

Identifying and characterizing scientific authority-related misinformation discourse about hydroxychloroquine on twitter using unsupervised machine learning

Big Data & Society
January–June: 1–15
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517211013843
journals.sagepub.com/home/bds
 SAGE

Michael Robert Haupt¹ , Jiawei Li^{2,3} and
Tim K Mackey^{2,3,4}

Abstract

This study investigates the types of misinformation spread on Twitter that evokes scientific authority or evidence when making false claims about the antimalarial drug hydroxychloroquine as a treatment for COVID-19. Specifically, we examined tweets generated after former U.S. President Donald Trump retweeted misinformation about the drug using an unsupervised machine learning approach called the biterm topic model that is used to cluster tweets into misinformation topics based on textual similarity. The top 10 tweets from each topic cluster were content coded for three types of misinformation categories related to scientific authority: medical endorsements of hydroxychloroquine, scientific information used to support hydroxychloroquine's use, and a comparison group that included scientific evidence opposing hydroxychloroquine's use. Results show a much higher volume of tweets featuring medical endorsements and use of supportive scientific information compared to accurate and updated scientific evidence, that misinformation-related tweets propagated for a longer time frame, and the majority of hydroxychloroquine Twitter discourse expressed positive views about the drug. Metadata from Twitter accounts found that prominent users within misinformation discourse were more likely to have media or political affiliation and explicitly expressed support for President Trump. Conversely, prominent accounts within the scientific opposition discourse primarily consisted of medical doctors or scientists but had far less influence in the Twitter discourse. Implications of these findings and connections to related social media research are discussed, as well as cognitive mechanisms for understanding susceptibility to misinformation and strategies to combat misinformation spread via online platforms.

Keywords

Misinformation, scientific evidence, twitter, machine learning, computational social science

This article is a part of special theme on Studying the COVID-19 Infodemic at Scale. To see a full list of all articles in this special theme, please click here: <https://journals.sagepub.com/page/bds/collections/studyinginfodemicatscale>

Introduction

COVID-19 infodemic

On 27 July 2020, former U.S. President Donald Trump retweeted to his more than 84 million Twitter followers an online video featuring Houston doctor Stella Immanuel, who with others from a group affiliated with conservative activists called “America’s Frontline Doctors,” falsely claimed that the

¹Department of Cognitive Science, University of California San Diego, La Jolla, USA

²Global Health Policy Institute, San Diego, USA

³S-3 Research, San Diego, USA

⁴Department of Anesthesiology and Division of Infectious Diseases and Global Public Health, University of California, San Diego School of Medicine, San Diego, USA

Corresponding author:

Tim K Mackey, 8950 Villa, La Jolla Drive, A124 La Jolla, CA 92037, USA.
Email: tmackey@ucsd.edu



combination of the antimalarial drug hydroxychloroquine, zinc, and the antibiotic Zithromax was a cure for COVID-19 (Morrison and Heilweil, 2020). Twitter subsequently removed President Trump's retweets of the video as a violation of the platform's misinformation policy, the first time it had taken such action on a tweet from the President for a COVID-19-related topic (Morrison and Heilweil, 2020). The incident was not the first, nor the last time the President would tout the benefits of the unproven drug, which a randomized control study carried out across 34 U.S. hospitals found as ineffective in improving clinical outcomes for COVID-19 (Self et al., 2020). Even before this study, concerns about the drug's safety profile and lack of evidence of efficacy for treating COVID-19 led to the halt of clinical trials by the U.S. National Institutes of Health and the World Health Organization (WHO) (Mahase, 2020; NIH, 2020). Despite these concerns, as early as March and as late as August 2020, President Trump actively promoted the drug in media appearances and during White House briefings, and even claimed he had used it as a prophylactic measure to prevent COVID-19 infection (Cathey, 2020).

A central concern among medical professionals and public health experts has been the potential impact of hydroxychloroquine misinformation to influence knowledge, attitudes, and behavior among the broader public, particularly among social media users. A prior study found that the President's misinformation tweets about hydroxychloroquine generated large volumes (1 million +) of misinformation posts and general support for the drug, in stark comparison to a much smaller volume of posts expressing concern about the drug and its safety/efficacy (Mackey et al., 2021). Another study reported a much higher number of impressions for President Trump's hydroxychloroquine post compared to his average tweets, greater airtime about the drug on conservative TV networks, and increases in Google searches immediately following the President's retweet (Niburski and Niburski, 2020). Data queried from the Stanford Cable TV News Analyzer (n.d.) (see supplementary material), which continuously transcribes near 24-7 broadcast recordings of CNN, Fox News, and MSNBC also shows that mentions of hydroxychloroquine spiked from an average of 1.39 minutes on the week of 20 July to 13.78 minutes on the week of 27 July, and then dropped to .67 minutes by 10 August, indicating that the majority of media coverage related to hydroxychloroquine occurred shortly after Trump's retweet.

Importantly, despite Twitter's removal of the President's retweet, the "America's Frontline Doctors" video went viral on other social media platforms such as YouTube and Facebook, where it was

subsequently shared and viewed by millions of other online users (Frenkel and Alba, 2020). Hence, despite the growing body of scientific evidence disproving hydroxychloroquine's use for COVID-19, the volume and diversity of misinformation about the drug increased substantially due to this event involving a person of power and influence, feeding into the larger COVID-19 "infodemic" where other rumors, stigmas, and conspiracy theories about the virus continue to circulate and rapidly spread online (Islam et al., 2020).

