

Characteristics of Tweets from @realDonaldTrump in Early 2020 (with a Mid-Summer Update)

Cynthia Whissell

Psychology Department,
Laurentian University

ABSTRACT

This research examines the tweeting behavior of US president Donald Trump during the early crisis months of 2020. A study of 1507 tweets posted in January-April 2020, and a comparison of these to tweets from the first year of his presidency (2017) led to several statistically significant conclusions. Overall, the language of the tweets remained somewhat positive or pleasant. Trump's tweets were both longer and much more frequent in 2020 (the rate rose from 7 to 30 per day). The pleasantness of the language in tweets was negatively related to their popularity (popular tweets used relatively unpleasant language). The president's tweeting behavior modified somewhat (e.g. fewer hourly tweets) in conjunction with the coronavirus crisis and the abrupt decline of the markets in March. The tweets gave evidence of a distinctly Trumpian vocabulary that favored words and phrases such as "great", "hoax", and "fake news media". A mid-summer data update with 151 tweets indicated that Trump's tweets had dropped even further in positivity, falling below the norm for everyday English.

Keywords: Tweets, Trump, Contents, Emotion, Popularity.

INTRODUCTION

This research is part of a continuing effort to study the tweets of US President Donald Trump. Earlier work compared Trump's tweets to those of Obama as President (Whissell, 2018a), and also examined factors related to the popularity of Trump's tweets (Whissell, 2018b). This work focuses on tweets posted by @realDonaldTrump in the first three months of 2020, a time of national crisis. At the beginning of the year the issue was Trump's impeachment and the Senate's reaction and in March the COVID-19 health crisis began to take center stage.

This crisis was associated with a world-wide pandemic and had serious consequences such as widespread unemployment in the US and a steep drop in the markets. Tweets were studied in terms of language (e.g., vocabulary), style (e.g., tweet length), popularity (being marked as favorites) and emotional tone (pleasantness or positivity).

The questions of the research fell into three major categories. First, it was queried whether results from early 2020 tweets would confirm previously published results (Whissell, 2018a, b). If so, Trump's tweets should be mildly pleasant in emotional tone, and his less pleasant tweets should be more popular. The second research question addressed the predictability of the emotional tone of

language in Trump's tweets. It was expected that language and style measures could be employed to successfully predict positivity. The final question, addressed the predictability of tweet popularity, again on the basis of measures of language and style.

METHOD

Tweets Studied

Tweets were downloaded from the website trumptwitterarchive.com in late April of 2020. Retweets were excluded from this study. The first 1507 tweets of the year from the @realDonaldTrump account were downloaded. The time stamp employed was Eastern Daylight Time.

Measures of Tweets

There were seven vocabulary measures (all of which had been defined *a priori*) taken on each tweet. These were all binary and represented whether the author had or had not included words belonging to various families in the tweet. The families were *democrats* ("dems" or any word including the root "democrat-"), *media* (the word "media"), *virus* (the words "virus" or "COVID" or "corona" or any word including these), *I* (the word "I"), *we* (the word "we"), *great* (the word "great") and *impeachment* (any word including the root "impeach-"). Some of these vocabulary families had been employed in previous research and others reflected the issues of the day.

The Dictionary of Affect in Language (Whissell 2009) is a technique developed for the sentiment analysis of texts. It compares words employed in a text to emotional scores generated for these words by raters. Scores for pleasantness were employed in this research. High scores for pleasantness indicate a positive emotional tone and low scores a negative tone. The everyday English normative pleasantness, established on the basis of a standardization sample of more than a third of a million words from different sources, is 50. Pleasantness scores were computed for each tweet by averaging over all the words in it.

