COVID-19 Tweets of Governors and Health Experts: Deaths, Masks, and the Economy

Tim Hua[1], Chris Chankyo Kim[2], Zihan Zhang[3], and Alex Lyford[1]

[1]Middlebury College, Middlebury, VT, USA
[2]Stanford University, Stanford, CA, USA
[3]University of Chicago, Chicago, IL, USA
DOI: https://doi.org/10.47611/jsr.v10i1.1171

ABSTRACT

As COVID-19 spread throughout the United States, governors and health experts (HEs) received a surge in followers on Twitter. This paper seeks to investigate how HEs, Democratic governors, and Republican governors discuss COVID-19 on Twitter. Tweets dating from January 1st, 2020 to October 18th, 2020 from official accounts of all fifty governors and 46 prominent U.S.-based HEs were scraped using python package Twint (N = 192,403) and analyzed using a custom-built wordcount program (Twintproject, 2020). The most significant finding is that in 2020, Democratic governors mentioned death at 4.03 times the rate of Republican governors in their COVID-19 tweets. In 2019, Democratic governors still mentioned death at twice the rate of Republicans. We believe we have substantial evidence that Republican governors are less comfortable talking about death than their Democratic counterparts.
We also found that Democratic governor's tweet about masks, stay-at-home measures, and solutions more often than Republicans. After controlling for state-level variations in COVID-19 data, our regression model confirms that party affiliation is still correlated with the prevalence of tweets in these three categories. However, there isn't a large difference between the proportion of COVID-19 tweets, tweets about the economy, tweets about vaccines, and tweets containing "science-like" words between governors of the two parties.
HEs tweeted about death and vaccines more than the governors. They also tweeted about solutions and testing at a similar rate compared to governors and mentioned lockdowns, the economy, and masks less frequently.

Introduction

The COVID-19 pandemic has shaped our lives and was the defining event of 2020. In the United States, state governments became central to the pandemic effort, and governors found themselves in the spotlight. Google Trends data reveals that web and news searches for Governor Jay Inslee of Washington, Governor Andrew Cuomo of New York, and Governor Gavin Newsom of California—three states that witnessed early coronavirus outbreaks—surged significantly since early March (Google Trends, n.d.). This popularity spike is also reflected in the number of followers these governors gained on Twitter. For example, Gov. Inslee of Washington added 663 followers in February and 82,594 in March, Gov. Newsom added 2,137 followers in February and 101,401 in March, and Gov. Cuomo added 1,472 followers in February and 633,652 in March (Social Blade, n.d.).

Health experts (HEs) saw similar increases in popularity on Twitter. Individuals such as Robert Redfield, Scott Gottlieb, and Eric Feigl-Ding all saw their Twitter accounts surge in popularity as the pandemic broke out (Social Blade, n.d.). This trend has not eluded the media (Wells, 2020). In fact, there exist multiple online guides recommending readers whom to follow on Twitter to learn more about the coronavirus pandemic (e.g., Brown, 2020; Elemental Editors, 2020). The rising popularity of governors and HEs on Twitter warrants a closer look, especially given that

Twitter has seen a record-breaking increase of users between April and June (Wise, 2020). The dissemination of COVID-19 related information is essential to an effective pandemic response. This study seeks to examine COVID-19 tweets from HEs and governors—tweets many users would have seen. We focus on the frequency in which these groups invoke the following subjects in their tweets: COVID-19, death, economy, solutions to the pandemic (e.g., mask, testing), and technical language (e.g., test positive rate, aerosols). The full list of keywords is available in Appendix 1.

We scraped tweets from the Twitter accounts of governors and from 46 influential U.S.-based HEs dated from January 1st, 2020 to October 18th, 2020 using the Twint API (N=192,845) (Twintproject, 2020). We marked all tweets that contained words relating to COVID-19 (N = 90,169) and deleted the non-COVID-19 related tweets. We then looked at the prevalence rates of keywords—such as words related to death or the economy—within the subset of COVID-19 tweets. We also looked at all tweets mentioning solutions to the pandemic, and the prevalence rates of words relating to different solutions (e.g. testing, masks, lockdown) within this subset of solution tweets. The full list of HEs is available in Appendix 2.

We found several differences between Democratic governors, Republican governors, and HEs. Democratic governors mentioned COVID in 42.29% of their tweets and Republicans in 41.48%. With respect to death, HEs mentioned it in 12.14% of their COVID tweets, Democrats in 6.63%, and Republicans in 1.65%. With respect to the economy, HEs mentioned it in 1.85% of their COVID Tweets, Democrats at 8.31%, and Republicans at 8.68%. With respect to solutions to the pandemic, HEs mentioned them in 41.17% of their COVID-19 tweets, Democrats in 49.83%, and Republicans in 42.18 %. Within the solutions tweets subset, HEs mentioned masks 21.59% of the time, Democrats 26.85%, and Republicans 22.77%. After controlling for three key state-level COVID data—deaths per capita, cases per capita, and tests per capita—party affiliation is still more strongly correlated with of how a governor tweets compared to the control variables when it comes to death, solutions, masks, and stay-at-home measures related tweets.
We were intrigued by the massive difference between Democrats and Republicans in the rate at which they mentioned death related words (die, dead, dying, death, coffin) and decided to see if this difference existed in 2019. We examined the ratio of total death tweets to total tweets from governors from January 15th, 2019 to October 18th, 2019 and compared it to 2020. In 2019, Democratic governors mentioned death in 1.82% of their tweets, which increased to 4.42% in 2020. For Republicans, the increase was from 0.91% to 1.59%. Republican governors are talking about death in their tweets in 2020 less often than Democrats in 2019, before the pandemic. One possible explanation for this phenomenon is moral foundations theory (Graham et al., 2013), which claims that liberals focus on the harm/care morality more than conservatives.

Our study contributes to the literature surrounding healthcare and political communications in the context of the pandemic and verifies the existing claim that Republicans tend to mention the human consequence of the pandemic less (e.g., Sullivan, 2020; Gittleson, 2020). With support from the 2019 tweet data, it also contributes to the political psycholinguistics literature which examines the differences in how liberals and conservatives talk (e.g., Sterling et al., 2020). Finally, we contribute to the emerging literature that examines healthcare communications during COVID-19 and how health care professionals address the pandemic on Twitter (e.g., Boğan et al., 2020).

The paper is organized in the following manner. We will first present additional background information and literature related to our study. We will then describe in detail how we obtained and categorized the tweets and the information we extracted. We conclude by providing some commentary on the findings, address the limitations of this study, and point to future topics that researchers should explore.

Literature Review

As early as February 2020, editors of The Lancet concluded that “There may be no way to prevent a COVID-19 pandemic in this globalized [sic] time, but verified information is the most effective prevention against the disease of panic.” Indeed, a different editorial in Patient Education and Counseling stressed that “Political leaders and health experts have a special responsibility to provide us with accurate information” (Finset et al., 2020). There are many guides to communications during a pandemic (such as a dedicated chapter in CDC Field Epidemiology Manual, or the article by Vraga and Jacobsen (2020), who illustrated three key challenges in pandemic communication), but normative discussions are not the focus of this study (Trumpey et al., 2018). We do not offer advice on how to tweet about COVID-19 in this paper and seek only to investigate and describe the messages that are currently sent by governors and health experts. To our knowledge, there currently exist two studies looking at governors' COVID-19 messaging. Grossman et al. (2020) specifically examined the effects of stay-at-home orders and messaging and found that they resulted in significant mobility reductions, with larger reductions from Democratic states and counties. Sha et al. (2020) used dynamic topic modeling to examine the topics mentioned in the tweets of U.S. governors and cabinet officials. Studies on COVID-19 tweets by HEs are rarer. One study by Boğan et al. (2020) examined 251 Twitter accounts by emergency medicine physicians and residents in Turkey and classified their tweets into categories such as comments and suggestions (42.4%) and institutional announcements (18.6%).