Psychological and cognitive effects of misinformation

Previous research has demonstrated the damaging impact of misinformation and the psychological effects that allow misinformation to propagate (Lewandowsky et al., 2020). Not only does a democracy rely on a well-informed public in order to function properly (Kuklinski et al., 2000), but misinformation also has a direct impact on public health, as seen with the increase in cases of vaccine-preventable diseases due to anti-vaccination movements (Poland and Spier, 2010). Even just the exposure to misinformation in the form of a conspiracy theory can have a negative effect on trust in government services and institutions (Einstein and Glick, 2015). Misinformation has also been shown to be psychologically pervasive and difficult to correct once an individual adopts an incorrect belief. As demonstrated in the Illusory Truth Effect (Hasher et al., 1977; Henkel and Mattson, 2011), repeated information is more likely to be judged as true because of its familiarity, while the continued influence effect (Hamby et al., 2020; Johnson and Seifert, 1994) shows that people continue to rely on inaccurate information in their memory even after a credible correction has been made. It is also likely that these psychological effects are exacerbated on social media platforms that encourage sharing in sometimes sparse networks of like-minded users and exhibit echo chamber effects concerning political issues (Barberá et al., 2015).

Recent research also examines cognitive mechanisms for understanding misinformation spread and susceptibility to false beliefs. Mosleh et al. (2021) shows that Twitter users who score higher on cognitive reflection (Frederick, 2005) – a measure designed to assess an individual's propensity to engage in self-reflective, analytic thinking over initial intuitive responses – shared news content from more reliable sources and were more selective about who they follow. In contrast to a motivated reasoning explanation of misinformation that claims people spread false claims because they select and dismiss evidence based on their group identity (Kahan, 2017), Pennycook and Rand (2019, 2020) provide evidence arguing that misinformation spread is

driven rather by a lack of critical thinking and “reflexive open-mindedness” (i.e. a tendency to be overly accepting of weak claims). Within the context of COVID-19, Pennycook et al. (2020) showed that participants with higher cognitive reflection scores were better at discerning between true and false news content. They also showed that being prompted to rate the accuracy of a non-COVID-related headline before being asked whether they would share a COVID-related headline nearly tripled participants’ level of discernment between sharing true and false headlines. Building on these approaches, this study will analyze affiliations of influential Twitter users and the level of replicated messages per topic using a measure called an Echo score (Haupt et al., 2021) to examine whether the hydroxychloroquine discourse reflects either motivated reasoning or reflexive open-mindedness accounts of misinformation spread related to scientific credibility.

Use of scientific credibility and authority

While it appears that accurate scientific evidence and recommendations from established medical organizations may have limited impact in mitigating misinformation spread across online platforms, it is worth noting that some of the more viral types of misinformation explicitly evoke scientific authority and credentials when making their claims. This is true in the case of President Trump’s misinformation retweet, as the source of the misinformation came from a group called “America’s Frontline Doctors” where Dr Stella Emmanuel and other clinicians used their backgrounds as medical professionals as source credibility to promote hydroxychloroquine as a COVID-19 treatment. This is certainly not the first instance in which scientific authority was used to back questionable claims concerning important public health topics. For example, in response to increasing evidence linking cigarette smoking to lung cancer, the Tobacco Industry Research Committee was formed in 1953 with the main goal of promoting scientific research contradicting the consensus that smoking kills, and by 1986 had spent over \$130 million on sponsored research resulting in 2600 published articles (O’Connor and Weatherall, 2019). In order to change U.S. eating habits to include bacon as a staple of an “American Breakfast,” Edward Bernays, also known as the “father of public relations,” reported that he found a physician willing to state publicly that a hearty breakfast which included bacon was healthier than a lighter breakfast with the full knowledge that a large number of people will “follow the advice of their doctors” (Bernays, 1928).

More recently, celebrity physician Dr Mehmet Oz (“Dr Oz”), who in 2018 was appointed by President Trump to the President’s Council on Sports, Fitness, and Nutrition, used his medical credentials as a

professor of surgery at Columbia University to promote health treatments on his daytime television talk show despite findings showing that only 46% of his recommendations were supported by scientific evidence (Korownyk et al., 2014). However, online sales for health products can spike up to 12,000% after being featured on his program (Belluz and Hoffman, 2013), which has subsequently led to Dr Oz having a current net worth of \$100 million (Celebrity Net Worth, 2019). While media programs can serve a beneficial role by communicating important health behavior change and scientific findings in accessible ways, it is also important to recognize the impact of these communications when delivered by individuals who both rely on their scientific and medical credibility, but also have incentives to raise viewership ratings and promote views or health product for economic considerations.

Applying this to the COVID-19 pandemic, which is now a global health crisis that is also characterized by its parallel “infodemic,” an assessment of how representations of scientific authority are being skewed to shape public perception for political purposes is needed. This is particularly important as this pandemic occurs in a post-digital era when the rapidity of information dissemination has been accelerated due to the ubiquity of Internet and social media use. This means that undue political influence on scientific dissemination and public health communication can quickly spread and become viral, directly enabling acceptability of interventions that may lack credible evidence. In fact, concerns about “superspreader” events involving influential medical professionals promoting suspect COVID-19 testing kits have also been reported (Mackey, 2020). Previous work has also used survey data to investigate how deference towards scientific authority is influenced by media (Lee and Scheufele, 2006) and drives support of biotechnology (Brossard and Nisbet, 2007); however, there does not appear to be substantive research specifically examining the use scientific authority to promote misinformation particularly in the context of COVID-19.

Building on prior research, this study examines misinformation that evokes scientific authority to promote unsubstantiated claims about treatment for COVID-19. More specifically, this study identifies and characterizes hydroxychloroquine misinformation tweets generated after President Trump’s retweet and classifies them into three scientific credibility and authority categories using an inductive coding approach. These categories include medical endorsements, scientific evidence that supports hydroxychloroquine, and for comparison purposes scientific evidence that opposes hydroxychloroquine use. Propagation of the three scientific credibility types is compared, and affiliations of prominent accounts associated with each type of tweet are also analyzed.