Additional measures of tweets included *length* (in number of words), and *traffic density* (the number of tweets posted in the particular hour in which a tweet appeared). Tweets were characterized in terms of the number of readers who had marked them as *favorites*, and the number of readers who had *retweeted* them. Log transformations of these positively skewed variables were used in their stead in various correlational analyses. A measure external to the tweets themselves was the *Dow Jones Industrial Average* (DJIA) for the week in which the tweet was posted. This was accessed from the site statista.com/statistics/1104278/weekly-performance-of-djia-index/, and represented the health of the markets. A high value is associated with a healthier market and a drop in value (which took place in March, for example) with a fall in the markets.

The variables of the research are summarized and exemplified in Table 1 which shows detailed results for a single tweet posted on January 2, 2020. Some tweets were not in English and others contained only reference addresses, so there were 1364 tweets with complete scores for all variables.

Table 1 Information for a single tweet in terms of all the research variables.

Tweet posted 1am January 2nd:

"Best equipment & finest military in the World. On site quickly! <https://t.co/cDUtqFLLYB>"

Variable	Score (and Explanation)
Democrats	0 (not mentioned)
Media	0 (not mentioned)
Virus	0 (not mentioned)
I	0 (not used)
We	0 (not used)
Great	0 (not used)
Impeachment	0 (not mentioned)
Pleasantness	64.02 (well above average, positive emotional tone)
Length	10 words (excluding "&" and the http address: shorter than usual)
Favorites	90,120 (number of times marked as favorite)
Retweets	19,763 (number of times retweeted)
Tweet Traffic	4 (four messages posted in that hour)
DJIA	28,584 (strong market: one of the higher values in the data set)

RESULTS

The first question of the study which dealt with the level of pleasantness in the tweets and its relationship to popularity was answered by consulting individual variables. When it came to the predictive schemes, ten variables served as predictors in the analyses (the seven vocabulary measures, length, traffic density, and the DJIA) while three variables were considered as criteria (pleasantness, retweets, and favorites). Because there were intercorrelations among predictors, a principal components analysis with a varimax rotation was employed as a data reduction technique. Five components (the last two with eigenvalues below 1 but above .9) were employed to embody the full set of variables. The components explained 72% of the total variance.

They were interpreted as representing *current issues* (mentions of media [.88], impeachment [.85], and virus [.85]), *long tirades* (use of "I" [.72], length [.69], and mention of democrats [.63]), *high tweet traffic and high markets* (high tweet traffic [.73] and DJIA [.72]), *community* (the use of "we" [.91]) and *praise* (the use of "great" [.95]). Sample tweets with high and low scores on each component are provided in Table 2. Scores computed to represent the five components were orthogonal to one another and would therefore not overlap or hide one another in a predictive scheme.

Table 2: Examples of tweets representing the low and high poles of each component from the principal components analysis

Component (description)	Sample Low Scoring Tweet	Sample High Scoring Tweet
Current Issues (High scoring tweets refer to media, impeachment, and/or the virus)	So true! https://t.co/Ds9MSUIWXA Feb 6	Do Nothing Democrats were busy wasting their time on the Impeachment Hoax, & anything they could do to make the Republican Party look bad, while I was busy calling early boarder & flight closings, putting us way ahead in our battle with the Coronavirus. Dems called it very wrong! Feb 27
Long Tirades (High scoring tweets are longer, use "I" and refer to democrats)	(implied continuation from previous tweet) THE UNITED STATES IS OPEN FOR BUSINESS! Feb 18	As hard as I work, & as successful as our Country has become with our Economy, our Military & everything else, it is ashame that the Democrats make us spend so much time & money on this ridiculous Impeachment Lite Hoax. I should be able to devote all of my time to the REAL USA! Jan 4
High Tweet Traffic, High Markets (High scoring tweets are one of several tweets posted in the same hour, during a time of high DJIA)	I am giving consideration to a QUARANTINE of developing "hot spots", New York, New Jersey, and Connecticut. A decision will be made, one way or another, shortly. March 28	Great photos from a SOUTHERN BORDER WALL Briefing with Secretary of Defense, Mark @EsperDoD! https://t.co/rpGHpdola6 Jan 30
Community (High scoring tweets contain the word "we")	I JUST GOT IMPEACHED FOR MAKING A PERFECT PHONE CALL! Jan 16	There will be a News Conference by the CoronaVirus Task Force today at the White House, 5 P.M. @VP We are working closely with the Governors of our Country's great States, which are a very big factor. They are working hard, along with us, to get the job properly done... (implied continuation to next tweet) March 15
Praise (High scoring tweets contain the word "great")	"We demand fairness" shouts Pelosi and the Do Nothing Democrats, yet the Dems in the House wouldn't let us have 1 witness, no lawyers or even ask questions. It was the most unfair witch-hunt in the history of Congress! Jan 13	So, what the hell has happened to @FoxNews. Only I know! Chris Wallace and others should be on Fake News CNN or MSDNC. How's Shep Smith doing? Watch, this will be the beginning of the end for Fox, just like the other two which are dying in the ratings. Social Media is great! Jan 28