Outside of COVID-19, there is sparse research on Twitter as a healthcare communications venue for HEs. A few studies focus on tweeting during professional conferences (Salzmann-Erikson, 2017; Lemay et al., 2019; Ziemba et al., 2020), and one study that examined the #TipsForNewDocs hashtag (Rashid et al., 2018). Research on the differences in Republican and Democrat politicians are more plentiful. Neiman et al (2015) analyzed the language used by U.S. federal-level politicians in speeches, press appearances, and presidential debates. They found that political elites do not systematically differ when it comes to value-related language. Although our analysis focuses on politicians who are explicitly related to a party, we believe that literature that examines the difference between language used by conservatives and liberals also add context to our findings. For example, Robinson et al.(2015) examined language use on conservative and liberal websites, and Brundidge et al. (2014) looked at conservative and liberal blogs. There also exist several studies on how partisans use Twitter. A comprehensive study on tweets can be found at Sterling et al. (2020), who tested 27 hypotheses on how the "linguistic styles" of liberals and conservatives differ. The ideology of a user is calculated based on who they follow. They used tweets from 25,000 users and found a significant difference in 23 cases. They also found that political extremists use language differently compared to moderates in 17 different dimensions. Finally, Sylwester and Purver (2015) found seven linguistic differences between Twitter users who exclusively follow @GOP and those who exclusively follow @TheDemocrats, the two parties' official Twitter accounts. They found that Democrats are more likely to use swear words, feeling-related words, positive sentiment, and first-person singular pronouns, whereas Republicans are more likely to use religion-related words and first-person plural pronouns.

Due to the novel nature of COVID-19, there are large gaps in our knowledge about its portrayal in the media. This paper is intended to provide a starting point for future research on health and political communication during the COVID-19 pandemic and answer a few basic questions. We chose to focus on communication on Twitter for several reasons. Twitter seems to be the major platform for receiving and distributing political information due to its low complexity, making it the preferred platform over other media platforms such as Facebook (Parmelee & Bichard, 2011). As previously stated, Twitter gained a record-breaking number of accounts during the pandemic, and many users use the site to follow health experts and governors. Finally, the availability of public APIs such as Twint makes it very easy for researchers to study tweets (Twintproject, 2020). Twint is an opensource tweet scraping tool that allows researchers to download unlimited amounts of tweets without rate limitations that come with many other APIs.

We start by first looking at how often the groups and individual users mention the pandemic, which we believe to be a good indicator of the importance one puts on the issue and how comfortable a person is with the subject. We grouped the governors based on party not only because of convention but also due to how politics is driving the COVID-19 response. For example, the decision on whether or not to reopen public schools in the fall is correlated with Trump's vote share in 2016 in the county (Hartney & Finger, 2020). Indeed, all else equal, Republican governors were slower (by 2.04 days) to implement social distancing measures at the beginning of the pandemic (Adolph et al., 2020). The divide exists outside of government at the individual level as well. Democrats are more likely to avoid public spaces and social gatherings compared to Republicans (Allcott et al., 2020). Next, we examine how these users talk about the pandemic's consequences and solutions, what we believe to be the two central themes surrounding coronavirus messaging. For consequences, we focused on the differences between words relating to the economy and words relating to deaths since there is an ongoing debate on COVID-19 restrictions and its effects on the economy (Samuels & Klar, 2020; Holland & Hunnicutt, 2020). For solutions, we looked at the prevalence rate of a range of words relating to different solutions (masks, testing, vaccine, etc.) in the tweets of governors and HEs and the rate at which these groups mention any solutions in their tweets.

We start by first looking at how often the groups and individual users mention the pandemic, which we believe to be a good indicator of the importance one puts on the issue and how comfortable a person is with the subject. We grouped the governors based on party not only because of convention but also due to how politics is driving the COVID-19 response. For example, the decision on whether or not to reopen public schools in the fall is correlated with Trump's vote share in 2016 in the county (Hartney & Finger, 2020). Indeed, all else equal, Republican governors were slower (by 2.04 days) to implement social distancing measures at the beginning of the pandemic (Adolph et al., 2020). The divide exists outside of government at the individual level as well. Democrats are more likely to avoid public spaces and social gatherings compared to Republicans (Allcott et al., 2020). Next, we examine how these users talk about the pandemic's consequences and solutions, what we believe to be the two central themes surrounding coronavirus messaging. For consequences, we focused on the differences between words relating to the economy and words relating to deaths since there is an ongoing debate on COVID-19 restrictions and its effects on the economy (Samuels & Klar, 2020; Holland & Hunnicutt, 2020). For solutions, we looked at the prevalence rate of a range of words relating to different solutions (masks, testing, vaccine, etc.) in the tweets of governors and HEs and the rate at which these groups mention any solutions in their tweets.

We also wanted to examine the degree to which governors mimicked the language of HEs. We believe that this can be seen as a measure of how much a governor wishes to be associated with the scientific community, given that a subset of conservatives has little trust in science and have lower social distancing intentions (Koetke et al., 2020). We attempt to measure this by looking at the prevalence of technical language (e.g., "R0," "asymptomatic," "aerosol") in tweets. Finally, we want to see if the differences between governors still exist after accounting for state-level variations in COVID-19 data.

To reiterate, we seek to answer the following questions:

RQ1: Do Democratic and Republican governors differ in how often they tweet about COVID-19?

RQ2: Do Democratic governors, Republican governors, and health experts (henceforth referred together as DRHs) differ when talking about the consequences of the pandemic?

RQ3: Do DRHs differ when talking about solutions to the pandemic?

RQ4: Do DRHs differ in the prevalence of technical language in their COVID-19 related tweets?

RQ5: If there are differences between Democrat and Republican governors in the aforementioned research questions, are the differences more due to state-level COVID-19 data, or is party affiliation the driving force behind the differences in how members of the two parties' tweet?

Methodology

Selecting Twitter Accounts

We first found the most followed Twitter account of each governor. Some governors have multiple accounts for personal, campaign, and official uses. For example, the official account of Pennsylvania's governor's office is @GovernorsOffice, but it only has around 48,784 followers compared to @GovernorTomWolf with 268,713. We believe that it is more appropriate to select the latter account, since it is most likely that more Pennsylvanians get their information from @GovernorTomWolf as opposed to @GovernorsOffice. (All follower numbers were recorded on February 22nd, 2021.)

In this paper, health experts refer to users who hold the following qualifications: Doctor of Medicine (MD), Master of Public Health (MPH), a doctorate or professorship in epidemiology, virology, biology or related field. We chose these criteria because we are looking for users who are seen as experts and a voice from the medical field. We are not concerned with the current professions of the HEs as long as they are a source of COVID-19 information for their followers.