Methods

Data collection and analysis using unsupervised machine learning

A total of 2,771,730 tweets were collected from the public streaming Twitter API using “hydroxychloroquine” and “chloroquine” as keywords for filtering between 21 and 30 July 2020. We chose this period as it represents a study time frame immediately prior to and after an event (i.e. Trump’s misinformation retweet) that generated a large volume of hydroxychloroquine discussion from twitter users, including misinformation tweets, which could be analyzed for key misinformation themes of interest associated with scientific authority and credibility (Mackey et al., 2021). From this corpus of tweets, the top 100 most retweeted tweets were coded for misinformation using a binary coding scheme based on existing COVID-19 misinformation themes previously reported in the literature (Islam et al., 2020; Lu, 2020). See Table A1 in supplementary material for the full list of criteria used to code misinformation. Due to the sharp increase in the volume of hydroxychloroquine-related tweets immediately after President Trump’s misinformation retweet (see Figure A1 in supplementary material), 2,289,441 tweets (82.6% of the entire dataset) from the period after 27 July were chosen for further analysis to detect misinformation themes related to scientific authority using an inductive coding scheme (further discussed in the Content Coding of tweets for scientific authority themes section). As seen in Figure A2 in the supplementary material, the volume of hydroxychloroquine-related tweets dropped sharply after 30 July, indicating that the majority of the Twitter discourse concerning the drug occurred during the study time frame. In order to analyze the large volume of tweets collected in this study, we used the biterm topic model (BTM), an unsupervised machine learning approach using natural language processing (NLP), to extract themes from text of tweets as used in prior studies examining COVID-19 topics on social media (Mackey et al., 2020a, 2020b), including political protests in response to COVID-19 stay at home and quarantine measures (Haupt et al., 2021).

Unsupervised topic modeling strategies, such as BTM, are a method particularly well suited for sorting short text (such as the 280-character limit for tweets) into highly prevalent themes without the need for pre-determined coding or a training/labeled dataset to classify specific content. This is particularly useful in characterizing large volumes of unstructured data where predefined themes are unavailable, such as in the case of emerging social movements, novel disease outbreaks, and other emergency events where

information changes rapidly. The corpus of tweets containing the hydroxychloroquine keywords was categorized into highly correlated topic clusters using BTM based on splitting all text into a bag of words and then producing a discrete probability distribution for all words for each theme that places a larger weight on words that are most representative of a given theme (Kalyanam et al., 2017). While other natural language processing approaches use unigrams or bigrams for splitting text, BTM uses “biterms,” which is a combination of two words from a text (e.g. the text “go to school” has three biterms: “go to,” “go school,” “to school”) and models the generation of biterms in a collection rather than documents (Yan et al., 2013).

BTM was used for this study because biterms directly model the co-occurrence of words which increases performance for sparse-text documents such as tweets. Conducting BTM analysis is done initially by setting the BTM topic number (k) and “ n ” words (for the first round of analysis we set at $k = 10$, $n = 20$ to cover several possible misinformation topics that might be present in the corpus). A coherence score is then used to measure how strong the top words from each topic correspond to its respective topic. For this study, the model with $k = 30$ was chosen because it had the highest coherence score compared to other iterations tested. All data collection and processing were conducted using the programming language Python. See supplementary item 7.7 for additional information on how BTM is operationalized for this dataset.

Calculating echo measure

Replication of tweet text was measured to observe message resonance throughout each misinformation topic in the network of Twitter users examined in this study. Tweets were grouped by text content, tweet ID and BTM topic to produce the number of unique tweets. A metric introduced in a previous Twitter COVID-19 study called “Echo” (Haupt et al., 2021) was used to represent the ratio of total tweets per topic by the number of unique tweets (i.e. represented by the number of unique tweets where retweets replicating the same message were removed), as depicted in the following formula:

$$E = T/u$$

where E = represents the Echo measure, T = total number of tweets by topic, and u = number of unique tweets by topic. An Echo closer to 1 signifies a higher number of unique tweets within a given topic, while an Echo with larger values reflects higher replication levels of the same message. In the context of Twitter, replicated text among tweets suggest that users are

retweeting without inserting additional comments, their own opinion, or expressing unique sentiment, while retweets that add commentary to the original source are expressions of more unique opinion. Echo was important to measure in the context of assessing the replication and propagation of tweets, including tweets containing unaltered misinformation, across the broader Twitter network. See supplementary material for example of how Echo is calculated.

Content coding of tweets for scientific authority themes

In order to characterize highly prevalent misinformation topics in the corpus, the top 10 most retweeted tweets from all BTM topic outputs were extracted and manually coded for relevance first using our binary deductive coding scheme based on existing COVID-19 misinformation themes from the literature, then sub-coded for scientific credibility and authority themes, and then manually annotated for sentiment with specific emphasis on understanding the context of the tweet text in relation to hydroxychloroquine. Sentiment analysis was coded for positive and negative reactions toward the use of hydroxychloroquine for COVID-19, with 1 indicating positive sentiment, -1 indicating negative sentiment, and 0 if the tweet only reports information about hydroxychloroquine without stating an opinion or exhibited neutral user sentiment. Three categories of scientific credibility and authority type (*Medical Endorsement*, *Scientific Support*, and *Scientific Opposition*) were used to sub code tweets related to misinformation spread after binary coding was completed. An inductive coding approach was used to create the scientific credibility and authority categories, which was also partially informed by a combination of existing COVID-19 misinformation categories and prior studies on scientific credibility use per the definitions below (Islam et al., 2020; Lu, 2020) (see Table 1 for examples of tweets for each category):

1. A tweet was coded as a *Medical Endorsement* if it evokes medical authority or expertise when supporting the use of hydroxychloroquine as a COVID-19 treatment (i.e. mentions a doctor or medical expert endorsing hydroxychloroquine). While it is possible for organizations such as medical associations or scientific societies to also use their reputations as the basis for an endorsement, we found that in this study scientific organizations were more likely to include evidence when making a claim instead of explicitly relying on their organization's reputation.
2. A tweet is categorized as *Scientific Support* if it presents scientific evidence and findings that support

hydroxychloroquine use. We defined scientific evidence as findings obtained by scientific method or scientific instruments based on scientific theories used to prove a specific claim (Walton and Zhang, 2013). While *Medical Endorsement* relies on a call to the scientific credibility and qualifications of a person supporting hydroxychloroquine use, *Scientific Support* relies on the credibility of the evidence such as a study or other scientific information to support its claim.