Comparison of findings from 2020 to Trump's 2017 tweets

According to earlier research (Whissell, 2018a) Trump's tweets in 2017, the first full year of his presidency, had a mean pleasantness of 54, and were more positive in tone than Obama's tweets as president (51), as well as being more positive than the normative average for everyday English (50). The tweets from early 2020 had a mean pleasantness of 52.84. According to *t* tests, this value was significantly higher than that for Obama's tweets as president ($p < .01$), and than the everyday English average ($p < .01$). It was significantly lower, however, than the value for 2017 tweets. One difference in the present sample is that it did not include retweets, while the 2017 sample did.

A second difference lies in the crises faced by the president and the country in early 2020. However, it can still be concluded that Trump's tweets were generally pleasant and positive in tone. Tweet length increased significantly from 2017, moving from 22 to 24 words per tweet ($p<.01$). The monthly frequency of tweeting also increased, moving from 217 tweets per month in 2017 to 411 per month in the first three full months of 2020. When retweets were included, Trump's tweeting rate was a staggering 900 tweets per month or roughly 30 tweets + retweets per day. The frequency of tweets from the president's account had more than tripled! As had been noted for the 2017 tweets, popularity was negatively correlated with pleasantness (-.07 for favorites, $p<.02$, and -.12 for retweets, $p<.01$): Trump's most popular tweets were again his less pleasant ones, even though the relationship was a weak one.

Predicting tweet pleasantness and popularity

A multiple regression was employed to predict how pleasant a tweet would be from the five components described in Table 2. This prediction was successful, and every single component contributed significantly to the prediction ($R=.42$, Table 3). More pleasant tweets were less likely to deal with current issues, they were less likely to be long tirades, they were less likely to occur in heavy tweet traffic or during high markets, and they were less likely to include the word "we" but more likely to include the word "great".

Table 3 Results of linear regressions predicting tweet pleasantness, popularity, and retweets from the five components

Component	Pleasantness Beta Weight	Favorites Beta Weight	Retweets Beta Weight
Current Issues	-.13**	-.06*	-.00
Long Tirades	-.28**	.10**	.12**
High traffic/markets	-.10**	-.21**	-.19**
Community	-.14**	.04	.05*
Praise	.13**	-.04	-.04
$R(p)$	.42 (<.001)	.25 (<.001)	.23 (<.001)

* $p<.05$, ** $p<.01$

A second multiple regression was employed to predict how often a tweet was marked as a favorite (log transformed). This regression was also successful (Table 3), but with a lower multiple R (.25) and only three significant predictors. Tweets marked as favorites were less likely to deal with current issues, and to be posted in high traffic and when the markets were strong. They were more likely to be long tirades. A third and final multiple regression indicated that retweets (log transformed) were predicted at a similar level as favorites ($R=.23$).