We took a convenience sample of HEs tweeting about COVID-19 by first starting with eight arbitrarily chosen HEs (see "Initial HEs" list in appendix 2) who are publicly known to tweet about COVID-19 and have over 100,000 followers. We then looked at everyone followed by these eight users and filtered for those who had specific health related keywords in their username, Twitter handle, or bio (see appendix 2 for full list of keywords). We only included users with at least 20,000 followers. This first round resulted in 341 users. We then downloaded all tweets sent by these users between April 1st and April 30th—the height of the pandemic. We excluded all accounts (n=126) without at least 30 COVID-19 related tweets (defined as having a word from the COVID-19 keywords dictionary in Appendix 1) and accounts who do not tweet about COVID-19 in at least 33% of their tweets (n= 105). Finally, we manually went through to exclude organizations and checked that all individuals are based in the US and hold the qualifications listed above. This resulted in a final list of 46 health experts, which can be found in Appendix 2.

We believe that a user who does not tweet about COVID-19 at least once a day and has 20,000 followers does not reach enough users with their COVID-19 tweets to count as influential and merit our attention. We excluded those who do not tweet about COVID-19 at least 33% of the time because those users might not be seen as health experts or followed because of their credentials. Despite the possible setbacks of using a convenience sample, we believe that this search method is the best one available. If we had followed a predetermined list of Twitter accounts such as the one posted by Elemental Editors (2020), we would be subject to the bias of the curators of that list. Selecting a few HEs and searching through all of their followers allows us to get a large enough sample to cover a substantial amount of the population, while at the same time be a small enough list that makes it feasible to manually examine each account to verify the user's credentials.

Obtaining Tweets and Conducting Analysis

We downloaded all state governors' and HEs' tweets from January 1st, 2020 to October 18th, 2020 using the Twint Python API (Twintproject, 2020). This totaled 56,110 tweets from governors and 136,293 tweets from HEs.

We conducted a content analysis using a wordcount program written in Python. The program reads a dictionary and checks if a word from that dictionary is present in a tweet. It makes no distinction between tweets that mention a word once or multiple times. We then looked at the number of tweets that contained words from a specific dictionary for that user. We opted for this binary method instead of a percentage score for each tweet—which would be similar to the LIWC algorithm by Pennebaker et al. (2015)—due to the 280-character limit for tweets. This binary count is also used in other tweet-oriented text analysis papers such as Sterling et al. (2020).

We used custom dictionaries in Appendix 1 in conjunction with our wordcount program to answer our research questions. Unless otherwise stated, the following analysis will utilize prevalence rates as opposed to the absolute number of tweets. Here, prevalence rates refer to the proportion of tweets that contain at least one word from a given dictionary—for example, what proportion of COVID-19 related tweets contain at least one word from the Solutions dictionary. When we say, "prevalence rate of y with respect to x", we meant the following percentage: $\frac{y}{x}$.

For RQ1, we looked at the prevalence rate of tweets that contain words from the "COVID-19 Keywords" dictionary with respect to the total number of tweets. For RQ2, we removed all non-COVID-19 tweets and looked at the prevalence rate of tweets that contained words relating to the economy with respect to the total number of COVID-19 tweets as well as the prevalence rate of tweets containing words relating to deaths with respect to the total number of COVID-19 tweets using the "Consequences Keywords" dictionary. For RQ3, we measured the prevalence rate of tweets that contained words from the "Solutions" dictionary (a subset of "COVID-19 Keywords") with respect to total COVID-19 tweets. We also measured the prevalence rate of tweets containing individual solutions (e.g. "mask," "vaccine") with respect to all solutions-related tweets. For RQ4, we measured the prevalence rate of tweets that contained words from the "Technical Language" dictionary (a subset of "COVID-19 Keywords") with respect to total COVID-19 tweets. All of the prevalence rates were calculated for each individual governor, the two parties, and HEs.

Since we collected all tweets from all governors between January 1st and October 18th, our data represent a census. We also obtained nearly all tweets from HEs who fit our criteria (20,000 followers, U.S.-based, and tweets about COVID-19 at least 33% of the time). Thus, this sample is sufficiently large and nonrandom such that we chose to treat it as a census. Given our census of data, statistical hypothesis tests—such as t-tests and ANOVA—are not appropriate for our analysis. Hypothesis testing is useful when we collect sample statistics (such as sample means or sample proportions) to estimate a given population parameter. In our case, we have calculated the exact parameter values for the population. Thus, all differences are statistically significant—it is up to the reader to determine if the given differences are practically significant.

To answer RQ5, we built linear regression models that try to account for the variations between how different governors tweet based on not just their party affiliation, but also cases, tests, and deaths per capita of their state. We obtained the data for total cases and deaths from the CDC and used the U.S. Census Bureau's 2019 population estimate for each state (2020; 2019). We used state level testing data from the COVID-19 tracking project (2020). Case, death, and testing numbers are from Oct 18th, 2020.

Results

Table 1 and Figure 1 display the total number of tweets containing words from each dictionary. Note that non-COVID-19 related tweets are removed from analysis after RQ1. Thus, if a tweet read "Justice Ginsburg died," it would not be counted as a COVID tweet and thus will not be counted in the "Total Death Tweets" column. It is also important to keep in mind that a tweet like "Masks stop aerosols" will be counted in the technical language section, the solutions section, and the masks section.

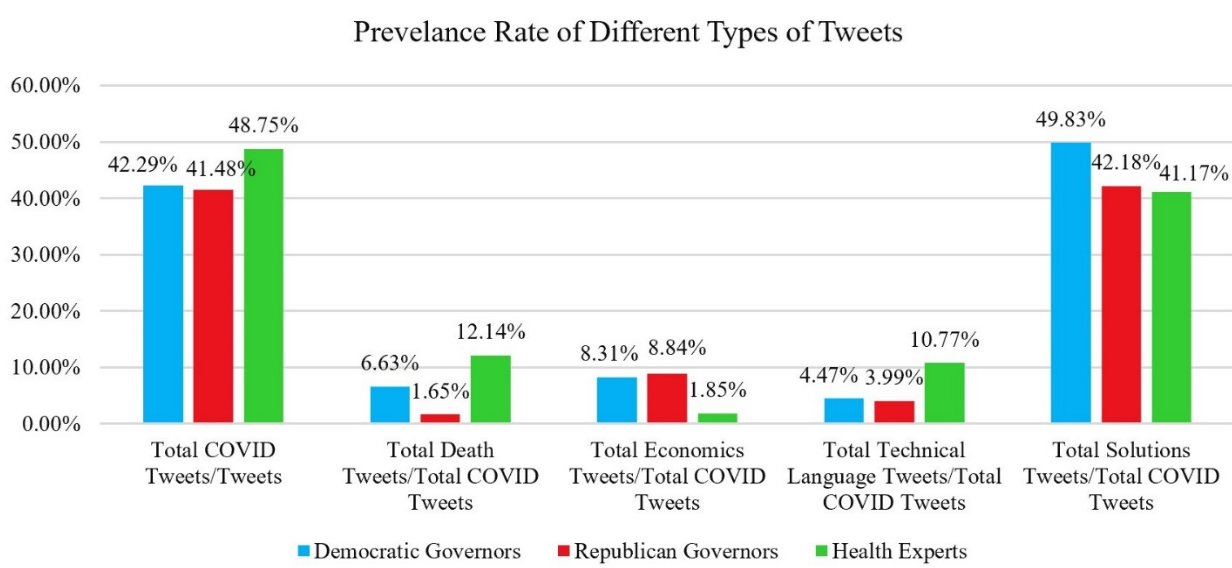

Figure 1. Prevalence rate of different types of tweets.