3. Lastly, a tweet was classified as *Scientific Opposition* if it uses scientific evidence and findings to discourage the use of hydroxychloroquine.

In order to focus the analysis on misinformation spread in the context of scientific credibility and authority, only topics that had at least 50% of the top 10 retweeted tweets (a majority of content in the BTM topic output) confirmed and labeled as associated with the three categories above were chosen for further analysis (see Table A2 in supplementary material for example tweets related to each topic). First and last authors coded posts independently for scientific credibility and authority misinformation topics and achieved a high intercoder reliability ($\kappa = 0.92$). For inconsistent results, authors met and conferred on correct categories and themes.

User accounts were also grouped together by sentiment scores, which were calculated by taking the average sentiment of each user's tweets associated with misinformation categories. Users with sentiment scores greater than .75 were classified as "Strongly Positive," (i.e. twitter accounts expressing strong support for hydroxychloroquine use for COVID-19 treatment on average) scores between .25 and .75 were classified as "Positive," scores between -.25 and .25 were categorized as "Mixed," scores between -.25 and -.75 were classified as "Negative," and scores less than -.75 were classified as "Strongly Negative."

Content coding of twitter account profiles

Twitter profiles of accounts that produced the top 10 most retweeted tweets for each BTM scientific authority and credibility topic output were content coded to investigate publicly self-reported occupation and political affiliations among these prominent and influential users who were active in the Twitter hydroxychloroquine online discourse during the study period. Publicly available metadata from user account profiles were retrieved and coded to determine whether descriptions stated affiliations with the media (e.g. a journalist, TV, or radio personality), that they were a medical doctor or scientist, or were a politician or worked for a political organization. If an account worked in the

media, the name of the media outlet (e.g. CNN) was also recorded. Since our study examined a period when hydroxychloroquine discourse was driven primarily by President Trump's original 27 July misinformation retweet supporting hydroxychloroquine use (Mackey et al., 2021), accounts were also coded for whether they explicitly mentioned support for Trump on their profile page. An account was considered supportive of President Trump if they used phrases such as "Trump 2020" or "MAGA" in their profile descriptions or featured a picture with President Trump as the cover photo of their profile.

Results

Content analysis and characterization

Out of a total of 2,771,730 tweets collected from 21 to 30 July, the initial coding used in this study extracted the top 10 most retweeted tweets from both the period immediately preceding President Trump's

misinformation retweet (prior to 27 July, $n = 174,600$ tweets and retweets) and the period immediately following the retweet (on and after 27 July, $n = 1,063,399$) accounting for a total of 44.7% ($n = 1,237,999$) of the entire dataset of tweets collected. Of the 1,237,999 tweets coded with the binary coding scheme, 1,044,794 (84.4%) were classified as general COVID-19 misinformation categories, including an observation that the pre-period had an overall lower proportion (78.4%, $n = 136,963$) of misinformation tweets compared to the post-period (85.4%, $n = 907,831$), though generally misinformation themes were much more prevalent than tweets that expressed concern about hydroxychloroquine (5.4%, $n = 56,879$) or were neutral/unrelated to misinformation (Mackey et al., 2021). Results from this prior study identified general hydroxychloroquine misinformation themes that included claims of treatment efficacy (including rumors, examples of clinical treatment, and use of non-traditional medicine), calls for authority action demanding drug access, claims of censorship of

Table 1. Examples of content coded tweets (paraphrased and redacted to retain anonymity).

Information types	
Hydroxychloroquine sentiment	Medical endorsement: 1. Texas Doctor reports how he was intimidated and bullied by his local Medical Board for SUCCESSFULLY treating patients with #Hydroxychloroquine. HCQ works but there's no money in it for Big Pharma. Dr [REDACTED] Infectious Disease Specialist #HydrochloroquineWorks
	Scientific support: 1. More evidence presented for why hydroxychloroquine should be made available in a new court filing by the Association of American Physicians & Surgeons 2. Treatment with hydroxychloroquine cut death rate significantly in COVID-19 patients @HenryFordNews health system study shows. "The findings have been highly analyzed and peer-reviewed," Dr [REDACTED] "#Hydroxychloroquine is the Beginning of the End of the Pandemic," Dr [REDACTED] announces hydroxychloroquine study that is, "Game Changer," in battle against coronavirus
	Scientific opposition: 1. Just a reminder: we now have four major trials all congruent in demonstrating hydroxychloroquine provides NO benefit in COVID-19. 1. Post exposure ppx: [LINK 1], 2. RECOVERY (UK): [LINK 2], 3. ORCHID (USA): [LINK 3], 4. SOLIDARITY (WHO) 2. Debunking hydroxychloroquine (again), that viral HCQ video today it's time to bump up this thread on the mega RECOVERY randomized trial of HCQ with 4700 people showing NO benefit for mortality and even higher risk of ventilator+mortality. And no subgroups benefit. #COVID19
	Positive sentiment: 1. Zinc, Vitamin D, Vitamin C, and hydroxychloroquine are not the cures but a vaccination that was quickly made and not tested on animals is? 2. Viruses mutate. A vaccine is a fools errand. Treat the symptoms. Prescribe #Hydroxychloroquine. Work on herd immunity. Keep high risk isolated from herd.
	Negative sentiment: 1. The drug doesn't work as a prophylaxis, as a preventative it doesn't work in the treatment of early disease and it very convincingly does not work in hospitalized patients. . . We can definitively say hydroxychloroquine doesn't work 2. Hydroxychloroquine with or without azithromycin in mild-to-moderate Covid-19. "the use of hydroxychloroquine alone or with azithromycin did not improve clinical status at 15 days as compared with standard care" [LINK]

favorable evidence for hydroxychloroquine, and other conspiracy theories about the drug, but did not specifically assess claims of scientific or medical credibility or authority (Mackey et al., 2021).