Significant predictors of frequent retweets were being longer, poor markets and low tweet traffic, and the employment of "we". Transformed scores for favorites and retweets were strongly correlated ($r=.96$, $p<.001$). The largest difference in prediction between favorites and retweets was that avoiding current issues was important to the prediction of favorites, but not the prediction of retweets (t test comparing b weights=3, $p<.01$).

Additional findings

Some additional analyses were performed *post hoc*, after data had been gathered and the main questions of the research had been answered. The first additional analysis investigated whether Trump, as a tweet author, had been sensitive to changes in the markets. The nine other tweet-descriptive variables were correlated with the DJIA. As the markets fell, Trump tweeted less often per hour. He also tended to mention democrats and to use the word “we” and impeachment-related words less often. On the other hand, mentions related to the virus increased, as did mentions of media. Correlations were all weak, ranging from $\pm .07$ to $\pm .18$, $p < .001$.

An increased mention of the virus and a decrease in references to impeachment matches the historical time frame as impeachment was the early issue of 2020, and the pandemic coincided with the fall of the markets. The reduced tweeting rate, the avoidance of “we” and mentions of democrats, and the increased mention of the media are signs of changes in Trump’s tweeting behaviors in association with the market crisis.

A further question focused on the possible existence of authors other than Trump where the tweets were concerned. It had been suggested that tweets posted after 11 am and before 7 pm could have been published under Trump’s authority, while not being written by Trump himself (Feinberg, 2017) (Actually, the hours suggested were 10:30 and 6:30, but a broader range was used to ensure that few Trump tweets were misclassified.) This argument was based on the fact the Trump was often otherwise occupied between the hours of 11 am and 7 pm. Although this division of labor was merely speculative, it was decided to compare tweets published during “Trump hours” ($N=1209$) to those posted during “non-Trump hours” ($N=298$) in terms of the ten measures of the research. A multitude of strong differences would lend support to the conclusion of different authorship. Significant differences between Trump and non-Trump tweets were present for five variables (t values ranging from 2.13 to 3.53, $p < .05$; r^2 between .002 and .008). Tweets posted in Trump hours were more pleasant (53.17 vs 51.58) and shorter (23.7 vs 26.76 words).

They made more mentions of the virus (16% vs 11% of each tweeter’s messages), of impeachment (15% vs 9%) and of the media (13% vs 6%). The last three variables all loaded strongly on the component labeled “current issues”, which had the most explanatory power in the principal components analysis. The author of tweets posted during non-Trump hours focused less on current issues and wrote at greater length in less pleasant terms. Because of the weak effect sizes, there is no guarantee, in these results, of different authorship: it is possible that Trump’s style varied across the hours of the day, as has been previously suggested (Whissell, 2018b).

It was decided to examine the language of all the tweets in a search for famous Trump-isms. The existence of these oft-repeated phrases has led several analysts of the tweets to conclude that Trump has a very distinct tweeting style

(<https://filiph.github.io/markov>, <https://medium.com/swlh/analyzing-trumps-tweets-5368528d2c90s>).

In a data set of 36,631 tokens and 4181 types, the phrase “lamestream media” occurred 15 times, and the phrases “fake news media” or “fake media” 15 times. The word “hoax” appeared 63 times. Democrats were characterized as “do nothing” 36 times, and as leftist or socialist 7 times. The impeachment was referred to as “totally partisan” 6 times, and characterized by other adjectives

likely intended as derogatory (phony, ridiculous, partisan, corrupt, their, and Pelosi's) 39 times. The word "great" appeared 330 times, and members of the family "America" "American", "Americans", and "America's" 157 times. Trump talked a lot about himself: "I" was used 330 times, "me" 57 times, and "my" 123 times, although, to be fair, "you" was employed 365 times. Here is a fictional tweet peppered with Trumpisms of the types exemplified in these results:

"The impeachment is a hoax. The do nothing Democrats and their corrupt lamestream media are to blame! Only I can make America great again and I am already doing a perfect job."

