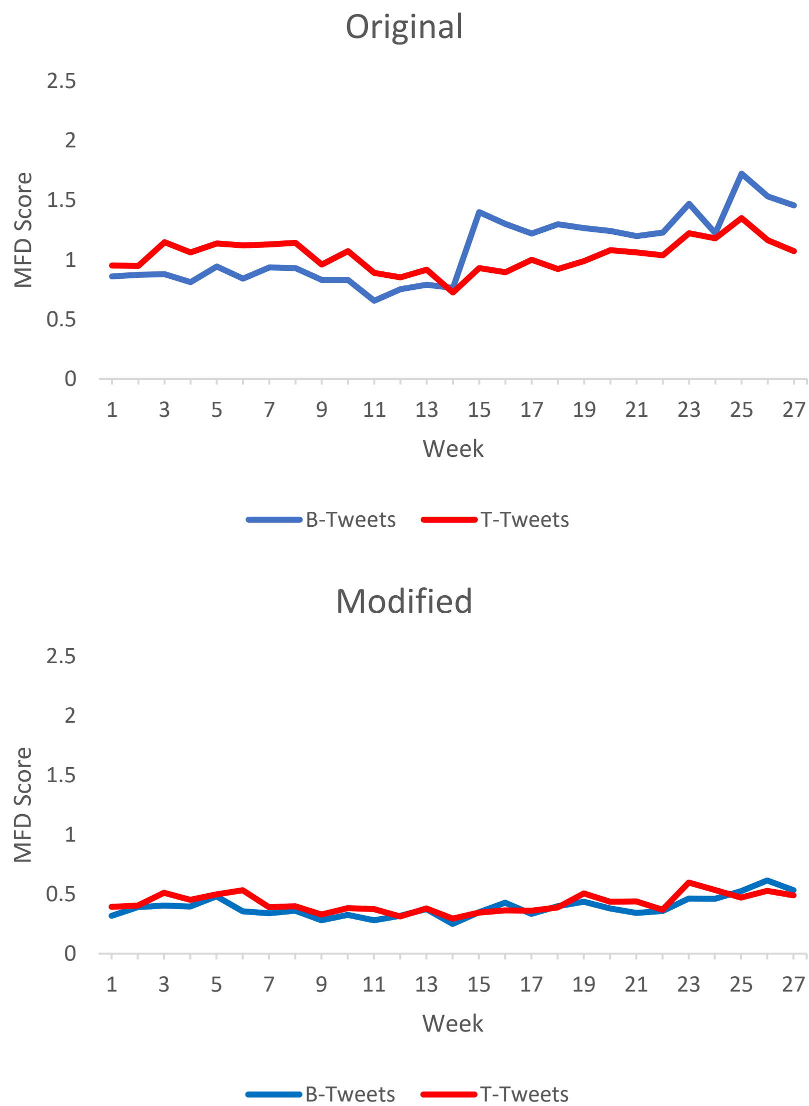
The Authority, Loyalty, and Sanctity foundations showed differences between their virtue and vice components (See Table 5 in the Appendix). The Authority foundation was more frequently used in B-Tweets after the election, while T-Tweets did not show much of a change. The Loyalty foundation was hardly used at all as a vice in either group. The Sanctity foundation was used more as a vice than a virtue.

Modified Moral Foundations

The effectiveness of the Moral Foundations Dictionary is restricted by the list of words used for each foundation. One foundation may appear frequently simply because it is relevant to the topic of interest, while another might appear sparingly because the text used proper nouns instead of the generic terms listed in the dictionary. We modified two of the moral foundation categories: Authority (Virtue) and Sanctity (Vice).

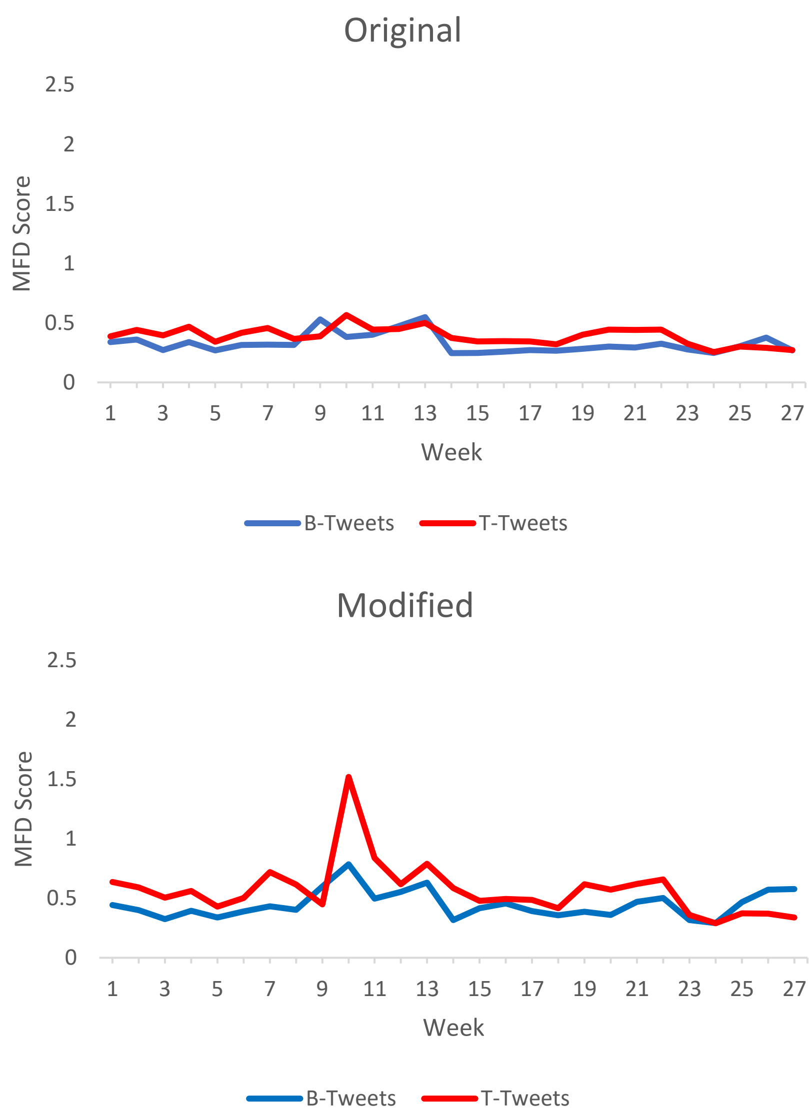
Our initial analysis of the Authority (Virtue) foundation showed an increase in the frequency from before the election to afterwards in B-Tweets. When we examined the weekly trends, we noted a large spike in the frequency of use that occurred around the election. We hypothesized that this spike was the result of the use of the word "president" and other similar words. Joe Biden won the 2020 Presidential Election and was thereby referred to as "President Biden" or "President-Elect Biden". To test our hypothesis, we reran the Moral Foundations Dictionary, this time excluding the words "President", "Presidential", and "Vice President". The original analysis showed a difference in the pattern of change over time between B-Tweets and T-Tweets. Most notably was the sharp increase in the frequency of the foundation that occurred around the election in Week 14 (See Figure 7). When we remove those words, the two groups follow a more similar trend.

Figure 7. Authority (Virtue) Foundation – Comparison of Original and Modified Versions



In contrast, our initial analysis of the Sanctity (Vice) foundation showed that the frequency decreased in both B-Tweets and T-tweets. However, users on Twitter may have been discussing this foundation more frequently than could be identified by the Moral Foundations Dictionary. The Sanctity (Vice) foundation includes words such as “pandemic” and “virus”. The 2020 Presidential Election occurred amidst the COVID-19 pandemic. Users discussing the pandemic may have referred to it as “covid” or “COVID-19”, neither of which appear in the dictionary. As a result, the Sanctity (Vice) foundation may be underrepresented within our dataset. To account for this, we reran the Moral Foundations Dictionary, this time including the words "covid", "COVID-19", and "coronavirus". As expected, adding the additional terms to the dictionary increased the frequency of the Sanctity foundation. Notably, there was a spike in the frequency in the use of words in this foundation that occurred around Week 10, which corresponds to when Donald Trump was hospitalized with COVID-19 (See Figure 8).

Figure 8. Sanctity (Vice) Foundation – Comparison of Original and Modified Versions



DISCUSSION

Based on previous research, we expected to find differences in the sentiment of tweets about Donald Trump and Joe Biden. The outcome of the election was polarizing, with little cross-party support for either candidate. This, combined with evidence on political homophily, suggested that we would see differences between the candidates in the sentiment of tweets and the expression of the moral foundations. We noted several instances where changes in the sentiment and frequency of the moral foundations coincided with real world events. We identified differences between the candidates on the emotions of anger, anxiety, and sadness.

This was a highly contentious election with record turnouts, with the highest and second highest vote totals for an individual candidate. During the election cycle, there was civil unrest (e.g., BLM protests), the COVID-19 pandemic, and an abundance of conspiracy theories and misinformation about voter fraud. The current study lays the foundation for future research by providing an overview of the sentiment of tweets about the two candidates during the election cycle. The information we gathered is beneficial both for the scale of the dataset and the timeframe in which the data was collected. There were several unanticipated events that occurred during data collection; the hospitalization of Donald Trump with COVID-19, the insurrection at the U.S. Capitol on January 6th to disrupt the certification of the electoral votes, and the first time that a president was impeached twice. The timeframe of this study provides insight into the sentiment of political discussions during 2020 election cycle.

Our analysis showed that, on average, there was an overall negative sentiment in the tweets. Each candidate individually had a negative average sentiment, with tweets referencing Trump having lower emotional tone than tweets referencing Biden. When we compared the frequency of positive and negative emotion words, the differences between the candidates became more apparent. While both candidates had similar proportions of negative emotion words in their tweets, Biden had higher levels of positive emotion words. We can see than that frequency of positive and negative emotion words did not change much in tweets referencing Trump from before the election to afterwards. In contrast, the proportion of positive emotion words increased in tweets referencing Biden after the election, while the amount of negative emotion words decreased. Breaking the data down into weekly averages further demonstrates the differences between the candidates in the frequency of emotion words before and after the election. We can see a divergence in the emotional tone that occurs around Election Day.

The results of our study showed differences in the sentiment and the expression of the moral foundations in tweets about Biden and Trump. We found that tweets about Biden had more positive tone on average than tweets about Trump. However, the tone overall was negative for both candidates. Our findings are supported by Yaqub et al. (2017), who measured the sentiment of tweets during the 2016 Presidential election between Hillary Clinton and Donald Trump and found that the sentiment of tweets was negative for both candidates.

Regarding the moral foundations, we found that Care, Fairness, and Authority were the most frequently used foundations. This is similar to Kaur and Sasahara (2017), who also found that Care was one of the more dominant foundations. We found notable spikes in the frequency of certain foundations that coincided real-world events. Most of the changes occurred around the week of the election. There was a sharp increase in the frequency of tweets using Authority (Virtue) words for tweets referencing Joe Biden and an increase in the frequency of Sadness words in tweets about Donald Trump. The tone of the tweets was similar between the groups until it diverged after the election. The tweets about Biden became more positive and the tweets about Trump became more negative over the course of this study.

Our results showed several spikes in the frequency of the Authority (Vice) foundation that coincided with real-world events. Frimer et al (2013) speculated that participants interpreted the foundation of authority as "authority that I believe to be legitimate". This election was marred by allegations of voter suppression and voter fraud. Donald Trump held a rally on the National Mall on January 6th, 2021, to rile his supporters with false claims that the election was stolen, and he encouraged them to march down the street to prevent Congress from counting the electoral votes. This ultimately culminated in the storming of the U.S. Capitol on January 6th, 2021, temporarily disrupting the certification of the electoral college votes. During that same period, Biden supporters decried Trump's efforts as an illegal grab for power. The spikes we saw in our results may be reflective of the notion of illegitimate authority.

One concern with using social media as a research tool, particularly Twitter, is the lack of consistent demographic information. A limitation of the *rtweet* package is that the data collected is restricted to whatever the creator of a tweet provides, and the account is marked as publicly available. With fully public accounts, *rtweet* is an excellent tool that can provide the geocoordinates where the tweet was posted, the number of friends for the account, etc. This restriction can potentially lead to biases within the sample, as it may exclude users that are more privacy conscious. This may also cause issues if researchers want to compare tweets from different geographic locations.

The demographics of Twitter users do not perfectly reflect the population of voters. Twitter users tend to be younger, higher educated and more politically liberal than the general population (Mellon & Prosser, 2017). According to Auxier and Anderson (2021), 72% of Americans say they ever use social media sites. Nearly 84% of adults ages 18 to 29 say they ever use any social media sites, which is similar to the number of those ages 30 to 49 who use social media (81%). By comparison, a smaller proportion of those ages 50 to 64 (73%) say they use social media sites, while fewer than half of those 65 and older (45%) report using them. We also could not identify the political leaning of users from individual tweets and are therefore unable to control for this factor. Additionally, we couldn't identify whether a user participated in the election, which means that our sample includes people who may have been unable or unwilling to vote.