Further analysis based on reviewing outputted word groupings by BTM topic clusters allowed us to identify specific misinformation themes in the entire corpus of Twitter hydroxychloroquine discourse during the study period, including specific scientific credibility and authority hydroxychloroquine topics, user affiliations, and measure replication rates of messages associated with each topic of interest. As stated previously, only topics that had at least 50% of the top 10 retweeted tweets that were coded as including categories associated with *Medical Endorsements*, *Scientific Support*, or *Scientific Opposition* were chosen for further analysis in this study. Thirty topic clusters in total were produced by BTM, and eight were selected based on these criteria that first manually annotated all top 10 tweets in BTM clusters to assess relevance to the study aims. For additional details about all topics produced in this study, see Table A3 in the supplementary material.

Table 2a shows all BTM topics that are composed predominantly of tweets related to *Medical Endorsements*, *Scientific Support*, or *Scientific Opposition*. Columns “%Med Endorse,” “%Sci Support,” and “%Sci Oppose” indicate the percentage of tweets categorized in each respective category by topic weighted by number of retweets, and the column “Info type” categorizes each topic based on

which type of information is most prevalent (e.g. Topic 9 has 98.2% of its top 10 most retweeted tweets as medical endorsements compared to the other categories so it is labeled accordingly). In total, five topics were classified as *Medical Endorsements*, two as *Scientific Support*, and only one as *Scientific Opposition*. The column “Avg sentiment” is based on the average sentiment scores of the top 10 retweeted tweets for each topic. Additionally, the column “Retweet top 10%” indicates the percentage of total tweets for each topic that is accounted for by the top 10 retweeted tweets. As shown in Table 2a, at least 50% of the total number of tweets for each topic are composed of the top 10 retweets, indicating that the top 10 retweets characterizes the majority of tweets assigned to each topic. It is not surprising that the top 10 retweets can account for the majority of tweets within each topic as previous social media research shows that many online discussions are guided by a small number of very active users (Papakyriakopoulos et al., 2020).

Table 2b shows the average scores of each topic grouped by information type. As shown below, both misinformation types (i.e. *Medical Endorsements* and *Scientific Support*) have a much higher volume of total tweets compared to tweets expressing concern or citing evidence against the use of hydroxychloroquine in the *Scientific Opposition* category. *Medical Endorsements* in particular had a total of 413,072 tweets, which consists of 18.04% of the total corpus

Table 2a. Analysis by misinformation scientific credibility and authority topic.

Topic	Info type	Tweet count	Unique tweets	Echo	Avg sentiment	%Med Endorse	%Sci Support	%Sci Oppose	Retweet top 10%
9	Med endorse	83,399	2803	29.75	.96	98.2%	0.0%	1.8%	80.6%
11	Sci support	16,868	1216	13.87	1.0	0.0%	76.9%	0.0%	77.2%
13	Med endorse	114,764	5563	20.63	1.0	87.6%	0.0%	0.0%	77.0%
15	Med endorse	132,104	11158	11.84	.92	64.5%	7.3%	2.7%	53.8%
17	Sci support	13,891	2616	5.31	.68	0.0%	100.0%	0.0%	62.2%
20	Med endorse	20,873	2613	7.99	1.0	58.6%	31.7%	0.0%	70.6%
23	Med Endorse	61,932	3696	16.76	.52	51.8%	0.0%	0.0%	74.1%
24	Sci Oppose	11,601	2263	5.13	−.89	0.0%	5.0%	94.4%	66.6%

Note: Top 10 tweets were weighted by number of retweets.

Table 2b. Analysis by misinformation scientific credibility and authority sentiment and echo scores.

Information type	Number of tweets	% Total tweets	Avg sentiment	Avg echo	Number of topics
Med Endorse	413,072	18.04%	0.88	17.39	5
Sci Support	30,759	1.34%	0.84	9.59	2
Sci Oppose	11,601	0.51%	−0.89	5.13	1

Note: Since *Scientific Opposition* is only composed of one topic, it only shows figures related to Topic 24 and is not an average of other topics.