A Mid-Summer Update

The COVID-19 crisis continued in the US well into mid-summer of 2020 with deaths of close to 1000 people a day being reported, unemployment reaching record heights, and Trump losing ground in terms of approval ratings to his 2020 democratic opponent Joe Biden. Because of the volatility of the situation, it was decided to take an additional sample of tweets posted by the president. Eleven days of tweets (25th of July to 4th of August, inclusive) were sampled, in the manner described in the Method section of this article. The 151 tweets were evaluated in terms of their pleasantness using the Dictionary of Affect. The average pleasantness of tweet content had dropped to 48.77. This value indicates significantly less emotional positivity than had been present earlier in the year ($t_{36,298} = 10.44, p < .001, d \approx .20$) and less positivity than the everyday English norm of 50 ($t_{3,562} = 3.32, p < .01$). Postings occurred at roughly the same rate as previously (13.72 per day) and were of roughly the same length (27 words).

A search for the negatively toned words in the mid-summer tweets led to the discovery that "fake" was mentioned 15 times and that words such as "no" (12 times), "not" (19), and "nothing" (4) also appeared relatively often. All these words have low pleasantness scores in the Dictionary, as do "bad" or "badly" (which appeared 10 times) and "terrible" (2). Biden was mentioned 16 times and democrats were named 8 times. It is almost certain that Trump, in his mid-summer tweets, was responding to the ongoing negative events that surrounded him.

CONCLUSIONS AND DISCUSSION

Several conclusions were reached on the basis of the examination of Trump's tweets from early 2020 and the update sample from mid-summer. These are summarized here in the order in which they were treated within the text.

1. Trump's early 2020 tweets, though still positive in tone, were less positive than those posted in 2017.
2. 2020 tweets were somewhat longer and much, much more frequent than those from 2017.
3. As was the case in 2017, the popularity of the 2020 tweets was negatively related to their positivity or pleasantness. Less pleasant tweets were retweeted and liked more often.
4. A principal components analysis of tweet criteria yielded five components: current issues, long tirades, heavy traffic/strong markets, and the use of the words "great" and "we".
5. The pleasantness of tweets was successfully predicted on the basis of all five components.
6. Popularity, represented by both retweeting and liking, was also successfully predicted by a combination of components.
7. Trump's tweets did change – slightly – in conjunction with the fall in the markets evidenced by the Dow Jones Industrial Average and the COVID-19 crisis.

8. Trump's tweets posted in the evenings, nights, and mornings, were different – but not extensively different - from those posted during the working hours of the day, possibly by a person other than Trump himself.
9. Multiple examples of famous Trumpisms such as “lamestream media” and “impeachment hoax” were identified in the tweets; these would allow any interested party to manufacture Trump-like tweets.
10. Possibly as the result of the many political and economic crises faced by Trump, there was a significant drop in the pleasantness or positivity in the language of his mid-summer tweets.

President Trump tweeted at a very high rate in the first third of 2020. An early attempt to describe his tweeting behavior in terms of “tweetstorms” (Cassino, 2017) used the criterion of 7 messages in one day for several successive days as a “storm”. Even when we ignore retweets, Trump's current tweet production sits at more than 13 messages a day. When retweets are included the rate rises to almost 30: the storm has been exceeded by the norm. Such extreme behavior leads one to question the dynamics behind the president's tweeting.

Is Donald Trump an impulsive tweeter who does himself more harm than good with his barrages of tweets or is he a savvy politician manipulating the twitter-verse as part of a well-planned, highly effective, power-grab? Kreis (2017) proposes that Trump uses tweets “strategically” to achieve his aims and Ross and Rivers (2018) maintain that Trump's tweets are engines of intentional “mis- and dis- information” rather than being random ramblings. On the other side of the issue are analysts who point out that Trump's tweet storms are actually damaging to his popularity rating (Cassino, 2017) and those who liken Trump to a toddler with little impulse control (Drezner, 2019).