As expected, Authority was the most frequently used foundation. When we adjusted the dictionary, we saw that the gap in frequency was largely due to the topic of the tweets. Most of the Authority words were "president", "presidential", and "vice president", all of which are words that one would expect to see in discussion about a presidential election. Loyalty was the second most dominant foundation. This is also reasonable to expect within the context of an election, particularly one as polarizing as the 2020 election. The foundation was used almost exclusively as a virtue. Loyalty (Vice) was the least used foundation.

Sanctity (Vice) was underrepresented in our initial analysis. We anticipated that it would be among the more dominant foundations due to the COVID-19 pandemic. However, the narrow list of words in the Moral Foundations Dictionary 2.0 did not fully capture the topic. When we expanded the dictionary, we saw that the Sanctity (Vice) foundation was more prevalent than initially reported. The frequency of the foundation nearly doubled when we accounted for words commonly used to discussion the pandemic.

Technical Limitations

The "search_tweets" function operates by identifying tweets that contain the words specified in the search query. For example, we used the search query "Biden" to identify tweets that were likely to mention Joe Biden. This was done to strike a balance between having more focused data and having a sizable sample. Tweets are limited to 240 characters, which means that users are likely to use a shortened version of the candidates' names rather than their full names (i.e., Joe Biden being referred to as "Biden" or Donald Trump being referred to as "Trump"). This also does not capture when the candidates were called by nicknames (i.e., "Sleepy Joe", "Donald Drumpf", etc.). The search query was broad enough to identify as many tweets as possible, which could be cleaned later to remove misidentified individuals (e.g., Hunter Biden, Ivanka Trump, etc.). No videos or photos linked within the tweets were analyzed in this study, we focused solely on the text of the tweets.

LIWC and MFD do not take the context of the sentence into account. "I broke the law." is only scored in the Fairness (Virtue) category, even though the statement would imply a violation of authority. LIWC and MFD were selected for their overall validity and reliability as text analysis techniques, but their effectiveness is limited with short texts such as tweets. The frequency of certain foundations may be the result of subject of the tweets. Authority (Virtue) contains words that one would expect to see during an election (e.g., “Leader”, “President”, etc.).

Another concern is that the Moral Foundations Dictionary (and the MFD 2.0) only contain a small proportion of the words in the English language. Therefore, it is unreasonable to expect a high proportion of morality words to appear in any given body of text. In Lewis (2019), none of the moral foundations represented even 1% of the words in statements made by candidates in the 2015-2016 presidential primary. All the foundations together only totaled approximately 2.2% of the words (similar to the percentage found in Graham et al. 2009 and Lipsitz, 2018). They argued that the low averages obscure that fact that while many statements have zero words representing any foundation, some candidate statements use certain foundations in a much more pronounced way. We encountered similar issues with our data; many of the tweets used none of the moral foundations.

We were not able to collect data for one day due to a technical malfunction. However, bias resulting from this minimal lack of data is likely to be negligible. Due to the large sample size in our study, traditional methods of statistics cannot easily compare these groups because even the smallest effect is statistically significant. To be thorough, we reported the results of t-tests comparing the groups of tweets, but we recommend caution when drawing conclusions.

Comparison to Prior Studies

Hanschmidt and Kersting (2021) identified several issues limiting the generalizability of analyzing social media data. They collected tweets through the Twitter API which does not provide all tweets matching the search query but instead samples a collection of tweets. The researchers identified tweets using hashtag-based search queries to increase the relevance of tweets obtained via Twitter's API.

Bruns and Stieglitz (2014) argue that samples of tweets collected by hashtag-based searches may not be representative of discourse related to a specific topic. We encountered similar issues when using LIWC2015 and the Moral Foundations Dictionary 2.0. We were restricted to the list of words presented in each dictionary. This was particularly a problem when analyzing the Sanctity (Vice) foundation. The COVID-19 pandemic was ongoing throughout the entire period of data collection. As such, it was a frequently discussed topic. When we modified the dictionary by adding three words, we saw that the frequency nearly doubled.

Hanschmidt and Kersting (2021) also noted that the word count approach they used to estimate the prevalence of emotions could not identify negations and other contextual factors that changes a tweet's meaning (e.g., sarcasm). However, since they calculated rates of emotions based only on aggregated tweets (i.e., at the overall or weekly levels), the misclassification of a single tweet was unlikely to influence the overall results. We followed the same procedure when analyzing our data by reporting the aggregated results at the overall and weekly levels.

Our data collection procedures had similar limitations. We used a keyword-based search to identify tweets that referenced either candidate. The R package *rtweet* can only collect a sample of 18,00 tweets every 15 minutes, rather than collecting every tweet using those keywords.

Sánchez-Rada & Iglesias (2019) argued that information shared on social networks is not isolated. The meaning of a particular piece of content (e.g., a tweet, a Facebook post, etc.) may only be understood when its context is taken into consideration. This context includes visible information such as previous content that belongs to the same conversation, previous interactions between users, or people that interacted with the content (e.g., by liking it). For instance, some demographic factors such as age and gender correlate with sentiment and vocabulary.

Implications

Our study has major implications for the use of social media data in psychological and political research, as well as providing a broader understanding of individuals' reactions during this election. We built upon previous research into big data and linguistic analysis. While we did not develop new tools or procedures, we instead demonstrated how previously established techniques can be integrated into newer, revolutionary approaches.

Social media analytics is still a young field. The current popular platforms are less than two decades old. There simply has not been enough time to fully understand its effects. The concept of the Internet of Things seemed ludicrous several decades ago. How we communicate as a society has evolved. Through the Internet, we now have access to a virtually limitless source of information. We can talk to anyone in the world at any time instantaneously. We are no longer restricted to social interactions within our local communities, we are transitioning into a global society.

Social media came into prominence during the early 2000s with popular sites such as Myspace (2003), Facebook (2004), and Twitter (2006) among the most popular. The latter two are still widely used today. Despite their integration into modern society, these platforms are relatively young. Twitter, if it were a person, would still not be eligible to get a driver's license. Myspace, once the most visited website in the U.S., just recently reached the age of a legal adult. We are currently living in an era of exponential technological growth. According to a Pew Survey, 84% of Americans use some form of social media. This has shifted how people communicate. Whereas people were once restricted to communicating with those who were geographically close, we now have the capability to send and receive information instantaneously to anyone in the world at any time. This has massive implications within our society that have yet to be fully studied. Most of the population has yet to live in a time where social media has always existed. As such, it is important that we chronicle changes within our society during this transitionary period.

This study provides a snapshot into this period in history. Amid this era of technological progression, the United States has also experienced a period of political and social change. Since the launch of Twitter, the U.S. has elected its first African American president, legalized homosexual marriage, elected the first female Vice President, etc. Each of these milestones would have been radical to consider a decade or two earlier. The country has experienced periods of turmoil as well. The 2020 Presidential election occurred during movements of civil unrest, a global pandemic, and widespread accusations of election fraud. We can provide insight into the thoughts of average individuals and provide a foundation for future researchers and historians.

We chose LIWC because it is a widely used and repeatedly validated measure. This allows us to compare our results more easily to prior research. It also establishes a baseline we can use to test against newer analytical techniques. Traditional statistical analysis is unsuited for big data analytics. Any sufficiently large sample will produce statistically significant results, with little regard to the effect size. This hinders the predictive potential of the results. Machine learning is a possible solution to this problem. Previous studies have demonstrated the predictive power of machine learning with analyzing the sentiment of tweets (Hasan et al., 2018; Severyn & Moschitti, 2015; etc.) and the moral content of tweets (Garten et al., 2016; Xie et al., 2019; Hoover et al., 2020).

This study provided an exploratory overview about tweets from the 2020 Presidential election. Our goal was to identify possible differences in the sentiment and moral content of tweets referencing either candidate. Few studies have focused on tweets from presidential elections. Twitter (and social media in general) is still a relatively new technology. There have only been four presidential elections since its launch in 2006. We still do not fully understand how social media influences people during an election. The few studies on this topic have been limited in sample size and the duration of data collection. We addressed these issues by collecting a massive sample of 60-million tweets over a 6-month period. We wanted to establish a foundation for future researchers studying this period in history. Our data provides a snapshot into the reactions of millions of people during the contentious 2020 Presidential election.

We must also consider the ethical concerns inherent with new technology. How much information can be gathered and what can we learn from it? At what point does the data collection become invasive? Previous research has shown some of the potential perils of social media data. Machine learning algorithms can use public social media data to predict personality, sexual orientation, and political ideology. The algorithms are only going to improve over time. How are we, as a society, going to address these issues. We already allow companies to use social media data to generate personalized advertisements, but the same principles can be applied to other fields. Cambridge Analytical infamously used social media data in targeted misinformation campaigns during the 2016 Presidential election and other political campaigns (Confessore, 2018; Confessore & Hakim, 2017).

Careful consideration must be given to the design and implementation of social media research, striking a balance between the vast quantity of available data and the boundaries defined by the experimental design and hypotheses. The study was designed to be an exploratory analysis of the sentiment of tweets during the 2020 presidential election. The goal of this study was to track general trends over time rather than identifying changes in the attitudes of individual users. Research looking into the attitudes of individual users must take extra precautions to account for the prevalence of bots, troll farms, and parity accounts. These concerns were lessened in our study due to the generalized nature of our research objective. Despite these drawbacks, social media remains an important resource for modern psychological research.

The current study demonstrated the benefits of using social media (i.e., Twitter) data as a research tool. We collected approximately sixty million tweets over a seven-month period. It would have been logistically challenging to gather this quantity of data through traditional methods, but as shown in this study, it can be accomplished by a single researcher with a computer. When combined with the availability of open-source software such as R, collecting social media data is an accessible, cost-effective avenue for conducting psychological research.

We used well-established and validated methods for computerized text analysis (the Linguistic Inquiry and Word Count and the Moral Foundations Dictionary 2.0) and provided a comparison to prior research (i.e., Pennebaker et al., 2015). We classified the sentiment and moral content for millions of tweets, and because we collected the tweets daily over a seven-month period, we were able to identify temporal patterns of change. This study was not without its limitations, but it serves as a framework for future research using social media.

APPENDIX

Table 3. Major Events that Occurred During Data Collection

Week (Defined as Saturday to Friday)	Event Description	Event Start Date	Event End Date
Week 1	Data Collection Began on August 1st, 2020	8/1/2020	
Week 3	Democratic National Convention	8/17/2020	8/20/2020
Week 4	Republican National Convention	8/24/2020	8/27/2020
Week 9	1st Debate	9/29/2020	
Week 10	Donald Trump Hospitalized with COVID-19	10/2/2020	10/5/2020
Week 11	2nd Debate (Cancelled - Changed to Individual Town Halls)	10/15/2020	
Week 12	3rd Debate	10/22/2020	
Week 14	Election Day	11/3/2020	
Week 15	Joe Biden Declared the Winner	11/7/2020	
Week 20	The Electoral College Meets to Vote	12/14/2020	
Week 23	Jan. 6th Riot & Certification of Electoral College Results	1/6/2021	
Week 24	Second Impeachment of Donald Trump	1/13/2021	
Week 25	Inauguration of Joe Biden	1/20/2021	
Week 27	Data Collection Ended on February 5th, 2021	2/5/2021	

Table 4. Means for LIWC Categories for B-Tweets and T-Tweets

	Emotional Tone	
	M	*SD*
B-Tweets	42.43	35.72
T-Tweets	37.78	36.05
Overall	40.05	35.96

	Positive Emotions		Negative Emotions	
	M	*SD*	*M*	*SD*
B-Tweets	2.71	3.65	1.91	3.09
T-Tweets	2.61	3.58	2.55	3.60
Overall	2.66	3.61	2.24	3.37

	Anger		Anxiety		Sadness	
	M	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.78	2.05	0.24	1.02	0.34	1.23
T-Tweets	1.06	2.38	0.30	1.14	0.50	1.65
Overall	0.92	2.23	0.27	1.08	0.42	1.46

Table 5. Means for Moral Foundations for B-Tweets and T-Tweets

Virtues

	Care		Fairness		Loyalty		Authority		Sanctity	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.33	1.20	0.17	0.86	0.49	1.43	1.11	2.33	0.17	1.05
T-Tweets	0.37	1.32	0.22	0.97	0.47	1.40	1.04	2.24	0.20	1.17
Overall	0.35	1.26	0.19	0.92	0.48	1.41	1.08	2.28	0.19	1.11

Vices

	Care		Fairness		Loyalty		Authority		Sanctity	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.30	1.10	0.26	1.31	0.01	0.23	0.14	0.79	0.31	1.25
T-Tweets	0.45	1.43	0.31	1.22	0.02	0.29	0.19	0.91	0.38	1.51
Overall	0.38	1.28	0.28	1.26	0.02	0.26	1.08	0.85	0.35	1.39