Table 1. Number of tweets in each category.

	Total COVID Tweets/Total Tweets	Total Death Tweets/Total COVID Tweets	Total Economy Tweets/Total COVID Tweets	Total Technical Language Tweets/Total COVID Tweets	Total Solutions Tweets/Total COVID Tweets
Democratic Governors (n=29,297)	42.29% (n = 12,389)	6.63% (n=821)	8.31% (n=1,029)	4.47% (n=554)	49.83% (n=6,174)
Republican Governors (n=26,813)	41.48% (n= 11,121)	1.65% (n=183)	8.68% (n=983)	4.01% (n=446)	42.18% (n=4,691)
Health Experts (n=136,293)	48.75% (n=66,449)	12.14% (n=8,066)	1.85% (n=1,231)	10.77% (n=7,159)	41.17% (n=27,358)

Figure 1 and Table 1 address the majority of our research questions. For example, we can see that Democratic governors tweet marginally more in absolute numbers and percentage points than Republican governors about COVID-19. We also see that Republican governors talk about the economy marginally more than Democrats and around 4.7 times the rate of HEs. Conversely, Democratic governors mention death in their tweets at more than four times the rate of Republican governors, and HEs mention death at 7.36 times the rate of Republican governors.

Unsurprisingly, HEs have the highest proportion of technical language in their tweets. It appears that Democratic governors use technical language marginally more than Republican governors (0.46% more in absolute terms, or 11.4% more often than Republican governors).

We then examined tweets about solutions to the pandemic. Democratic governors talk about solutions 18% more often than Republicans, who talk about them at a similar rate as HEs. We decided to investigate a few subsets of the solutions dictionary: words relating to testing, words relating to staying at home, “mask,” and “vaccine.” The results of this analysis can be found in Table 2 and Figure 2. Note that the percentages here are derived by dividing the total tweets of that category by the total number of solution tweets (as opposed to the total number of COVID-19 tweets). This approach provides more information about the types of solutions offered by each of the three cohorts.

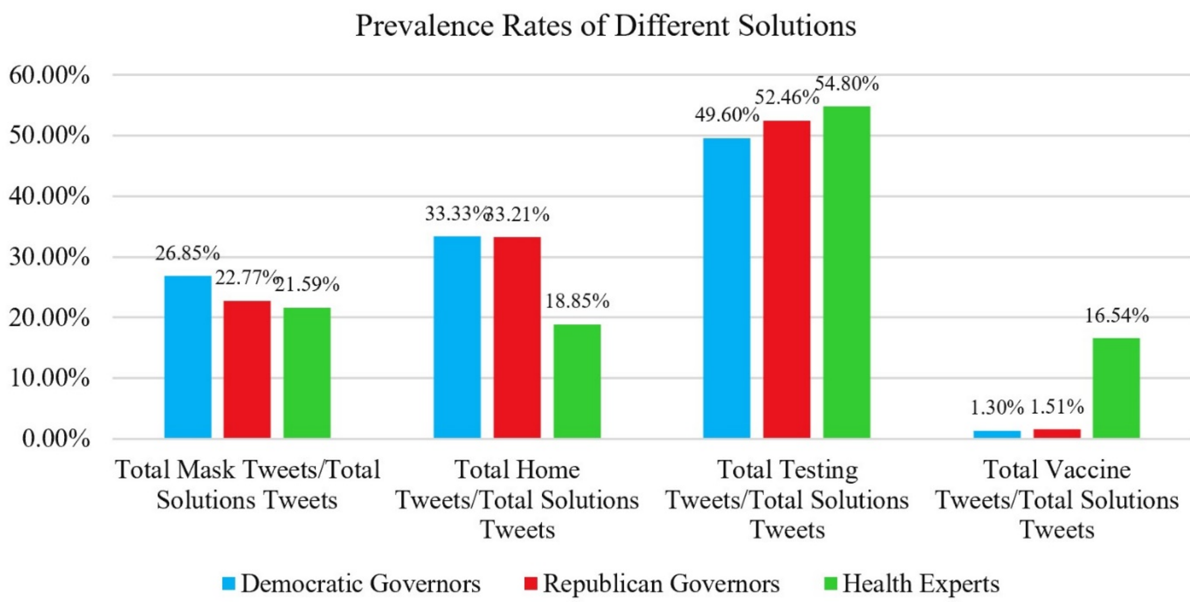

Figure 2. Prevalence rates of different solutions.

Table 2. Prevalence rate of difference solutions.

	Total Mask Tweets/Total Solutions Tweets	Total Home Tweets/Total Solutions Tweets	Total Testing Tweets/Total Solutions Tweets	Total Vaccine Tweets/Total Solutions Tweets
Democratic Governors	26.85% (n=1,658)	33.33% (n=2,058)	49.60% (n=3,062)	1.30% (n=80)
Republican Governors	22.77% (n=1,068)	33.21% (n=1,558)	52.46% (n=2,461)	1.51% (n=71)
Health Experts	21.59% (n=5,906)	18.85% (n=5,156)	54.80% (n=14,993)	16.54% (n=4,526)

There are minor differences between the three groups. Democratic governors talk about masks in their solutions tweets more frequently than Republicans. This gap becomes even larger if we consider the ratio between total mask tweets and total tweets, where Democratic governors (5.66%) mention "mask" in their tweets 42% more often than Republican governors (3.98%). When discussing solutions, Republican governors mention testing more in their tweets than Democrats. HEs mention masks and stay-at-home measures less frequently than governors, but they mention vaccines at a significantly higher rate.

Next, we investigated if the observed differences between Democratic and Republican governors' tweets were explained by their party affiliation or by differences in state-level COVID-19 data. To answer this, we use the following model:

$$Prevalence\ Rate_i = \beta_0 + \beta_1 * P_i + \beta_2 * C_i + \beta_3 * D_i + \beta_4 * T_i + \epsilon_i$$

Where i indexes all fifty states, $P_i \in \{0,1\}$ is the party affiliation of the governor (with $P_{Republican} = 1$), C_i is the number of cases per 100,000 residents, D_i is the number of deaths per 100,000 residents, and T_i is the number of tests per 100,000 residents. The C_i, D_i, and T_i variables serve as state-level controls to evaluate the effects of *only* party affiliation on each of the given prevalence rates.

Since we have a census of tweets, none of the beta values in the regression model are estimates—they are all exact population parameters values. In other words, the standard error for each of these beta values is zero, and any deviation in these betas from 0 is statistically significant. Table 3 shows the results of each of our nine regression models. The β_1 values are colored as a function of their magnitude, with red values representing negative slopes and green values representing positive slopes. We verified homoscedasticity on the residual plots.

Table 3. Regression model outputs.