It is difficult to hold on to two such opposing viewpoints simultaneously - one which presupposes an almost Machiavellian intellect and the other which assumes a serious lack of impulse control. Perhaps at this point it would be best to adopt, instead, a behaviorist's stance in answering the question of why the president tweets as often as he does. Behaviorists look for explanations of behavior in the environment rather than in the mind of the behavior. In fact, the mind is considered a “black box” and any explanation referring to it is dismissed by behaviorists as subjective.

A behaviorist would search for the reasons for behavior rates in the likely reinforcers to be found in the environment. By definition, a reinforcer is any condition or stimulus which is associated with a change in the rate of the behavior it follows (McLeod, 2020). If attention were the reinforcer for Trump's tweeting behavior, and especially if attention were defined in terms of both laudatory and critical responses, it could serve as the main reinforcer of tweeting behavior. In this situation “attention” can be defined in any number of objective ways – the number of selections of a tweet as a favorite, the number of retweets, headline inches devoted to the president's tweeting, minutes of national comedians' commentaries devoted to it, approval rating, or some combination of these.

Another topic of continuing interest is the presence of positive language in Trump's tweets. As noted previously (Whissell, 2018b) the Trump tweet corpus contains congratulatory tweets (mentioning winning teams or the work of approved groups and individuals) as well as condemnatory ones (naming and shaming of media outlets such as CNN or disapproved groups and individuals). Those of the first type tend to have positive language, those of the second negative. The “carrot and stick” approach seems to have been carried through to 2020. However, positive language is not an

immutable characteristic of Trump's tweets. The mid-summer update indicated a significant drop in tweet positivity, one considerably larger than the drop between 2017 and early 2020. By August 2020, Trump's tweets could no longer be considered emotionally positive in tone in comparison to everyday English.

Little in the twitter-verse is trustworthy. The author of a tweet may not be the stated author, and members of the tweet's audience might not be real people but rather programs engaged in retweeting messages to ensure their popularity ratings. In spite of these caveats, the data are the data. In this article the data have been described and analyzed in full awareness of the issues underlying author-ship and audience-ship.

References

- Cassino, D. (2017) President Trump's approval ratings are being driven down by his 'tweetstorms'. 2017-6-13, blogs.lse.ac.uk. <https://blogs.lse.ac.uk/usappblog/2017/06/13/president-trumps-approval-ratings-are-being-driven-down-by-his-tweetstorms/>
- Drezner, D. W. (2020) The toddler in chief: what Donald Trump teaches us about the modern presidency. Univ. of Chicago Press.
- Feinberg, A. (2017) How to tell when someone else tweets from @realDonaldTrump. Wired: Security 18-06-2017. <https://www.wired.com/story/tell-when-someone-else-tweets-from-realdonaldtrump/>
- Kreis, R. (2017) Trump used tweets as a "strategic instrument" of power politics. *Journal of Language and Politics*, 16, 607-618.
- McLeod, S. (updated, 2018) Skinner – operant conditioning. Accessed May 26, 2020. <https://www.simplypsychology.org/operant-conditioning.html>
- Ross, A. S. & Rivers, D. J. (2018) Discursive dissection: accusation of "fake news" and the spread of mis- and dis-information in the tweets of president Trump. *Media + Society*, April-June, 1-12.
- Whissell, C. (2018a) The medium may be the same but the message is different: comparing the tweets of U.S. Presidents Obama and Trump. *Advances in Social Sciences Research Journal*, vol. 5, issue 6.
- Whissell, C. (2018b) Trump tweets. *International Journal on Studies in English Language and Literature*, 6, 32-40.
- Whissell, C. (2009) Using the revised dictionary of affect in language to quantify the emotional undertones of samples of natural language. *Psychological Reports*, 105, 509-521.