Table 6. Comparison of the Original and Modified Versions of the Authority (Virtue) Foundation

Overall

	Original		Modified	
	M	*SD*	*M*	*SD*
B-Tweets	1.11	2.33	0.38	1.27
T-Tweets	1.04	2.24	0.43	1.38

Before the Election

	Original		Modified	
	M	*SD*	*M*	*SD*
B-Tweets	0.85	1.94	0.35	1.22
T-Tweets	1.03	2.18	0.37	1.22

After the Election

	Original		Modified	
	M	*SD*	*M*	*SD*
B-Tweets	1.36	2.59	0.43	1.34
T-Tweets	1.07	2.29	0.44	1.41

Table 7. Comparison of the Original and Modified Versions of the Sanctity (Vice) Foundation

Overall

	Original		Modified	
	M	*SD*	*M*	*SD*
B-Tweets	0.31	1.25	0.43	1.48
T-Tweets	0.38	1.51	0.55	1.78

Before the Election

	Original		Modified	
	M	*SD*	*M*	*SD*
B-Tweets	0.37	1.36	0.89	1.55
T-Tweets	0.43	1.56	0.49	2.17

After the Election

	Original		Modified	
	M	*SD*	*M*	*SD*
B-Tweets	0.28	1.19	0.49	1.45
T-Tweets	0.34	1.43	0.48	1.66

Figure 9. Mean Scores for the LIWC Summary Variable Emotional Tone

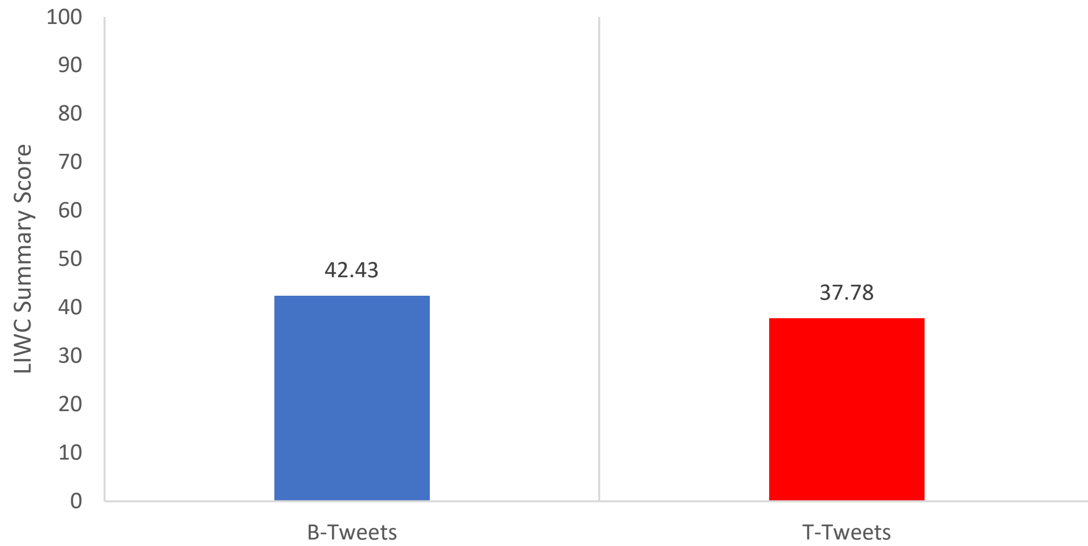


Figure 10. Mean Scores for Positive and Negative Emotions

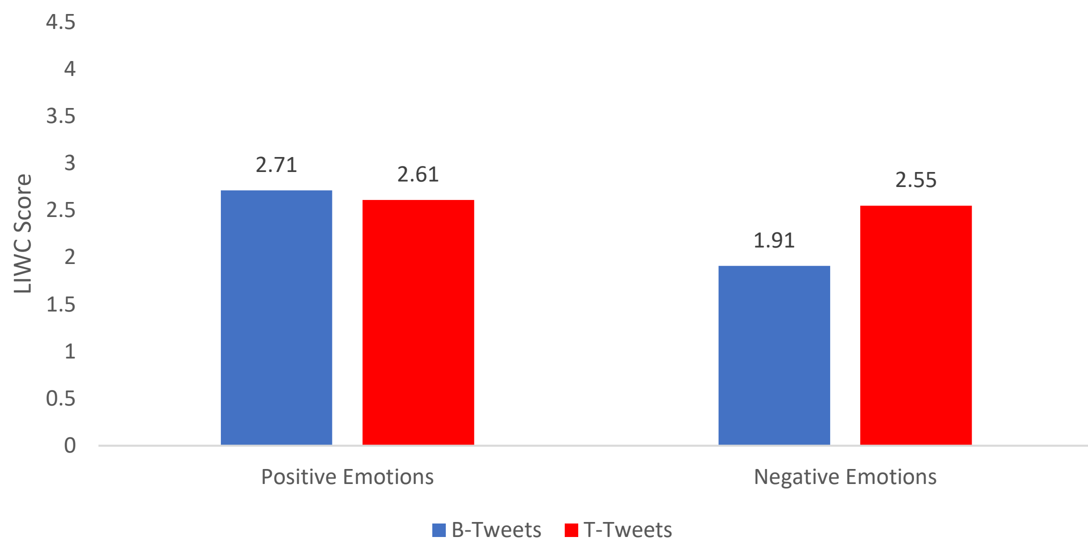


Figure 11. Mean Scores for the Negative Emotions Categories

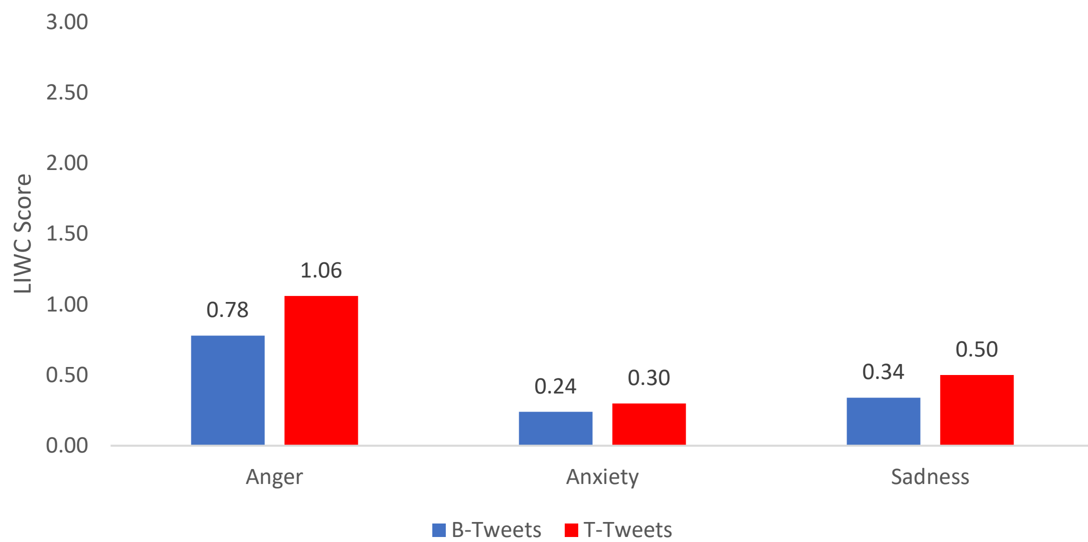


Table 8. Mean Scores for Emotional Tone Split by Group and Time

	Before the Election		After the Election	
	M	*SD*	*M*	*SD*
B-Tweets	39.26	35.27	44.04	36.40
T-Tweets	38.14	35.87	37.19	35.49

Figure 12. Mean Scores for Emotional Tone Split by Group and Time

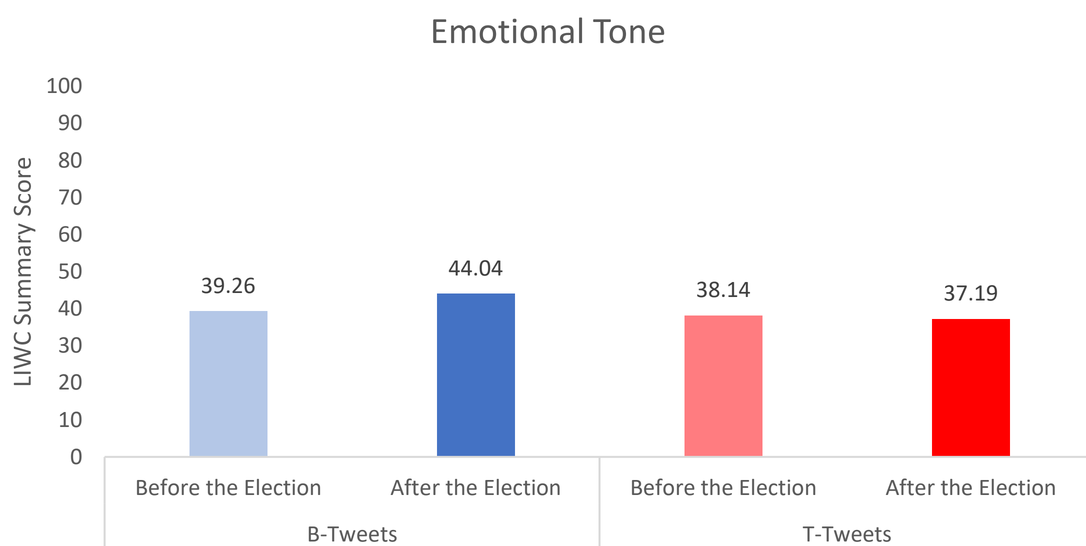


Table 9. Mean LIWC Scores for Positive Emotions and Negative Emotions Split by Group and Time.

Before the Election

	Positive Emotions		Negative Emotions	
	M	*SD*	*M*	*SD*
B-Tweets	2.45	3.45	2.07	3.15
T-Tweets	2.59	3.50	2.51	3.54

After the Election

	Positive Emotions		Negative Emotions	
	M	*SD*	*M*	*SD*
B-Tweets	2.85	3.68	1.86	3.08
T-Tweets	2.58	3.58	2.60	3.62

Table 10. Mean LIWC Scores for Negative Emotion Categories Split by Group and Time.

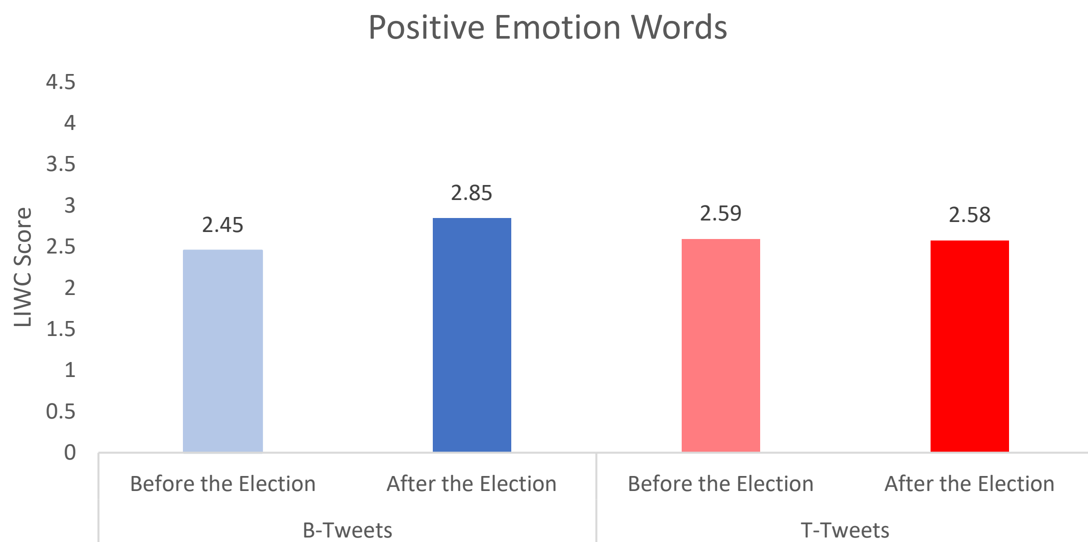
Before the Election

	Anger		Anxiety		Sadness	
	M	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.89	2.06	0.26	1.09	0.31	1.18
T-Tweets	1.11	2.44	0.23	1.13	0.37	1.39

After the Election

	Anger		Anxiety		Sadness	
	M	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.73	2.08	0.30	0.95	0.41	1.28
T-Tweets	1.02	2.31	0.31	1.15	0.57	1.81

Figure 13. Mean Scores for Positive and Negative Emotions Split by Group and Time