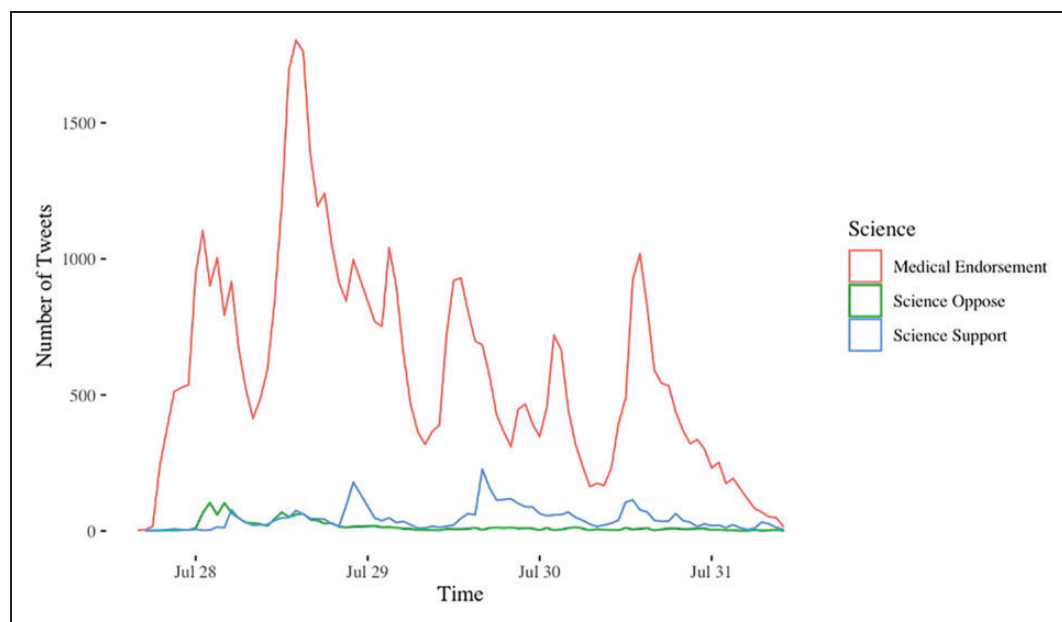


Figure 1. Number of tweets by information type.

collected during this entire study. While *Scientific Support* had a smaller volume at 30,759 tweets, it is still over double the amount of tweets compared to *Scientific Opposition*, which accounts for less than 1% (.51%) of the total volume of tweets collected. Both *Medical Endorsements* and *Scientific Support* also have higher average echo scores (17.34 and 9.59) compared to *Scientific Opposition* (5.13), indicating that messages within these topics are more often replicated unaltered throughout the Twitter network.

Figure 1 shows a line graph depicting the volume of tweets for topics grouped by scientific credibility and authority categories following the President's misinformation retweet. As shown in the figure, noticeable peaks in the volume and time periods of misinformation topics for *Medical Endorsement* and *Scientific Support* provide early indication that these specific misinformation themes may have co-occurred or created interaction among similar-minded users, including potentially amplifying overall pro-hydroxychloroquine messages at similar time intervals across different Twitter user groups. The highest volume of tweets were from *Medical Endorsement* before 29 July indicating that the peak of the hydroxychloroquine discourse occurred around two days after President Trump retweeted the misinformation video. In contrast, *Scientific Opposition*, which is represented by the green line, has two small peaks in the beginning of this time frame and then tapers off past 29 July and overall has a much smaller volume of engagement compared to both *Medical Endorsement* and *Scientific Support* misinformation topics that show multiple

peaks and high levels of engagement throughout the reviewed time frame. This may also indicate that these misinformation topics resonated and were amplified across Twitter user communities, whereas science-based information about the risks of the drug were drowned out by misinformation. These results show that not only do the misinformation related topics show higher volume in tweets and retweets compared to *Scientific Opposition*, but these discourses persist for longer periods of time among a diverse set of misinformation topics and more user networks.

User sentiment

In order to detect differences among users based on their attitudes and views towards hydroxychloroquine, accounts were grouped together by the average sentiment scores of their tweets as seen in Table 3 (see Table 1 in the Methods section for example of tweets by sentiment). This analysis included all twitter users in the dataset coded and extracted based on BTM and the top 100 retweeted tweets, not only the BTM topic clusters that were identified as relevant to *Medical Endorsement*, *Scientific Support*, and *Scientific Opposition* misinformation topics. Users classified as "Strongly Positive" had both higher mean number of tweets (5.12, standard deviation (sd) = 10.13) as well as a higher max number of tweets (310) throughout the five-day post-period compared to "Strongly Negative" users that had an average of 1.10 tweets (standard deviation (sd) = .634) and max of 40 tweets. This indicates that users who express strong support for

Table 3. Analysis by sentiment level.

	Strongly positive	Positive	Mixed	Negative	Strongly negative
Mean tweet	5.12	3.94	2.58	2.89	1.10
Stand dev	10.13	7.46	3.07	2.69	0.63
Median	2	2	1	2	1
75th percentile	5	4	3	3	1
90th percentile	13	9	6	4	1
99th percentile	52	36	15	9	3
Max	310	437	122	93	40
Users (n)	264,379	205,576	33,192	5281	23,771
Users (%)	49.7%	38.6%	6.2%	1.0%	4.5%

hydroxychloroquine were more active throughout the Twitter discourse. Additionally, 88.3% of users were classified as either “Positive” or “Strongly Positive” in contrast to 5.5% of users categorized as “Negative” or “Strongly Negative,” indicating that the vast majority of users examined in this study were engaged in topics that predominantly expressed support for using hydroxychloroquine to treat COVID-19, a result not surprising given that the total volume of misinformation topics and posts were much higher than concerns or evidence-based information about the drug.

Characteristics of user accounts

Accounts that produced the top 10 retweeted tweets for *Medical Endorsement*, *Scientific Support*, and *Scientific Opposition* misinformation topics were coded for occupation status and political affiliation in order to examine the background of prominent users driving the hydroxychloroquine discourse. As mentioned in the Methods section, accounts were classified based on if they worked in the media, were a physician or scientist/researcher, were a politician or worked for a political organization, and if they explicitly expressed support for President Trump in their profile. Figure 2 depicts differences in affiliations between accounts when grouped by misinformation type, showing the average percentage of affiliations of respective topics reviewed. Accounts within *Medical Endorsement* and *Scientific Support* topics were more likely to be affiliated with a media outlet (24% and 10%) and more likely to have a political affiliation (14% and 5%) compared to *Scientific Opposition*, which had no Twitter accounts that explicitly stated a link to media or politics. Misinformation topics were more likely to have accounts expressing support for Trump, with *Scientific Support* showing the highest average percentage at 50%, followed by *Medical Endorsement* at 36% compared to only 10% of the *Scientific Opposition* accounts (e.g. upon closer examination, the Trump

supporter account in the scientific opposition discourse expresses a minority opinion in the topic cluster by posting scientific support for hydroxychloroquine).