Response Variable	β_0	β_1	β_2	β_3	β_4	R^2
COVID-19 Tweets	0.400	2.42%	-3.15E-05	2.25E-04	0.116	13%
Death Tweets	0.050	-2.73%	-7.26E-06	1.07E-04	0.024	17%
Econ Tweets	0.138	-0.40%	9.90E-07	-4.52E-04	-0.022	12%
Tech Tweets	0.017	0.54%	-6.99E-06	3.21E-04	0.039	18%
Solution Tweets	0.445	-8.97%	-4.70E-06	6.28E-04	0.042	19%
Home Tweets	0.182	-6.23%	-6.43E-07	-3.83E-04	0.064	16%
Test Tweets	0.154	1.10%	-9.30E-06	7.87E-04	0.090	18%
Mask Tweets	0.150	-5.80%	9.32E-06	7.06E-05	-0.085	17%
Vaccine Tweets	0.009	0.10%	-1.09E-06	1.40E-05	-0.004	3%

None of these models has a substantially large R2 value. In other words, the amount of variation in the prevalence of a given response is only partially explained by the combination of party affiliation and state-level controls is typically around 15–20%. For some categories, party affiliation was strongly correlated with tweet prevalence. For example, the party-affiliation effect for Solution Tweets was -8.97%. After controlling for state-level variables, the prevalence rate of tweets about solutions to the pandemic was 8.97% lower for Republican governors than for Democratic governors. Other large differences included tweets about staying at home and tweets about masks. It's important to note that these differences are the absolute difference in prevalence rates and not the proportional difference in prevalence rates.

By examining the effect on R^2 by removing an explanatory variable, we can find out which of the variables has the strongest effect on the ability of the model to account for differences in prevalence rates. For example, if we removed party affiliation from the model predicting the prevalence rate of COVID-19 tweets, the R^2 value drops from 13% to 12%. However, if we removed cases per capita, it drops by 8% to 5%. (The drops for removing deaths per capita and tests per capita are 0% and 4% respectably). From this, we can see that cases per capita of a state is more important than party affiliation in determining how often a governor mentions COVID-19 in their tweets. In fact, party affiliation is only the most important factor for the rates of tweets mentioning death, solutions, stay-at home measures, and masks. This is not to say that party affiliation does not have an effect elsewhere—as long as $\beta_1 \neq 0$, party affiliation has a statistically significant effect on the prevalence rate of a given tweet. It just might not be as significant as the effect of other variables.

Further Analysis on Consequences Keywords

After seeing the large difference between how often Democratic governors and Republican governors mentioned death in their tweets, we were inspired to examine if this is a larger trend: Do Democratic governors talk more about death only in the context of COVID-19, or do Democratic governors talk more about deaths in general than Republican governors?

We obtained all tweets from governors' official accounts from January 15th, 2019, to October 18th, 2019 (N=32,279). We started on January 15th because many governors were inaugurated on January 14th. We also had to use different governor Twitter accounts for Kentucky (Republican) and Mississippi (Republican) since a different politician was in office in 2019. Another note is that Governor Bill Lee (Republican) of Tennessee was inaugurated on January 19th, 2019 even though we scrapped his tweets starting from January 15th. We once again used the death keywords from the consequences dictionary in conjunction with our wordcount program. The results are displayed in Table 4.

Table 3. Governors' Tweets in 2019

	Total Tweets	Total Death Tweets	Total Death Tweets/ Total Tweets
Democratic Governors	17,019	310	1.82%
Republican Governors	15,260	139	0.91%

Democratic Governors mentioned death twice the rate of Republicans, which is less than the difference between the two groups when it came to Total Death Tweets/ Total COVID-19 Tweets in 2020, where Democratic governors mentioned death at 4.03 the rate of Republican governors. For economy related tweets, Republicans talk about the economy 15.9% more than Democrats in 2019. In 2019 COVID-19 tweets, Republican governors talk about the economy 4.4% more than Democrats. It may appear that both Democratic and Republican governors talked about the economy at a higher percentage in 2019 than in 2020. However, this is because we removed non-COVID-19 related tweets before analyzing for mentions of economy and death in the 2020 dataset. When we reran the program on the entire 2020 tweet dataset for governors (N = 56,110) without first limiting our analysis to COVID-19 tweets, we get the results in Table 5.

Table 4. Governors' Tweets in 2020

	Total Tweets	Total Death Tweets	Total Death Tweets/ Total Tweets
Democratic Governors	29,297	1,294	4.42%
Republican Governors	26,813	427	1.59%

Comparing Table 4 and Table 5, we can see that the total number of tweets increased significantly between 2019 and 2020 for both parties (72% for Democrats and 76% for Republicans). However, the proportion of tweets that are about the economy barely changed. The proportion of death related tweets increased 74% for Republican governors and 143% for Democratic governors. However, note that the prevalence rate of death in 2020 for Republican governors is still lower than the prevalence rate of death for Democratic governors in 2019, before the pandemic.

The full dataset, including state-level COVID-19 data, data on tweets for individual governors, and all tweets scraped with Twint, can be found at https://github.com/zzhang-18/Tweet-Analysis

Discussion

This study provides a broad overview of how health experts (HEs), Democratic governors, and Republican governors tweet about COVID-19.

Starting with governors, we found that Democratic and Republican governors do not differ substantially in the rate at which they mention COVID-19 in their tweets, the rate at which they mention technical language in their COVID-19 tweets, and the rate at which they mention testing, vaccines, and stay-at-home measures in their tweets about solutions to the pandemic. We believe that the technical language dictionary is a good proxy of how "science-like" a group sounds since HEs use them at a much higher rate. Thus, we claim that Republican governors and Democratic governors are about the same in terms of how science-like they speak given their similar technical language usage rates.

We find substantial differences between the prevalence rates at which governors mention masks, stay-at-home measures, and solutions with respect to COVID-19 tweets. Democratic governors talk about these three subjects more than Republican governors, and our regression model confirms that in these three cases party affiliation is more correlated with of how a governor tweets than state-level COVID-19 data. This supports media reports of the politicization of mask-wearing in the United States (Aratani, 2020). One interpretation regarding the differences in solution tweets is that Republican governors might be more inclined to think that state governments could not do anything more against COVID-19, thus there is less need to talk about solutions to COVID-19 (Falconer, 2020; Marley, 2020). The weakness of this interpretation is that it would also imply that Republican governors will tweet about COVID-19 less frequently than Democrats. In reality, Democratic governors only mention COVID-19 in their tweets 1.7% more than Republicans, and party affiliation is not as a strongly correlated with how often governors mention COVID-19 as cases per capita in a state.