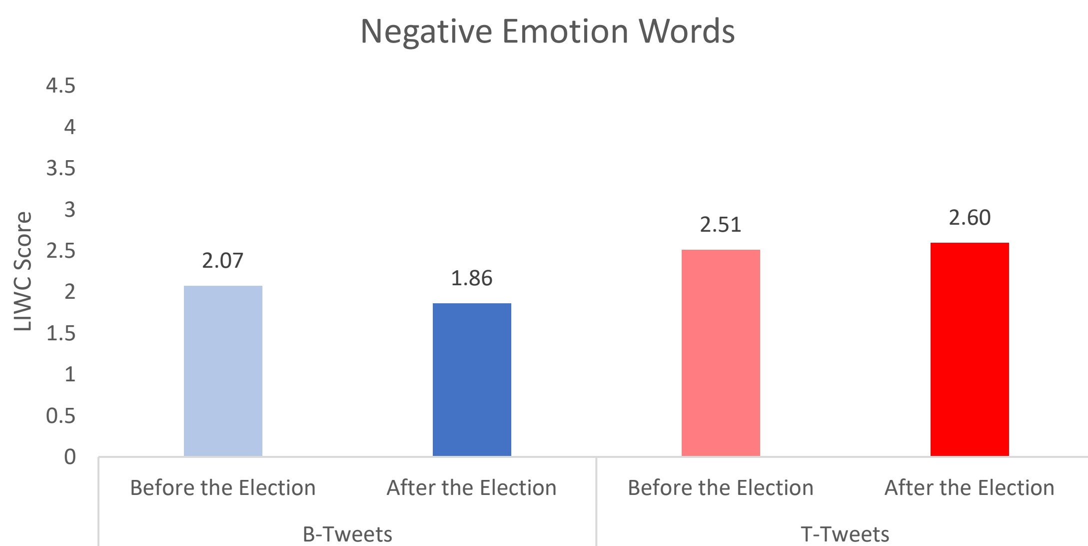


Figure 14. Mean Scores for Negative Emotion Categories Split by Group and Time

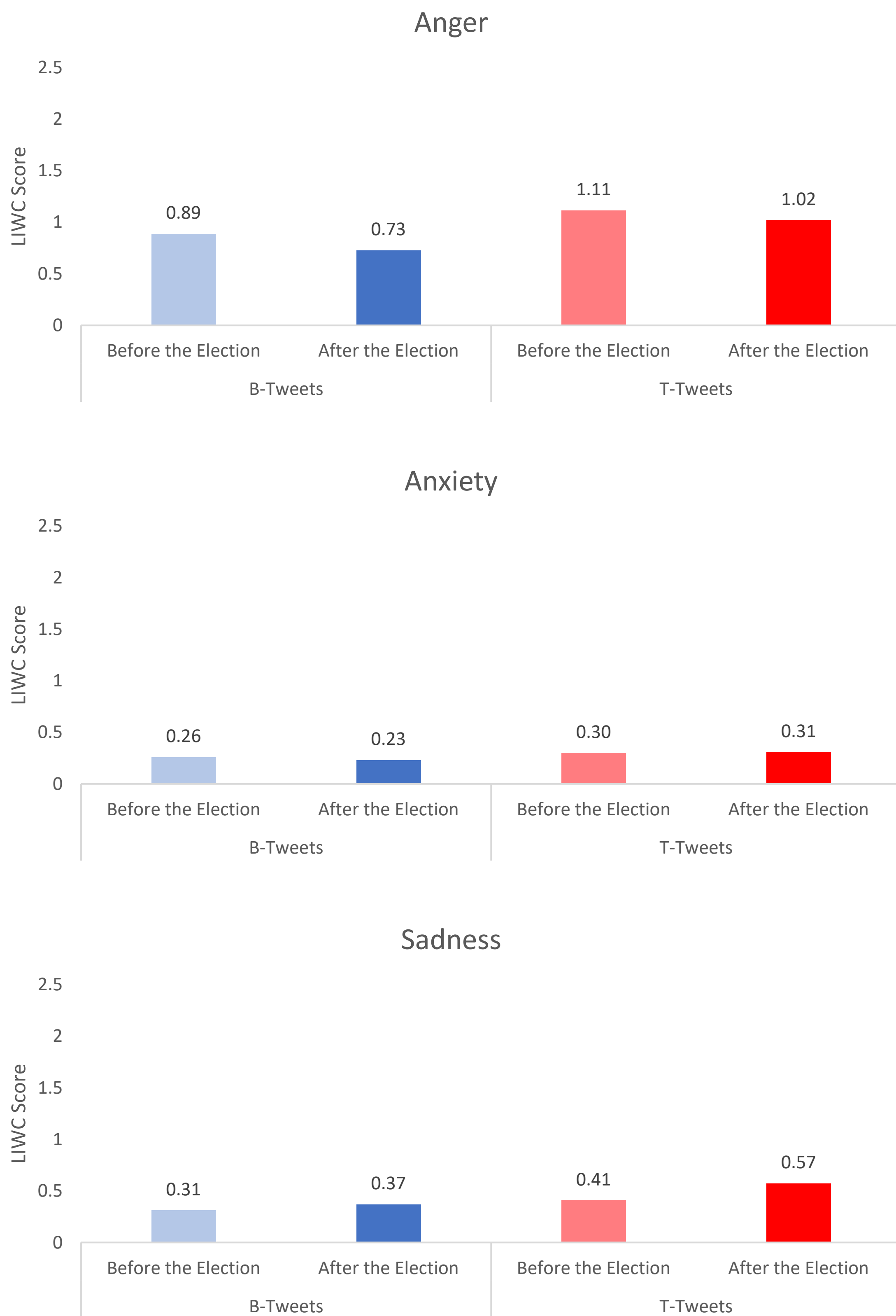
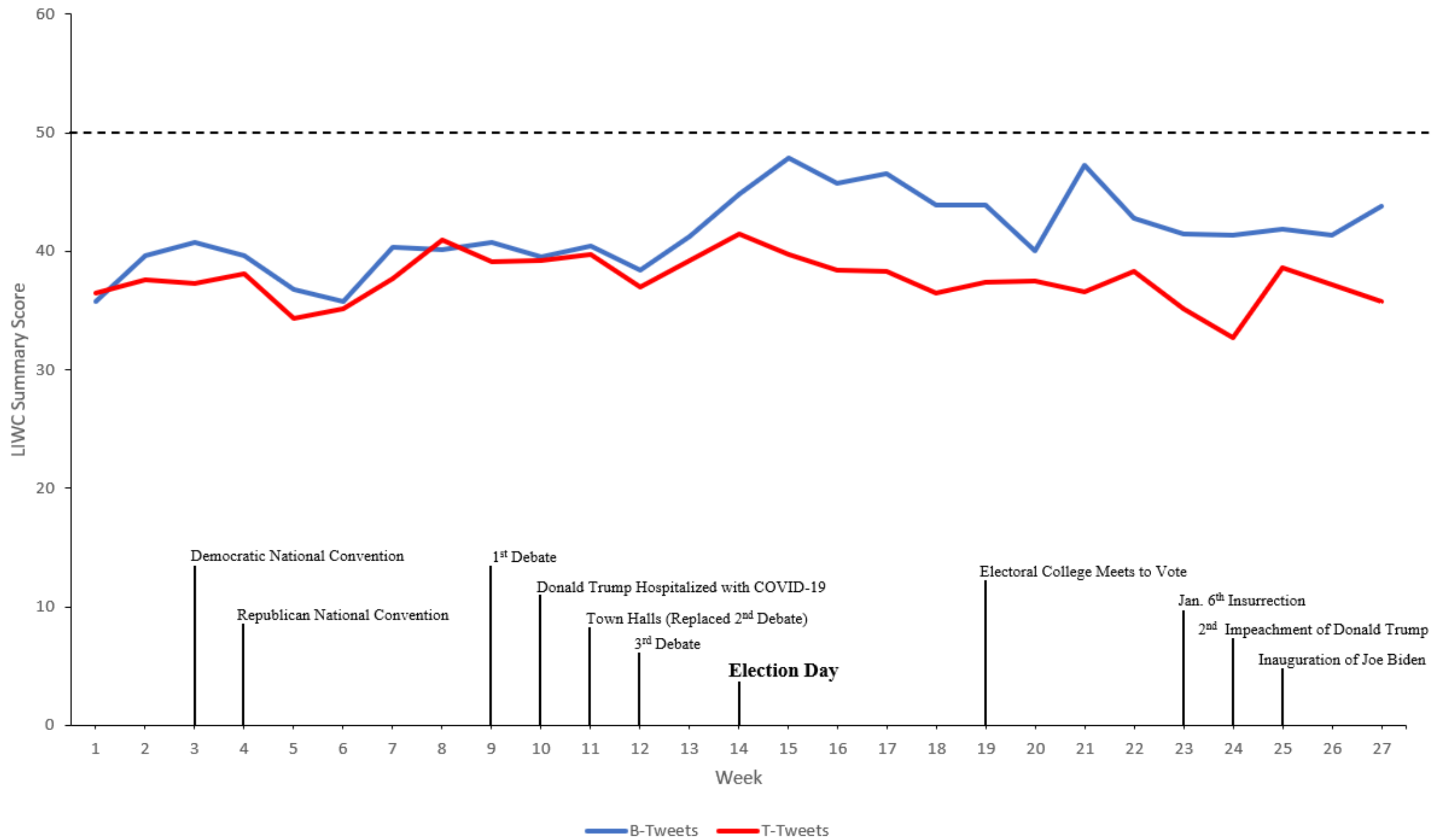


Figure 15. Weekly Trends for Emotional Tone (Annotated with Major Events)



Note. Emotional tone uses a different scale than the normal LIWC scores. The scores range from 0 to 100. Scores above 50 have a positive sentiment and scores less than 50 are considered negative.