Additionally, 80% of the top 10 retweeted tweets within *Scientific Opposition* came from a single account that stated they were a medical doctor or research scientist in contrast to lower overall representation of doctors or scientists in *Medical Endorsement* (8%) and *Scientific Support* (5%) categories. These results implicate media and political-affiliated Twitter user groups as those that were predominantly propagating misinformation-related topics in comparison to those claiming to be doctors/scientists, and included explicit support for President Trump as a prevalent theme among users who promoted misinformation. This is in contrast to the *Scientific Opposition* discourse which had no involvement from the media or political groups and was primarily composed of scientists or doctors, though the overall volume and diversity of users in this category was low.

Table 4 shows additional information related to the Top 10 retweeted accounts in categories reviewed, including average follower count, number of unique accounts, number of accounts that were suspended during the time profile data was collected, and associated media outlets among users who worked in the media. Accounts within the *Medical Endorsement* and *Scientific Support* topics had a higher average number of followers compared to users in the *Scientific Opposition* topic, indicating that tweets from these accounts within misinformation-related topics reached a larger scale of Twitter users. Among users working in the media within the *Medical Endorsement* topics, the majority were associated with conservative news outlets such as *News Max*, *Fox News*, and *One America News Network*. The one non-conservative media outlet (MSNBC) found within the *Medical Endorsement* category was associated with the 2.7% of scientific opposition found within topic 15 (see Table 2a). Results from both Figure 2 and Table 4 indicate higher media involvement, particularly from conservative

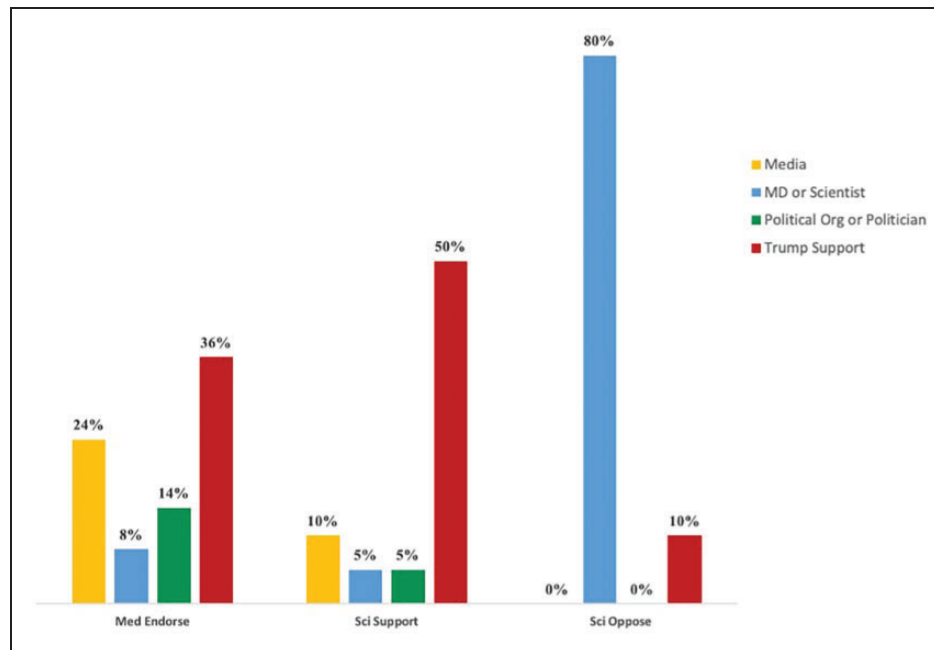


Figure 2. Affiliations among top 10 most retweeted accounts by information type (average of topics).

Table 4. Analysis of top 10 retweeted accounts by information type.

Account information	Med Endorse	Sci Support	Sci Oppose
Avg follower count	533,798	465,224	122,215
Median	541,917	465,224	18,383
75th percentile	450,897	468,594	288,383
90th percentile	576,744	470,616	347,045
99th percentile	652,253	471,830	347,045
Unique accounts	49	20	7
Suspended	7	3	0
Media outlet	News Max (2), Fox News (1), MSNBC (1), One America News Network (1)	Fox News (1)	

Note: Numbers in brackets () under Media outlet column indicate the number of times each outlet was present within the dataset. Profile information was collected on 1 December 2020, which is about four months after the tweets were collected for the main analysis.

media outlets, among *Medical Endorsement* topics, which might also account for the higher average number of followers within this category. See Table A4 in the supplementary material for additional information about the individual tweets produced by each media outlet.

Discussion

Misinformation and scientific evidence

The results from this study show that the majority of tweets that included discussion about hydroxychloroquine and were related to scientific credibility and authority topics following President Trump's 27 July

retweet comprised of misinformation, with only a small percentage of tweets communicating scientific evidence that raised well-established concerns about the drug's safety and efficacy for treating COVID-19. Not only did misinformation-related topics have a much higher volume of tweets (88.3% of users were engaged in topics expressing positive support for the use of hydroxychloroquine with *medical endorsements* topics comprising 18.04% of the total dataset), but these topics also had longer time durations and higher echo scores indicating that the messages were replicated unaltered throughout the social network. Examining prominent users who generated the top 10 retweeted tweets for each topic shows that accounts within misinformation-related topics were more likely

to work for conservative media outlets, political organizations or politicians, and explicitly expressed support for President Trump on their profile. These users were also more likely to have higher follower accounts, which coincides with previous work showing that journalists are more active on Twitter (Boldrini et al., 2018). Although the original source of misinformation originated from a group claiming medical and scientific credibility and authority (i.e. Dr Immanuel and “America’s Frontline Doctors”), the amplification of this misinformation primarily was disseminated by political (e.g. Trump) and media sources that arguably further lack scientific source credibility. In contrast, from the one topic that was composed primarily of scientific evidence in opposition to hydroxychloroquine’s use, no media or political accounts were involved and the majority of the most retweeted users were either medical doctors or scientists.