Finally, we investigated how each of the three groups talked about consequences of the pandemic. When examining COVID-19 tweets, Republican governors mention the economy 4.4% more than Democratic governors, who in turn mention death 303% more than Republican governors—the most substantial difference in the entire study. Party affiliation is the strongest predictor of difference in death tweets after controlling for state-level COVID data. Our data on COVID-19 tweets support the assertion that Republican governors tend to be more focused on the economic impact of COVID-19 (e.g. Nagourney & Peters, 2020) compared to the human impact of the pandemic (e.g., Sullivan, 2020; Gittleson, 2020) when compared to Democratic governors. With the data from 2019, however, we claim that this difference in death mentions could be reflective of deeper disparities between Democratic and Republican politicians. Our interpretation is that Democratic governors have always been more comfortable talking about death than Republican governors, who in turn like to focus on the economy. As a result of the Republican rhetoric surrounding COVID-19 and the economy, Democratic governors responded by mentioning the economy more in their COVID-19 tweets to match the rhetoric of Republican governors. However, although Republican governors talked about death at a higher rate than they did in 2019, their rate of Total Death Tweets (in all their tweets, not just the COVID-19 tweets) to Total Tweets ratio in 2020 is still lower than the Total Death Tweets to Total Tweets ratio of Democratic governors in 2019, before the pandemic. We believe that we have substantial evidence that Republican governors are more unwilling to talk about death in general compared to Democrats, a trend that extends to COVID-19. One possible explanation for this unwillingness is the moral foundation theory (Graham et al., 2013), which claims that liberals tend to focus on the moral foundations of Harm/Care and Fairness/Cheating more than conservatives, and the majority of Republicans are conservatives and Democrats liberals (Saad, 2021).

We included HEs for three reasons: to provide a comparison group to the governors, to see if governors are trying to mimic the language of HEs, and to investigate them as a group in their own right. Although we expected that the HEs would mention deaths and technical language more than the governors and less about the economy, other findings are unexpected. HEs mentioned solutions to the pandemic less frequently than governors. They also mentioned masks and stay at home measures less, while tweeting about vaccines at over ten times the rate of governors when we consider vaccine tweets/ total COVID-19 tweets. A possible explanation is that HEs mention masks and stay-at-home measures less since these are considered "obvious" solutions to the pandemic, and where they believe their expertise is needed is with regards to more complicated subjects such as vaccines. HEs also do not need to say the same things repeatedly to reinforce their message to constituents, which politicians need to do. HEs could also talk about vaccines more because they consider that to be a more medical solution as opposed to policy solutions such as stay-at-home measures. Finally, HEs could be following the politicians' rhetoric and selectively tweeting about subjects that they believe are not sufficiently covered by politicians. Nevertheless, our results provide basic information on how HEs tweet about COVID-19 on Twitter compared to Governors.

We believe that the most significant finding of this study is that Republican governors have consistently talked about death significantly less than Democratic governors, both in 2019 and in 2020. We believe that this death aversion from Republican governors could have implications beyond COVID-19 in policymaking and other arenas.

Limitations and Future Directions

We urge future researchers to investigate if the difference in death mentions is a phenomenon that extends to conservative and liberal politicians or media outlets in general and if any verifiable consequences result from this behavior. Another interesting question is whether death mentions is a proxy for death anxiety and mortality salience between conservatives and liberals in general, and if that a driver behind the different degrees of social distancing behaviors. Another way to investigate death anxiety is to conduct cross cultural analysis. For example, do Swedes, who are known for their lenient COVID-19 rules, have less death anxiety compared to their Norwegian counterparts, who have much stricter COVID-19 measures (Helsingen et al., 2020). Finally, researchers could examine if there is a way to manipulate death anxiety to increase compliance with social distancing measures.

A central limitation of the study is the word count methodology that we elected to use. Although analysis with wordcount programs is common and reveal useful information, there are limitations in the type of data they provide. For example, while Democratic governors and Republican governors mentioned stay at home measures at similar rates, we do not know the tweets' context. “Lockdowns need to stop” and “we need harsher lockdown measures” are identical to our analysis program. Furthermore, there are more sophisticated text analysis methods and programs such as Wordstat and machine learning algorithms. We elected to use a word count program for its simplicity and speed, since we believe that it is important to examine COVID-19 communications and report on it while the pandemic is still ongoing to inform policymaking and future research. However, future researchers should take advantage of the wide variety of methodologies available to examine these texts. For example, to measure the linguistic similarity between governors and HPs, we used the technical language dictionary in appendix 1, which we arbitrarily wrote. However, there are machine learning algorithms, such as topic modeling, that allow researchers to examine linguistic similarity in a more systematic fashion.

Another limitation is that we focus our analysis on tweet texts, whereas many governors (such as Governor Eric Holcomb of Indiana) post videos briefings on Twitter. Future research could also examine the content of these videos and other healthcare communication venues used by politicians and HPs. Our analysis of tweet texts is also limited by the tweet scraping package—Twint— used in this study. We chose Twint in response to Twitter recently changing their backend API, which made many existing tweet scrapers obsolete, with one user calming that they “can not complete [their] master thesis” as a result (Cairns & Shetty, 2020; herdemo, 2020). Most tweet scrapers, such as Rtweet, could only scrape the most recent 3,200 tweets from a given user. Moreover, Twitter charges a premium for those who wish to have access to the full archive (Twitter Developer, n.d.), which is why we chose to use Twint—a free API that provides unlimited downloads. Twint provides us with the text of each tweet for regular tweets. For replies and retweets, it provides us with only the reply text. It does not show us the text of the tweet which was replied to or retweeted. For example, if someone tweeted “Masks stop aerosols,” and a governor retweeted their tweet saying, “That’s right,” Twint would show the tweet as “That’s right” by the governor, and not the original “Masks stop aerosols” tweet. However, if the governor retweeted the original tweet without adding a comment, then the tweet will not be captured by Twint. Twint will also give a URL for any image or video content. Thus, photos or videos in a tweet will display as a URL at the end (i.e. “Look at this photo https://t.co/vMwfgtvFQj”) in our dataset.

This paper did not examine any potential correlation between governors’ tweets and the policies they implemented. Thus, questions such as “Do states with mask mandates have governors that talk about masks more often in their tweets?” are left for future researchers to answer.

Finally, our analysis looks at the period from January to October as a homogeneous block, even as cases and deaths varied wildly throughout. Future researchers could seek to conduct a time-sensitive analysis and see how the tweets by different users varied over time and if they are dependent on the location and time specific COVID-19 situation. We also treated HPs as a homogeneous block, even though there are disagreements between HPs (e.g. For information on criticism of Eric Feigl-Ding by Marc Lipsitch, see Bartlett (2020)).

Despite these limitations, we believe that our study adds valuable information to the literature and provides a foundation for future studies. Future researchers should seek to investigate the death aversion of Republican governors. They could also use more advanced tweet scraping methods and text analysis software to learn more about how political elites, influential healthcare professionals, and other opinion leaders talk about COVID-19 on Twitter and beyond.

Acknowledgements

We would like to thank Professor Eric Bleich for his feedback and support throughout the project. We would like to thank Cooper Kelly for help with the R code that filtered for HEs. We would like to thank Professor Julia Berazneva and Professor James Ralph for help revising the paper. Finally, we would like to thank twintproject for making tweets accessible to researchers.