Figure 16. Weekly Trends for Positive Emotion Words

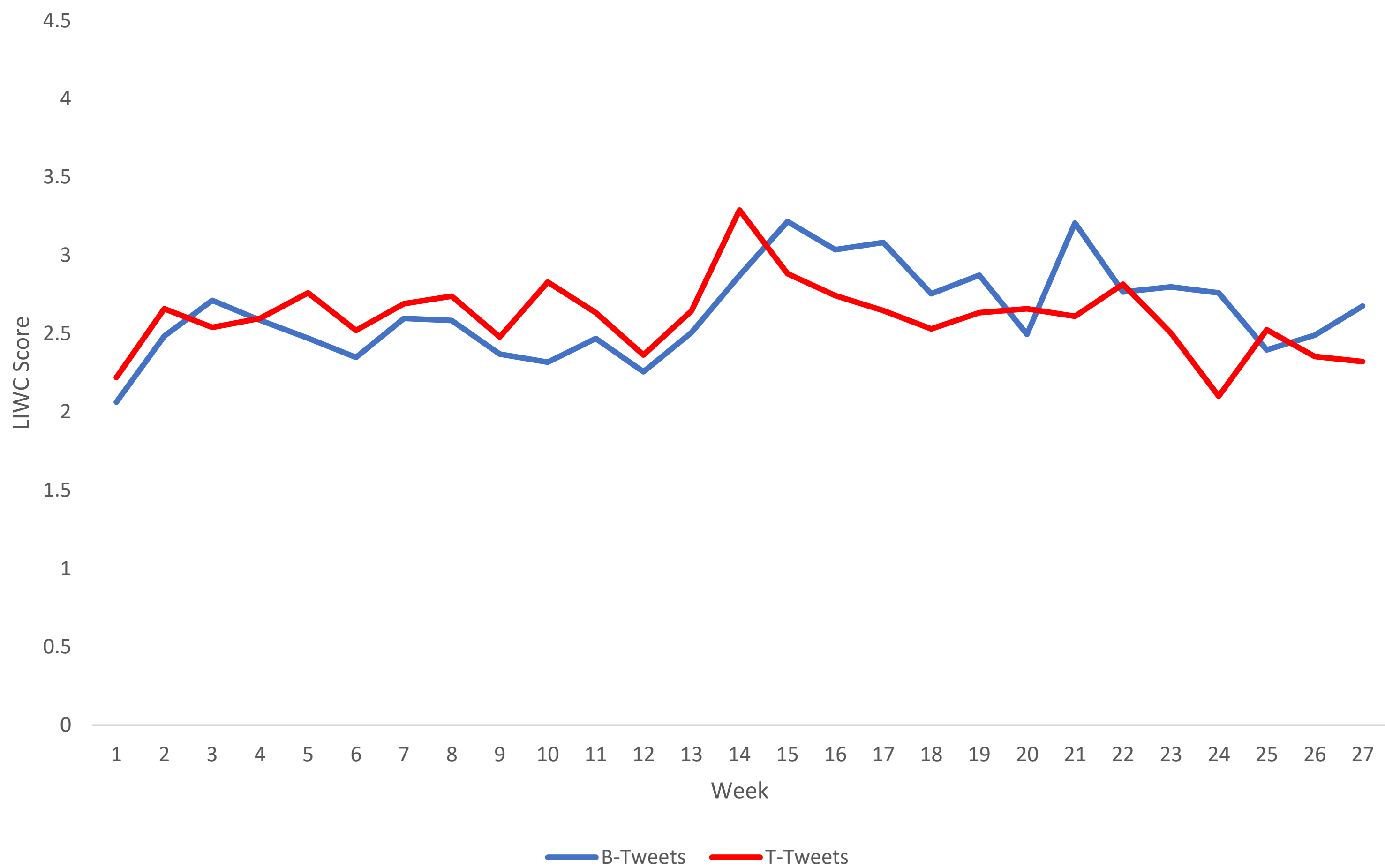


Figure 17. Weekly Trends for Negative Emotion Words

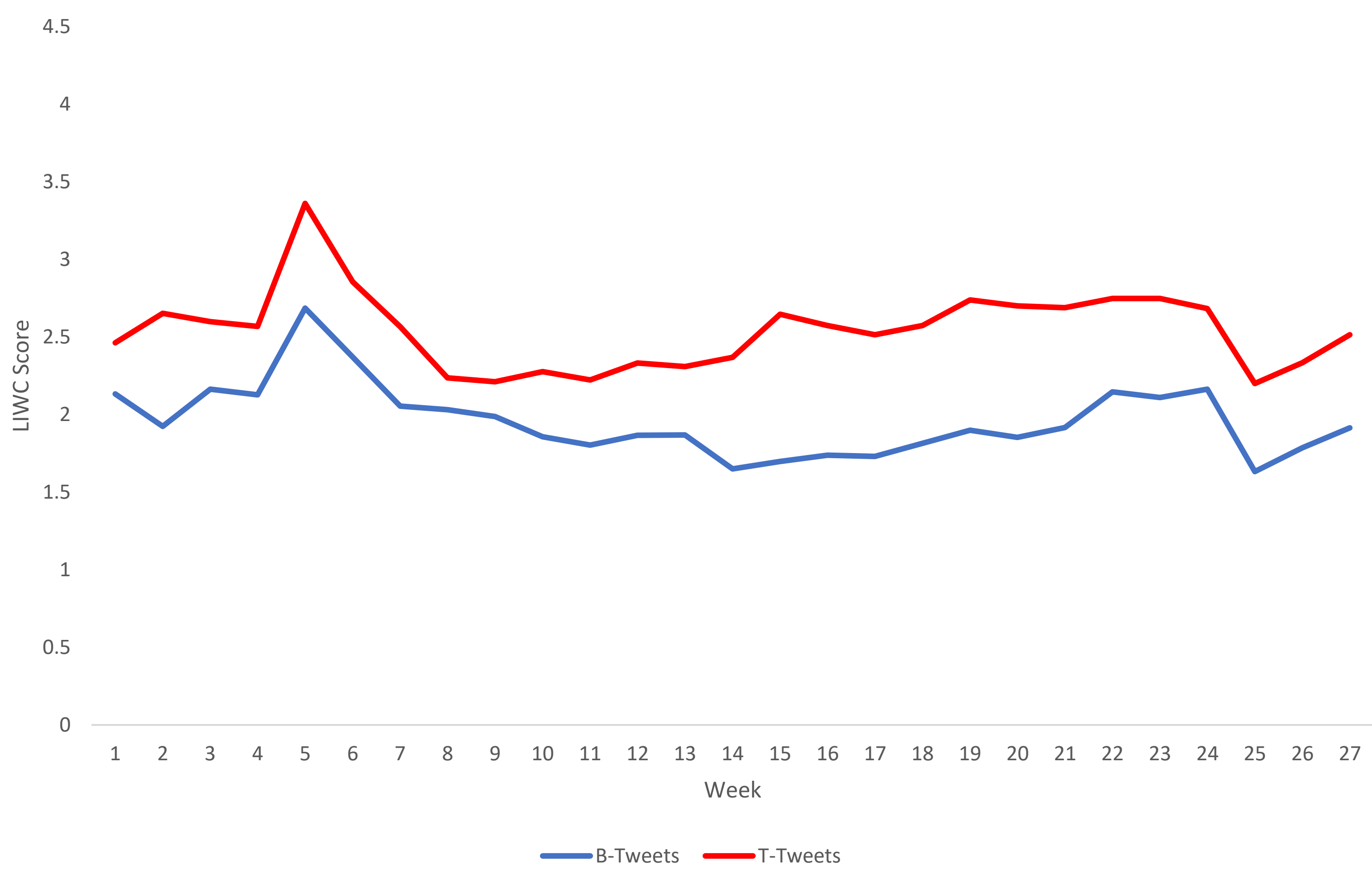


Figure 18. Weekly Trends for Anger Words

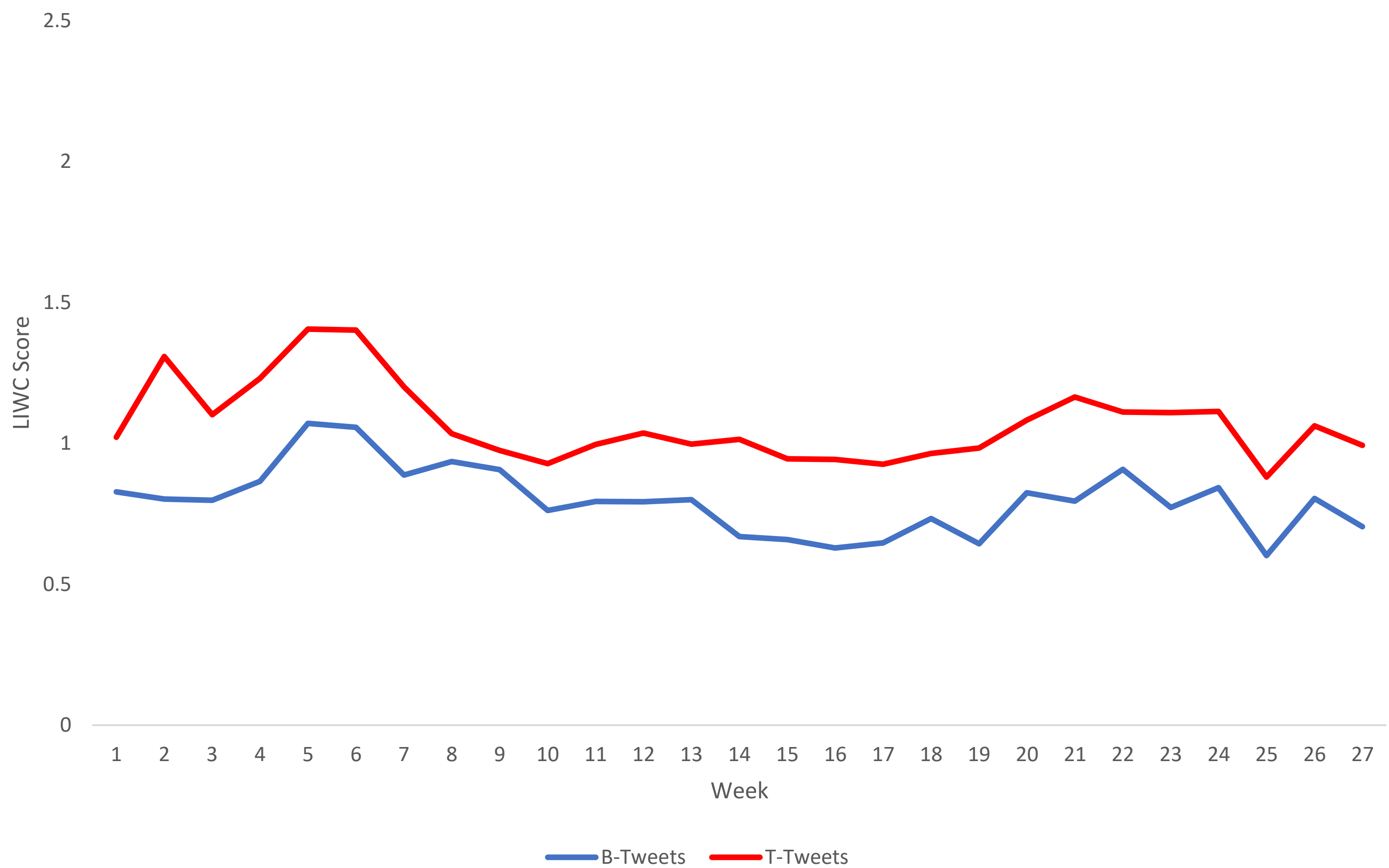


Figure 19. Weekly Trends for Anxiety Words

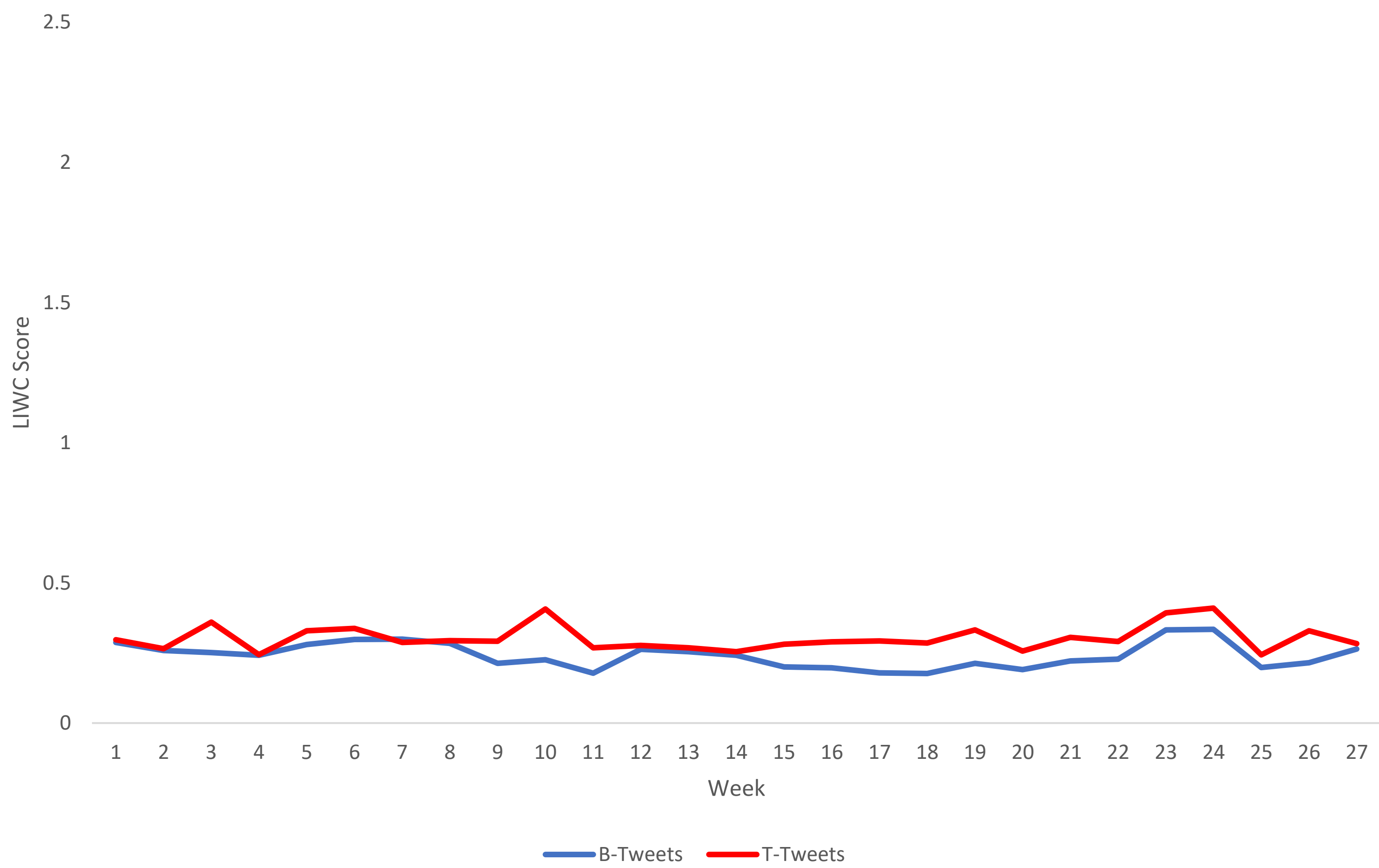


Figure 20. Weekly Trends for Sadness Words

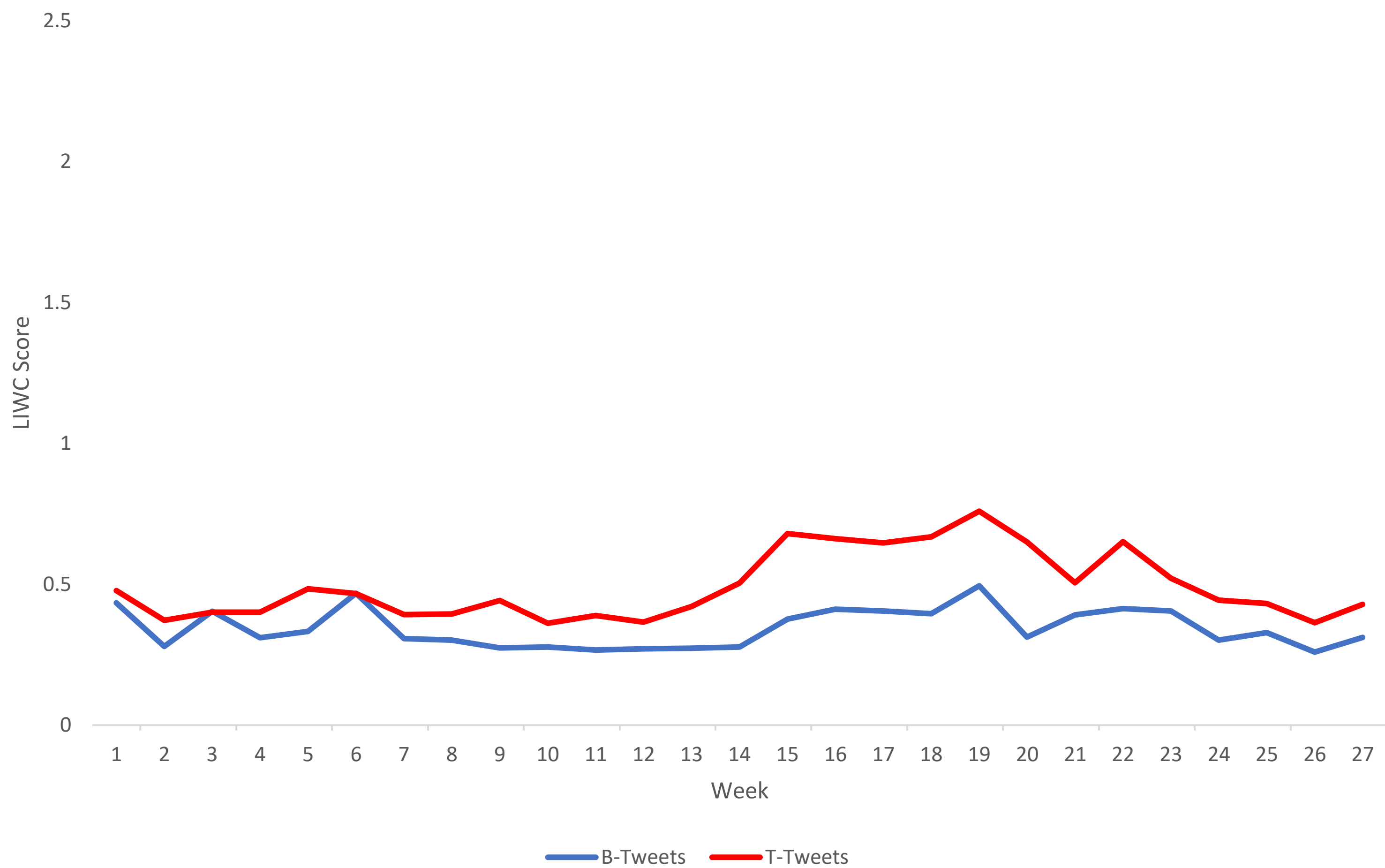


Table 11. Means for the Moral Foundations (Virtues) Split by Group and Time.

Before the Election

	Care		Fairness		Loyalty		Authority		Sanctity	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.35	1.24	0.14	0.84	0.48	1.44	0.85	1.94	0.20	1.13
T-Tweets	0.46	1.48	0.19	0.91	0.45	1.37	1.03	2.18	0.24	1.29

After the Election

	Care		Fairness		Loyalty		Authority		Sanctity	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.34	1.22	0.19	0.89	0.53	1.46	1.36	2.59	0.16	1.02
T-Tweets	0.31	1.19	0.24	1.01	0.49	1.42	1.07	2.29	0.17	1.06

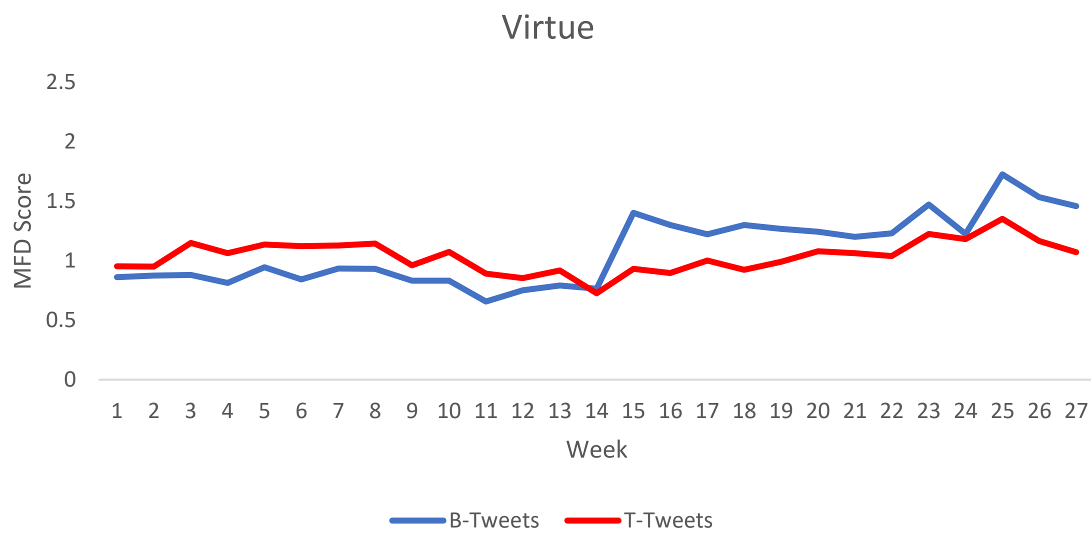
Note. Tweets from the week of the election (Week 14) were excluded.

Table 12. Means for the Moral Foundations (Vices) Split by Group and Time

	After the Election									
	Care		Fairness		Loyalty		Authority		Sanctity	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.27	1.03	0.29	1.50	0.01	0.21	0.17	0.83	0.28	1.19
T-Tweets	0.45	1.43	0.33	1.23	0.02	0.31	0.25	1.02	0.34	1.43
	Before the Election									
	Care		Fairness		Loyalty		Authority		Sanctity	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
B-Tweets	0.36	1.24	0.23	1.07	0.02	0.25	0.13	0.79	0.37	1.36
T-Tweets	0.47	1.44	0.28	1.21	0.02	0.27	0.12	0.72	0.43	1.56

Note. Tweets from the week of the election (Week 14) were excluded.

Figure 21. Weekly Trends for the Authority Foundation



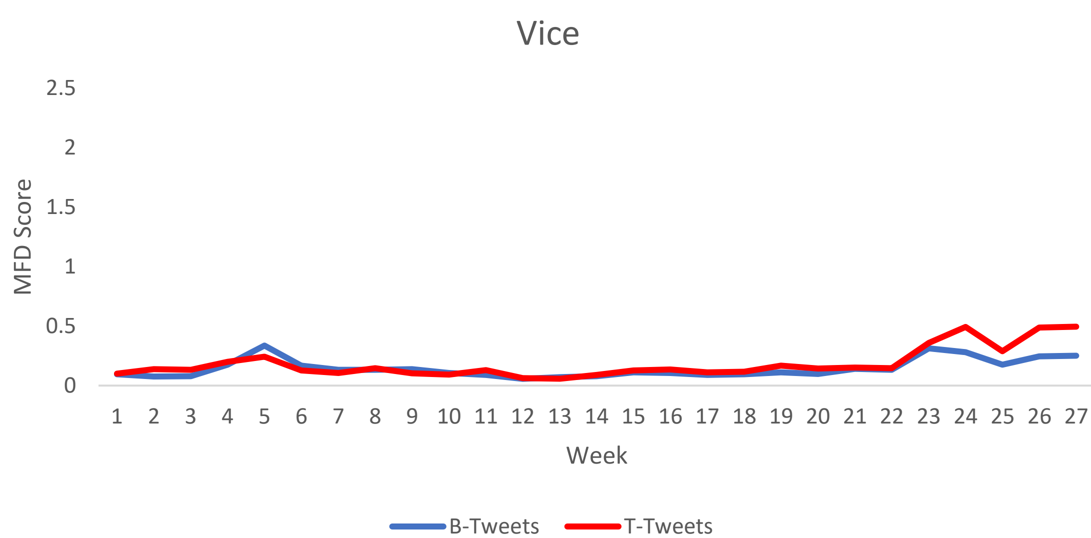
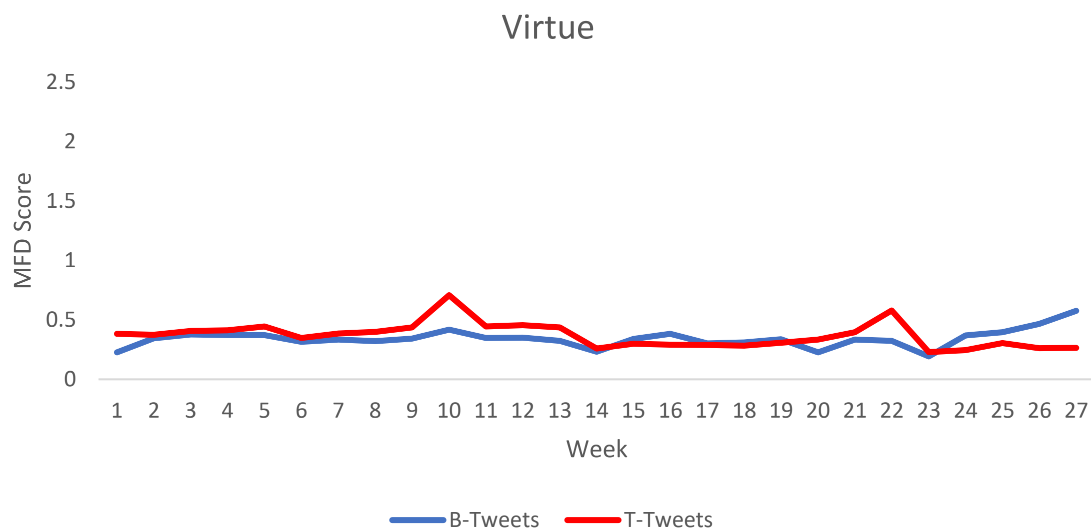


Figure 22. Weekly Trends for the Care Foundation



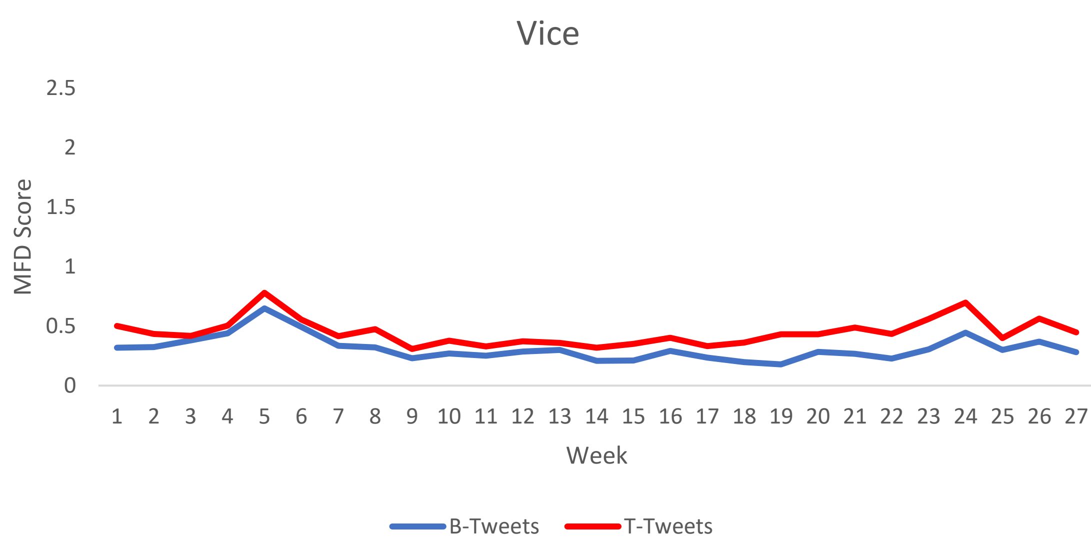
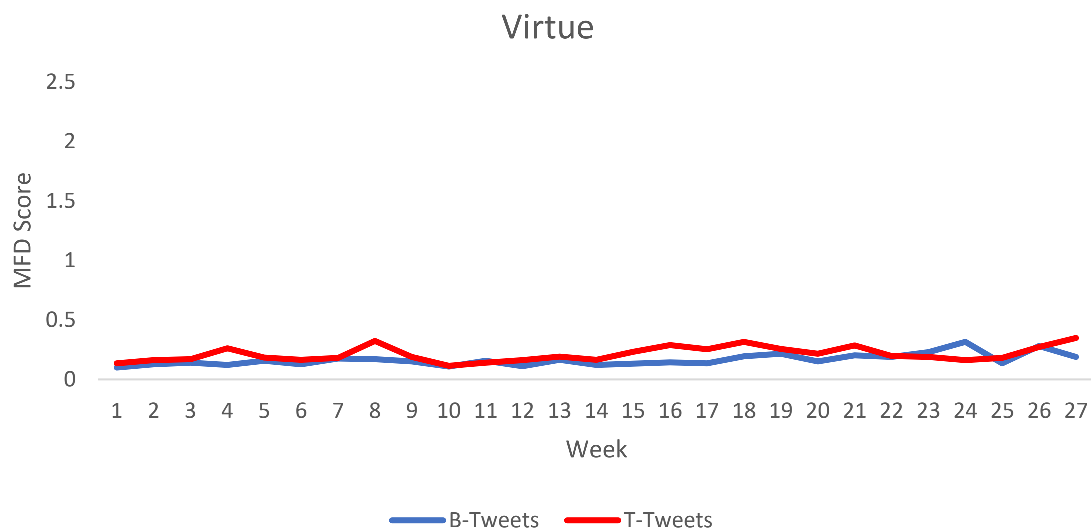