References

Adolph, C., Amano, K., Bang-Jensen, B., Fullman, N., & Wilkerson, J. (2020). Pandemic politics: Timing state-level social distancing responses to COVID-19. *J Health Polit Policy Law* 8802162. https://doi.org/10.1215/03616878-8802162

Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., & Yang, D. (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*, *191*, 104254. https://doi.org/10.1016/j.jpubeco.2020.104254

Aratani, L. (2020, June 29). How did face masks become a political issue in America? *The Guardian.* Retrieved from https://www.theguardian.com/world/2020/jun/29/face-masks-us-politics-coronavirus

Bartlett, T. (2020, April 17). This Harvard epidemiologist is very popular on Twitter. But does he know what he's talking about? *The Chronicle of Higher Education.* Retrieved from https://www.chronicle.com/article/this-harvard-epidemiologist-is-very-popular-on-twitter-but-does-he-know-what-hes-talking-about/

Boğan, M., Karadağ, M., & Boğan, F. (2020). Examination of emergency medicine physicians' and residents' Twitter activities during the first days of the COVID-19 outbreak. *International Journal of Travel Medicine and Global Health*, *8*(2), 46-50. Retrieved from http://www.ijtmgh.com/article_107217_a1d2c7a0269d86b72a61f1af19ff826b.pdf

Brown, A. (2020). Coronavirus: The most essential people to follow on Twitter during the COVID-19 outbreak. *Forbes.* Retrieved from https://www.forbes.com/sites/abrambrown/2020/03/14/coronavirus-the-most-essential-people-on-Twitter-to-follow-during-the-covid-19-outbreak/

Brundidge, J., Reid, S. A., Choi, S., & Muddiman, A. (2014). The "deliberative digital divide:" Opinion leadership and integrative complexity in the US political blogosphere. *Political Psychology*, *35*(6), 741-755. https://doi.org/10.1111/pops.12201

Cairns, I. & Shetty, P. (2020). *Introducing a new and improved Twitter API*. Retrieved from https://blog.Twitter.com/developer/en_us/topics/tools/2020/introducing_new_Twitter_api.html

CDC. (2020). *United States COVID-19 cases and deaths by state over time*. Retrieved from https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36

Elemental Editors. (2020). 50 Experts to trust in a pandemic. *Elemental*. Retrieved from https://elemental.medium.com/50-experts-to-trust-in-a-pandemic-fe58932950e7

Falconer, R. (2020, September 14). Trump to Woodward: "Nothing more could have been done" on coronavirus. *Axios*. Retrieved from https://www.axios.com/trump-woodward-covid-19-60-minutes-interview-4ef4035f-a08e-4709-ba57-013fc8c5ab04.html

Finset, A., Bosworth, H., Butow, P., Gulbrandsen, P., Hulsman, R. L., Pieterse, A. H., Street, R., Tschoetschel, R., & van Weert, J. (2020). Effective health communication - a key factor in fighting the COVID-19 pandemic. *Patient education and counseling*, *103*(5), 873–876. https://doi.org/10.1016/j.pec.2020.03.027

Gittleson, B. (2020, September 22). As nation surpasses 200,000 deaths, Trump keeps downplaying coronavirus. *ABC News*. Retrieved from https://abcnews.go.com/Politics/nation-surpasses-200000-deaths-trump-downplaying-coronavirus/story?id=73165811

Google. (n.d.). *Google trends*. Retrieved from trends.google.com

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In *Advances in experimental social psychology* (Vol. 47, pp. 55-130). Retrieved from doi.org/10.1016/B978-0-12-407236-7.00002-4

Grossman, G., Kim, S., Rexer, J. M., & Thirumurthy, H. (2020). *Partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States*, Retrieved from https://dx.doi.org/10.2139/ssrn.3578695

Helsingen, L. M., Refsum, E., Gjøstein, D. K., Løberg, M., Bretthauer, M., Kalager, M., & Emilsson, L. (2020). The COVID-19 pandemic in Norway and Sweden–threats, trust, and impact on daily life: a comparative survey. *BMC public health*, *20*(1), 1-10. https://doi.org/10.1186/s12889-020-09615-3

Herdemo. (2020) *HTTP Error, gives 4040 but the URL is working*. [Github Comment] Retrieved from https://github.com/Mottl/GetOldTweets3/issues/98#issuecomment-695340913

Hartney, M.T., & Finger, L. K. (2020). Politics, markets, and pandemics: Public education's response to COVID-19. (*EdWorkingPaper:* 20-304). Retrieved from doi.org/10.26300/8ff8-3945

Holland, S. & Hunnicutt, T. (2020, October 24). As COVID-19 hits swing states, Biden and Trump show sharp contrast. *Reuters*. Retrieved from https://www.reuters.com/article/usa-election/as-covid-19-hits-swing-states-biden-and-trump-show-sharp-contrast-idUSKBN27906J

Koetke, J., Schumann, K., & Porter, T. (2020). *Trust in science increases conservative support for social distancing*. https://doi.org/10.31219/osf.io/cngq8

Kruikemeier, S. (2014). How political candidates use Twitter and the impact on votes. *Computers in human behavior*, *34*, 131-139. Retrieved from doi.org/10.1016/j.chb.2014.01.025

Lemay, D.J., Basnet, R.B., Doleck, T., & Bazelais, P. (2019). Social network analysis of Twitter use during the AERA 2017 annual conference. *Education and Information Technologies 24*, 459–470. https://doi.org/10.1007/s10639-018-9787-2

Marley, P. (2020, October 20th). Rep. Joe Sanfelippo, who heads a health panel, contends there is nothing more the government can do to fight COVID-19. *Milwaukee Journal Sentinel.* Retrieved from https://www.jsonline.com/story/news/politics/2020/10/20/wisconsin-republican-says-lawmakers-have-done-all-they-can-do-covid/5998576002/

Nagourney A., & Peters, J. W. (2020, September 21). Denial and defiance: Trump and his base downplay the virus ahead of the election. *New York Times.* Retrieved from https://www.nytimes.com/2020/09/21/us/politics/trump-base-virus.html

Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015.* Austin, TX: University of Texas at Austin. Retrieved from dx.doi.org/10.15781/T29G6Z

Parmelee, J. H., & Bichard, S. L. (2011). *Politics and the Twitter revolution: How tweets influence the relationship between political leaders and the public.* Lexington books.

Rashid, M.A., McKechnie, D. and Gill, D. (2018). What advice is given to newly qualified doctors on Twitter? An analysis of #TipsForNewDocs tweets. *Medical Education 52*(7). 747-756. doi:10.1111/medu.13589

Robinson, M. D., Cassidy, D. M., Boyd, R. L., & Fetterman, A. K. (2015). The politics of time: Conservatives differentially reference the past and liberals differentially reference the future. *Journal of Applied Social Psychology*, *45*(7), 391-399. Retrieved from https://doi.org/10.1111/jasp.12306

Trumpey, A. J., Daigle, D., & Nwak, G. (2018) Communication during an outbreak of public health investigation. In A. S. Rassmussen & R. A. Goodman (Eds.), *The CDC field epidemiology manual.* Oxford University Press. Retrieved from https://www.cdc.gov/eis/field-epi-manual/chapters/Communicating-Investigation.html

Saad, L. (2021) Americans' political ideology held steady in 2020. *Gallup.* https://news.gallup.com/poll/328367/americans-political-ideology-held-steady-2020.aspx

Salzmann-Erikson, M. (2017). Mental health nurses' use of Twitter for professional purposes during conference participation using #acmhn2016. *International Journal of Mental Health Nursing, 27(2), 804–813.* doi:10.1111/inm.12367

Samuels, B. & Klar, R. (2020, March 23). Trump: 'We can't let the cure be worse than the problem itself'. *The Hill.* Retrieved from https://thehill.com/homenews/administration/488965-trump-hints-at-changes-to-restrictions-we-cant-let-the-cure-be-worse

Sha, H., Al Hasan, M., Mohler, G. Brantingham, P. J. (2020). *Dynamic topic modeling of the COVID-19 Twitter narrative among U.S. governors and cabinet executives.* Retrieved from arxiv.org/pdf/2004.11692.pdf

Social Blade (n.d.). *Social Blade Twitter Stats.* Retrieved from https://socialblade.com/Twitter/

Sterling, J., Jost, J. T., & Bonneau, R. (2020). Political psycholinguistics: A comprehensive analysis of the language habits of liberal and conservative social media users. *Journal of personality and social psychology*, *118*(4), 805–834. https://doi.org/10.1037/pspp0000275

Sullivan, P. (2020, September 6th). GOP uses debunked theory to downplay COVID-19 death toll. *The Hill.* Retrieved from https://thehill.com/policy/healthcare/515227-gop-uses-debunked-theory-to-downplay-covid-19-death-toll

The COVID Tracking project. (2020). Retrieved from https://covidtracking.com/

The Lancet (2020). COVID-19: fighting panic with information. *Lancet (London, England)*, *395*(10224), 537. https://doi.org/10.1016/S0140-6736(20)30379-2

Twintproject. (2020). *Twint.* Retrieved from https://github.com/twintproject/twint

Twitter Developer. (n.d.). *Product: search tweets: full archive.* Retrieved from https://developer.Twitter.com/en/pricing/search-fullarchive

U.S. Census Bureau. (2019). *State population totals and components of change: 2010-2019.* Retrieved from https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-total.html

Vraga, E. K., & Jacobsen, K. H. (2020). Strategies for effective health communication during the coronavirus pandemic and future emerging infectious disease events. *World Medical & Health Policy*, *12*(3), 233-241. https://doi.org/10.1002/wmh3.359

Wells, G. (2020, May 15). Doctors are tweeting about coronavirus to make facts go viral. *Wall Street Journal.* Retrieved from https://www.wsj.com/articles/doctors-are-tweeting-about-coronavirus-to-make-facts-go-viral-11589558880

Wise, J. (2020, July 23). Twitter usage surges amid coronavirus pandemic. *The Hill.* Retrieved from https://thehill.com/policy/technology/508670-Twitter-usage-growth-surges-amid-coronavirus-pandemic

Ziemba, Y. C., Razzano, D., Allen, T. C., Booth, A. L., Anderson, S. R., Champeaux, A., Feldman, M. D., Fitzhugh, V., Gittens, S., Grider, M., Gupta, M., Hanos, C., Kelly, K., Kothari, T., Laudadio, J., Lin, A. Y., Mirza, K. M., Montone, K. T., Prieto, V. G., … Markwood, P. S. (2020). Social media engagement at academic conferences: Report of the association of pathology Chairs 2018 and 2019 annual meeting social media committee. *Academic Pathology*. https://doi.org/10.1177/2374289520934019

Appendix 1: Custom Dictionaries.

None of the keywords are case sensitive and stemming is used for all keywords in appendix one and two. That is, “biolog” means that both “biology” and “Biologist” will be captured.

COVID-19 Keywords:

Virus names: coronavirus, 2019ncov, 2019-ncov, hcov-19, hcov 19, covid, sarscov2, sars-cov
General: epidemi, pandemic, disease, outbreak, infection, flatten the curve, ventilator
Solutions: stay at home order, stay-at-home order, stay home, stayhome, quarantine, lockdown, distanc, 6 feet apart, six feet apart, mask, vaccine, test, trace, tracing, hand washing, wash hand
Technical Language: positivity, aerosol, rate of positive, positive rate, % positive, percent positive, asymptomatic, R0, reproduction number, sarscov2, sars-cov, N95, respirator
Famous figures and organizations: Fauci, Birx, CDC,
Note: the virus names, general, solutions, technical language, and famous figures keywords are subsets of the COVID-19 keywords dictionary.

Consequences Keywords:

Death related: " die", dead, death, dying, coffin
Economy related: employment, economy, jobs, business, stock market
Note: the whitespace in front of the word " die" is used to filter out other words such as "studied"

Appendix 2: List of Influential Healthcare Professionals

Table 5. Initial HEs

UserID	Twitter Handle	Username	Criteria for inclusion	Followers (On Oct.30, 2020)
40156330	ScottGottliebMD	Scott Gottlieb, MD	MD	430,088
86626845	EricTopol	Eric Topol	MD	338,084
18831926	DrEricDing	Eric Feigl-Ding	Professor of epidemiology	316,215
224896427	trvrb	Trevor Bedford	Professor of epidemiology	268,998
75937326	mlipsitch	Marc Lipsitch	Professor of epidemiology	215,013
426909329	CDCDirector	Robert R. Redfield	MD	218,098
389313566	FaheemYounus	Faheem Younus, MD	MD	173,294
3238448948	CT_Bergstrom	Carl T. Bergstrom	Ph.D. in biology	127,119

When searching through the followers of initial HEs, we consider the following criteria. If any of them are met, then the user is included in our list.
User's Twitter handle contains: Dr, MD, Doctor, MPH
User's username contains: Dr, Doctor, MD, M.D., MPH
User's bio contains: biolog, " dr ", dr., doctor, epidemiolog, " MD ", M.D., virolog, " MPH ", public health
The space around " dr " , " MPH " and " MD " are included to filter out words like "Andrew."

Table 6. List of all health experts included in the study

User ID	Username	Qualification
426909329	CDCDirector	MD
232193350	rajshah	MD
2904169317	SteveFDA	MD
389313566	FaheemYounus	MD
18831926	DrEricDing	Doctorate in Epidemiology
605153786	DrLeanaWen	MD
40156330	ScottGottliebMD	MD
11274452	kevinmd	MD
1651522832	DrDenaGrayson	MD
86970530	tmprowell	MD
950783972	RepBera	MD
38531995	DrOz	MD
227429355	AmeshAA	MD
348075929	VirusWhisperer	MD
230769694	syramadad	MD
508591081	AbraarKaran	MD
29328876	sandrogalea	MD
18170896	drsanjaygupta	MD
17297668	larrybrilliant	MD
1094762324097822720	michaelmina_lab	MD
17240190	daniel_kraft	MD
35815074	JenniferNuzzo	DrPH in epidemiology
239210681	Bob_Wachter	MD
831279407088148480	Craig_A_Spencer	MD
817869452473548800	DrAlGrossAK	MD
151965668	AliRaja_MD	MD
86626845	EricTopol	MD
28023025	chngin_the_wrld	Ph.D. in Epidemiology
284093185	Farzad_MD	MD

58006725	celinegounder	MD
487673211	cmyeaton	MPH
887363635692969984	hiral4congress	MD
953924228306305024	Cleavon_MD	MD
317593544	priteshgandhimd	MPH
75937326	mlipsitch	Ph.D. in Epidemiology
1089859058	RepRaulRuizMD	MD
564099543	drjudymelinek	MD
846801440706351104	DrRobDavidson	MPH
593289567	PeterHotez	MD
65497475	eugenegu	MD
139173680	drjohnm	MD
745824471689244672	dremilyportermd	MD
394087611	angie_rasmussen	Ph.D. in Microbiology
170877963	DrSidMukherjee	MD
2710796718	BhadeliaMD	MD
30844417	gregggonsalves	MPH