Figure 23. Weekly Trends for the Fairness Foundation



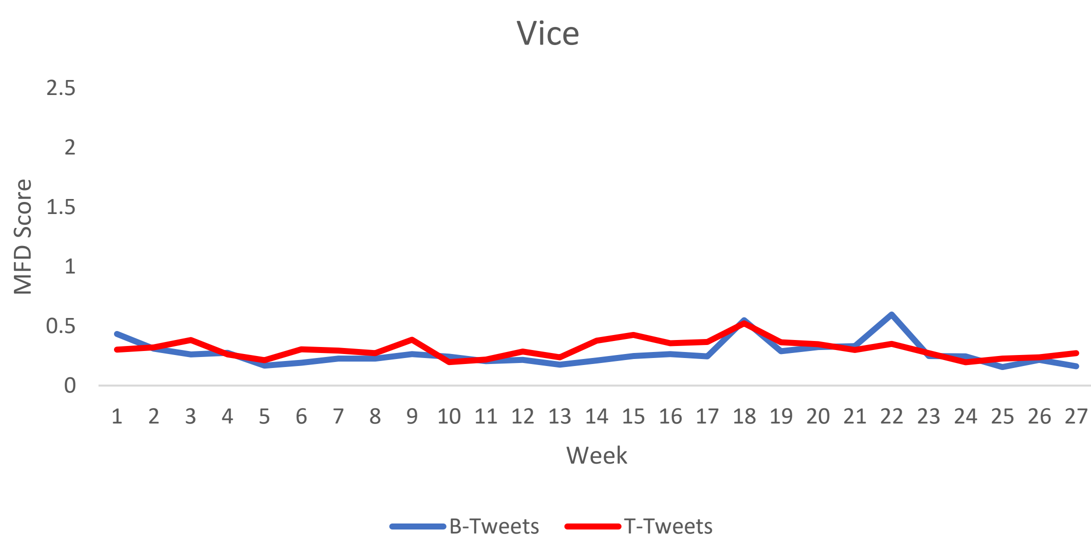
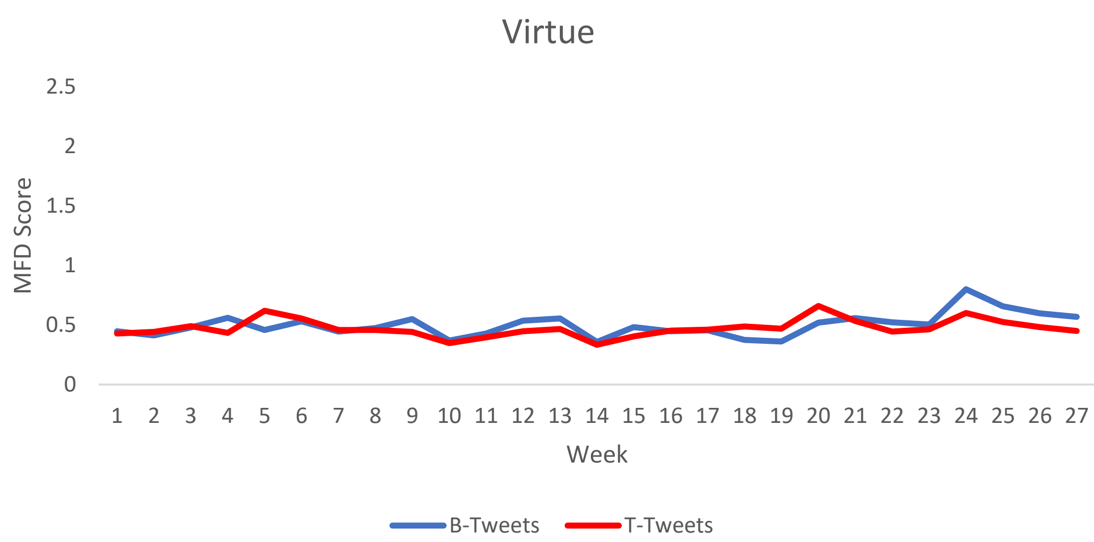


Figure 24. Weekly Trends for the Loyalty Foundation



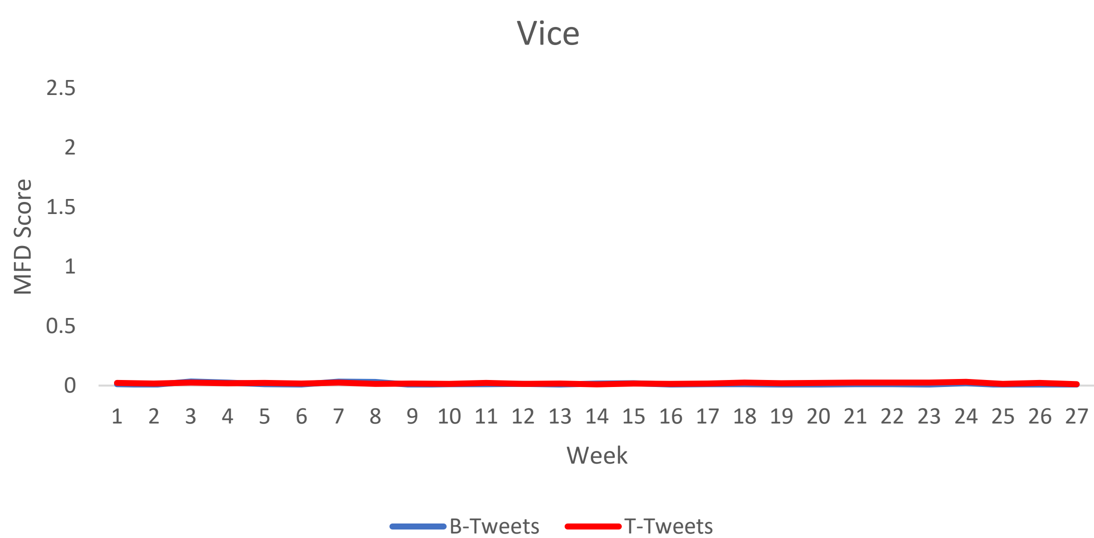
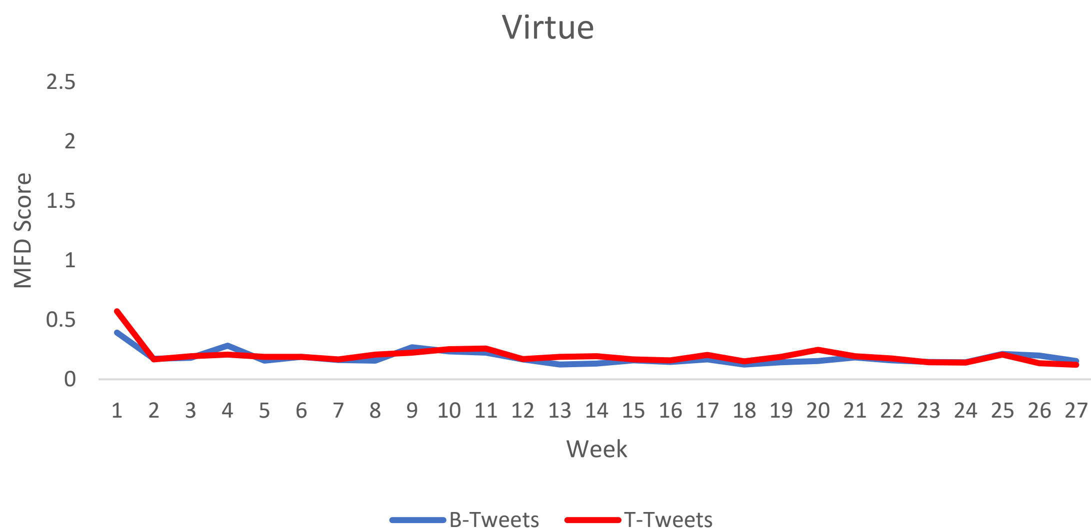


Figure 25. Weekly Trends for the Sanctity Foundation



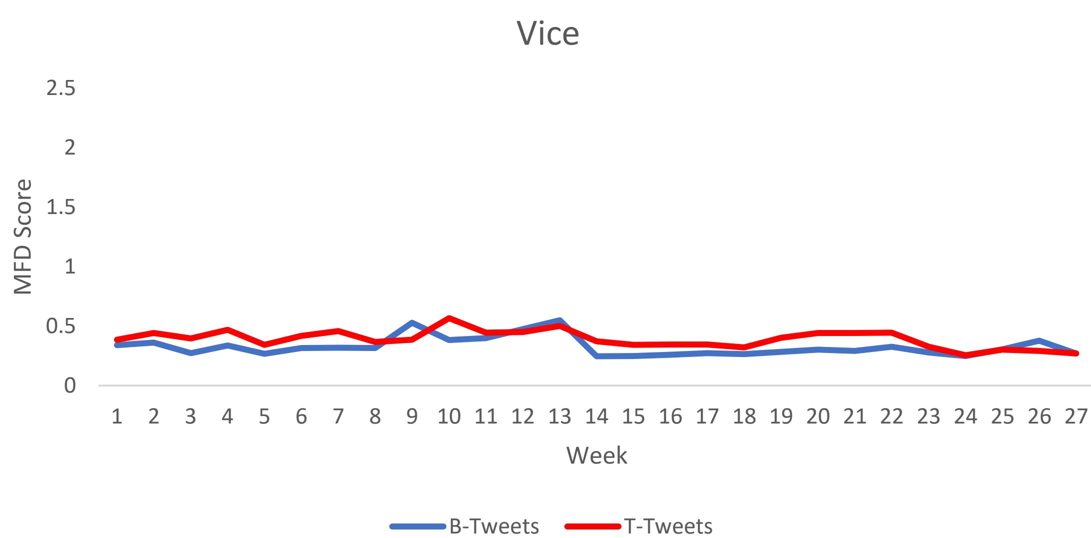


Table 13. Results of Welch t-Tests Comparing LIWC Scores in B-Tweets and T-Tweets

LIWC Categories	df	t	p	95% CI	Mean for B-Tweets	Mean for T-Tweets
Tone	63298442	516.82	$p < 0.001$	[4.639, 4.674]	42.43	37.78
Positive Emotions	63212500	112.39	$p < 0.001$	[.100, .104]	2.71	2.61
Negative Emotions	62822572	-766.56	$p < 0.001$	[-.646, -.642]	1.91	2.55
Anger	62883050	-501.41	$p < 0.001$	[-.280, -.278]	0.78	1.06
Anxiety	63224653	-234.42	$p < 0.001$	[-.064, -.063]	0.24	0.30
Sadness	60016062	-450.88	$p < 0.001$	[-.165, -.163]	0.34	0.50

Table 14. Results of Welch t-Tests Comparing the Moral Foundations in B-Tweets and T-Tweets

Moral Foundations	df	t	p	95% CI	Mean for B-Tweets	Mean for T-Tweets
Care (Virtue)	63384926	-110.50	$p < 0.001$	[-.036, -.034]	0.33	0.37
Care (Vice)	60862941	-494.82	$p < 0.001$	[-.158, -.157]	0.30	0.45
Fairness (Virtue)	63171395	-229.58	$p < 0.001$	[-.053, -.052]	0.17	0.22
Fairness (Vice)	62582043	-147.60	$p < 0.001$	[-.048, -.046]	0.26	0.31
Loyalty (Virtue)	63202516	50.85	$p < 0.001$	[.017, .019]	0.49	0.47
Loyalty (Vice)	61716286	-82.84	$p < 0.001$	[-.006, -.005]	0.01	0.02
Authority (Virtue)	63049858	119.53	$p < 0.001$	[.067, .070]	1.11	1.04
Authority (Vice)	63028606	-232.84	$p < 0.001$	[-.050, -.049]	0.14	0.19
Sanctity (Virtue)	63282599	-86.07	$p < 0.001$	[-.024, -.023]	0.17	0.20
Sanctity (Vice)	62262429	-207.31	$p < 0.001$	[-.073, -.071]	0.31	0.38

Table 15. Results of Welch t-Tests Comparing LIWC Scores Before and After the Election

LIWC Categories	df	t	p	95% CI	Mean Before the Election	Mean After the Election
Tone	51139102	180.70	$p < 0.001$	[1.701, 1.739]	38.66	40.38
Positive Emotions	52111008	184.97	$p < 0.001$	[.172, .176]	2.53	2.70
Negative Emotions	50967303	-766.56	$p < 0.001$	[-.646, -.642]	2.31	2.25
Anger	49715166	-210.69	$p < 0.001$	[-.127, -.125]	1.01	0.88
Anxiety	49118101	-29.22	$p < 0.001$	[-.009, -.008]	0.28	0.27
Sadness	56572488	304.24	$p < 0.001$	[.114, .116]	0.36	0.48

Table 16. Results of Welch t-Tests Comparing the Moral Foundations Before and After the Election

Moral Foundations	df	t	p	95% CI	Mean Before the Election	Mean After the Election
Care (Virtue)	46106413	-233.56	$p < 0.001$	[-.082, -.081]	0.41	0.33
Care (Vice)	48194573	-136.20	$p < 0.001$	[-.048, -.047]	0.42	0.37
Fairness (Virtue)	53609461	213.67	$p < 0.001$	[.051, .052]	0.17	0.22
Fairness (Vice)	55953529	160.40	$p < 0.001$	[.052, .053]	0.26	0.31
Loyalty (Virtue)	51638274	108.42	$p < 0.001$	[.040, .042]	0.47	0.51
Loyalty (Vice)	51928939	-0.98	$p < 0.001$	[-2.05^{-04}, 6.83^{-05}]	0.02	0.02
Authority (Virtue)	55687444	439.77	$p < 0.001$	[.260, .262]	0.95	1.21
Authority (Vice)	56838065	366.69	$p < 0.001$	[.081, .082]	0.13	0.21
Sanctity (Virtue)	45118491	-183.18	$p < 0.001$	[-.057, -.056]	0.22	0.17
Sanctity (Vice)	46992233	-235.48	$p < 0.001$	[-.090, -.088]	0.40	0.31

REFERENCES

Alper, S., & Yilmaz, O. (2019). How is the Big Five related to moral and political convictions: The moderating role of the WEIRDness of the culture. *Personality and Individual Differences*, *145*, 32-38. https://doi.org/10.1016/j.paid.2019.03.018.

Alpers, G. W., Winzelberg, A. J., Classen, C., Roberts, H., Dev, P., Koopman, C., & Taylor, C. B. (2005). Evaluation of computerized text analysis in an Internet breast cancer support group. *Computers in Human Behavior*, *21*(2), 361-376.

Bantum, E. O. C., & Owen, J. E. (2009). Evaluating the validity of computerized content analysis programs for identification of emotional expression in cancer narratives. *Psychological Assessment*, *21*(1), 79.

Bayrak, F., & Alper, S. (2021). A tale of two hashtags: An examination of moral content of pro-and anti-government tweets in Turkey. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.2763.

Bermingham, A., & Smeaton, A. (2011, November). On using Twitter to monitor political sentiment and predict election results. In *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011)* (pp. 2-10).

Biden, J. [@POTUS]. (2021, January 20). *There is no time to waste when it comes to tackling the crises we face. That's why today, I am* [Tweet]. Twitter. https://twitter.com/POTUS/status/1351946842838347776

Bos, L., & Minihold, S. (2021). The Ideological Predictors of Moral Appeals by European Political Elites; An Exploration of the Use of Moral Rhetoric in Multiparty Systems. *Political Psychology*. https://doi.org/10.1111/pops.12739

Brown, E. K., & Silver, J. R. (2020). The moral foundations of crime control in American presidential platforms, 1968–2016. *Punishment & Society*, https://doi.org/10.1177/1462474520966979

Carney, D. R., Jost, J. T., Gosling, S. D., & Potter, J. (2008). The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, *29*(6), 807-840.

Clifford, S., & Jerit, J. (2013). How words do the work of politics: Moral foundations theory and the debate over stem cell research. *The Journal of Politics*, *75*(3), 659-671.

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of communication*, *64*(2), 317-332.

Confessore, N. (2018, April 4). Cambridge Analytica and Facebook: The Scandal and the Fallout So Far. *The New York Times*. https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

Confessore, N., & Hakim, D. (2017, March 6). Data Firm Says ‘Secret Sauce’ Aided Trump; Many Scoff. *The New York Times*. https://www.nytimes.com/2017/03/06/us/politics/cambridge-analytica.html

Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011, July). Political polarization on twitter. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 5, No. 1).

Ditto, P. H., Liu, B. S., Clark, C. J., Wojcik, S. P., Chen, E. E., Grady, R. H., ... & Zinger, J. F. (2019). At least bias is bipartisan: A meta-analytic comparison of partisan bias in liberals and conservatives. *Perspectives on Psychological Science*, *14*(2), 273-291.

Doğruyol, B., Alper, S., & Yilmaz, O. (2019). The five-factor model of the moral foundations theory is stable across WEIRD and non-WEIRD cultures. *Personality and Individual Differences*, *151*, 109547.

Frimer, J. A., Biesanz, J. C., Walker, L. J., & MacKinlay, C. W. (2013). Liberals and conservatives rely on common moral foundations when making moral judgments about influential people. *Journal of Personality and Social Psychology*, *104*(6), 1040.

Frimer, J. A., Boghrati, R., Haidt, J., Graham, J., & Dehgani, M. (2019). *Moral Foundations Dictionary for Linguistic Analyses 2.0*. Unpublished manuscript.

Frimer, J. A., Skitka, L. J., & Motyl, M. (2017). Liberals and conservatives are similarly motivated to avoid exposure to one another's opinions. *Journal of Experimental Social Psychology*, *72*, 1-12.

Garten, J., Boghrati, R., Hoover, J., Johnson, K. M., & Dehghani, M. (2016, July). Morality between the lines: Detecting moral sentiment in text. In *Proceedings of IJCAI 2016 Workshop on Computational Modeling of Attitudes*.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366.

Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011, October). Predicting personality from twitter. In *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing* (pp. 149-156). IEEE.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political analysis, 21(3), 267-297. https://doi.org/10.1093/pan/mps028.

Grover, T., Bayraktaroglu, E., Mark, G., & Rho, E. H. R. (2019). Moral and affective differences in us immigration policy debate on twitter. *Computer Supported Cooperative Work (CSCW)*, *28*(3), 317-355. https://doi.org/10.1007/s10606-019-09357-w.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98-116.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55-66.

Halberstam, Y., & Knight, B. (2016). Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. *Journal of public economics*, *143*, 73-88.

Hoover, J., Portillo-Wightman, G., Yeh, L., Havaldar, S., Davani, A. M., Lin, Y., ... & Dehghani, M. (2020). Moral Foundations Twitter Corpus: A collection of 35k tweets annotated for moral sentiment. *Social Psychological and Personality Science*, *11*(8), 1057-1071. https://doi.org/10.1177/1948550619876629.

Hopp, F. R., Fisher, J. T., Cornell, D., Huskey, R., & Weber, R. (2021). The extended Moral Foundations Dictionary (eMFD): Development and applications of a crowd-sourced approach to extracting moral intuitions from text. *Behavior Research Methods*, *53*(1), 232-246.

Kahn, J. H., Tobin, R. M., Massey, A. E., & Anderson, J. A. (2007). Measuring emotional expression with the Linguistic Inquiry and Word Count. *The American journal of psychology*, 263-286.

Kalimeri, K., Beiró, M. G., Delfino, M., Raleigh, R., & Cattuto, C. (2019). Predicting demographics, moral foundations, and human values from digital behaviours. *Computers in Human Behavior*, *92*, 428-445.

Kaur, R., & Sasahara, K. (2016). Quantifying moral foundations from various topics on Twitter conversations. In *2016 IEEE International Conference on Big Data (Big Data)* (pp. 2505-2512). IEEE.

Kearney, M. W. (2019). rtweet: Collecting and analyzing Twitter data, Journal of Open Source Software, 4, 42. 1829. doi:10.21105/joss.01829 (R package version 0.7.0)

Kennedy, B., Ashokkumar, A., Boyd, R. L., & Dehghani, M. (2021). Text Analysis for Psychology: Methods, Principles, and Practices. PsyArXiv. https://doi.org/10.31234/osf.io/h2b8t

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. *Journal of Research in Personality*, *46*(2), 184-194.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the national academy of sciences*, *110*(15), 5802-5805.

Kraft, P. W. (2018). Measuring morality in political attitude expression. *The Journal of Politics*, *80*(3), 1028-1033.

Lewis, P. G. (2019). Moral Foundations in the 2015-16 US Presidential Primary Debates: The Positive and Negative Moral Vocabulary of Partisan Elites. *Social Sciences, 8*(8), 233.

Lipsitz, K. (2018). Playing with emotions: The effect of moral appeals in elite rhetoric. *Political Behavior*, *40*(1), 57-78.

Murthy, D. (2015). Twitter and elections: are tweets, predictive, reactive, or a form of buzz?. *Information, Communication & Society*, *18*(7), 816-831.

Neiman, J. L., Gonzalez, F. J., Wilkinson, K., Smith, K. B., & Hibbing, J. R. (2016). Speaking different languages or reading from the same script? Word usage of Democratic and Republican politicians. *Political Communication*, *33*(2), 212-240. https://doi.org/10.1080/10584609.2014.969465

O'Connor, B., Balasubramanyan, R., Routledge, B., & Smith, N. (2010, May). From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 4, No. 1).

Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015). Linguistic Inquiry and Word Count: LIWC2015. Austin, TX: Pennebaker Conglomerates (www.LIWC.net).

Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences*, *11*(1), 37-43.

Quercia, D., Kosinski, M., Stillwell, D., & Crowcroft, J. (2011, October). Our twitter profiles, our selves: Predicting personality with twitter. In *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing* (pp. 180-185). IEEE.

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., ... & Funder, D. C. (2014). The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, 107(4), 677.

R Core Team (2020). R: A language and environment for statistical computing. [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sagi, E., & Dehghani, M. (2014). Measuring moral rhetoric in text. *Social science computer review*, *32*(2), 132-144.

Sánchez-Rada, J. F., & Iglesias, C. A. (2019). Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison. *Information Fusion*, *52*, 344-356.

Sano, Y., Takayasu, H., Havlin, S., & Takayasu, M. (2019). Identifying long-term periodic cycles and memories of collective emotion in online social media. *PloS one*, *14*(3).

Serfass, D. G., & Sherman, R. A. (2015). Situations in 140 characters: Assessing real-world situations on Twitter. *PloS one*, *10*(11).

Sylwester, K., & Purver, M. (2015). Twitter language use reflects psychological differences between democrats and republicans. *PloS one*, *10*(9).

Stolerman, D., & Lagnado, D. (2020). The moral foundations of human rights attitudes. *Political Psychology*, *41*(3), 439-459.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, *29*(1), 24-54. https://doi.org/10.1177/0261927X09351676.

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, *62*(2), 406-418.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2011). Election forecasts with Twitter: How 140 characters reflect the political landscape. *Social Science Computer Review*, *29*(4), 402-418.

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In *Proceedings of the ACL 2012 system demonstrations* (pp. 115-120). Association for Computational Linguistics.

Wang, W., Hernandez, I., Newman, D. A., He, J., & Bian, J. (2016). Twitter analysis: Studying US weekly trends in work stress and emotion. *Applied Psychology*, *65*(2), 355-378.

Wang, S. Y. N., & Inbar, Y. (2021). Moral-language use by US political elites. *Psychological Science*, *32*(1), 14-26. DOI:10.1177/0956797620960397

Yang, J., & Leskovec, J. (2011). Patterns of temporal variation in online media. In *Proceedings of the fourth ACM international conference on Web search and data mining* (pp. 177-186).

Yaqub, U., Chun, S. A., Atluri, V., & Vaidya, J. (2017). Analysis of political discourse on twitter in the context of the 2016 US presidential elections. *Government Information Quarterly*, *34*(4), 613-626.

Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, *44*(3), 363-373.

Yilmaz, O., & Saribay, S. A. (2017). Activating analytic thinking enhances the value given to individualizing moral foundations. *Cognition*, *165*, 88-96.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, *112*(4), 1036-1040.

Zhao, N., Jiao, D., Bai, S., & Zhu, T. (2016). Evaluating the validity of simplified Chinese version of LIWC in detecting psychological expressions in short texts on social network services. *PLoS One*, 11(6), e0157947.