These findings align with previous research investigating rumor propagation and false news spread throughout social media networks. As demonstrated in Shin et al. (2018), misinformation on Twitter tended to reappear multiple times after its initial publication, while true facts did not. This is consistent with the results shown in the timeline analysis from Figure 1 where both misinformation topics continued for a longer time period than the *Scientific Opposition* topic which was more active in the early stages of the discourse. The high volume of misinformation and pro-hydroxychloroquine tweets are also consistent with findings analyzing approximately 126,000 rumors spread by around 3 million people on Twitter from 2006 to 2017 that show false news on Twitter spreads farther and more broadly than true stories, with false political news having the most pronounced effect (Vosoughi et al., 2018). While medical endorsements and scientific evidence supporting hydroxychloroquine use would traditionally be considered a domain of medical or public health professionals, the past and continued political polarization of COVID-19 is having an impact on science communication, making the promotion of hydroxychloroquine unique compared to more traditional, economically driven misuses of scientific authority as previously discussed. Previous politicized actions in the context of COVID-19 include Republican elected-officials stating that they would not wear masks during the pandemic (Brenan, 2020; *CNBC*, 2020), President Trump undermining the director of the National Institute of Allergy and Infectious Diseases Dr Anthony Fauci when administrating guidelines (Miller, 2020), and Trump actively retweeting misinformation related to hydroxychloroquine as a viable COVID-19 treatment that was then propagated and amplified by politically partisan Twitter users as demonstrated in this study. Subsequent public health

interventions for mitigating misinformation spread related to the COVID-19 infodemic should recognize the political elements inherent within these discourses in the digital sphere.

When considering cognitive mechanisms that could influence misinformation spread, the hydroxychloroquine discourse generated by influential users appears to be driven more by politically motivated reasoning since those who identified as having conservative values or expressed support for Trump were more likely to engage in misinformation topics. However, since we do not have data on the political affiliations for all users in the dataset, it is not possible to discern if less influential users are retweeting misinformation due to a lack of cognitive reflection or in order to defend a political identity. Overall, within the context of the hydroxychloroquine Twitter discourse it appears that influential users are more politically motivated when spreading misinformation, while typical Twitter users might share misinformation due to either motivated reasoning or reflexive open mindedness.

Mitigating misinformation spread during an infodemic

Multiple interventions have been proposed to address misinformation spread from a research perspective. Lewandowsky et al. (2020) produced a consensus-based handbook detailing a comprehensive overview of misinformation spread and discusses strategies for correcting misinformation such as “prebunking” claims (i.e. preemptively warning about potential misconceptions before they arise), setting your own talking points, and attacking underlying logical fallacies of misinformation claims. When examining how advocacy organizations stimulate public conversation about social problems on Facebook, Bail et al. (2017) have shown that more conversation is generated when producing emotional messages after prolonged rational debate and vice versa. When combating misinformation related to the COVID-19 infodemic, perhaps altering between messages containing scientific evidence or statistics with more emotional anecdotal accounts (e.g. personal accounts of COVID-19 illness and its health consequences) could encourage higher engagement with accurate information among online users. Torabi Asr and Taboada (2019) suggest using NLP to immediately detect fake news articles and announced a call to action for other researchers to make publicly available textual data with labeled fake news articles to build a comprehensive training dataset. While this approach can be effective in detecting fake news articles within specific contexts, it might not be able to address instances where outdated information or findings that were considered accurate at the time of

publication are then used misleadingly in contemporary scenarios, such as in the case of emerging but now more conclusive evidence around hydroxychloroquine. Since science itself is an iterative process and individual studies can produce conflicting results, solely relying on an NLP approach might not be able to account for the nuance and proper contextual knowledge required to consistently make accurate classifications without continuously updating training datasets.

Even though findings from Vosoughi et al. (2018) suggest that misinformation spread is attributed more to human behavior as opposed to social media bots, it could also be worth examining the utility of bots to counter-balance the typically higher volume of misinformation posts prevalent within these discourses by using “public health promotion” bots to more broadly disseminate accurate scientific information, especially since previous evidence shows that bots can still have an impact on online political discussions (Bessi and Ferrara, 2016). However, the messaging for these approaches should be met with caution, as past research has shown that users who follow bots with opposite political ideologies to their own beliefs can instead lead to increased polarization (Bail et al., 2018). Since public health discussions related to COVID-19 have become highly politicized, it is recommended that if public health bots are utilized as a counter-misinformation strategy, their deployment should focus messaging on how it pertains to the health and well-being of users, their families, and their communities in the context of correcting COVID-19 misinformation.

Other approaches to combating misinformation include recruiting media figures, influential social media personalities/celebrities/influencers, and politicians to help communicate accurate scientific information around issues where misinformation spread is prevalent. As shown in the results from this study, influential users in misinformation-related topics were more likely to be involved in the media or politics and had higher follower counts compared to users in the *Scientific Opposition* topic, which consisted mostly of doctors or scientists. Since users who are politicians or in the media are more likely to be influential users (Dubois and Gaffney, 2014), recruiting these influential people to combat misinformation spread on social media platforms where they have a high number of followers and access to user networks could significantly increase the volume of tweets with up-to-date scientific evidence across these discourses. While it is also encouraging to see that doctors and scientists were actively engaged in spreading accurate scientific evidence, the overall influence and reach of these accounts

was extremely limited compared to pro-hydroxychloroquine users.

Specifically, there are better ways to mobilize “public health” influencers by first engaging with users who possess the needed scientific and medical source credibility, and then coupling their messages with other influential twitter accounts (e.g. politicians, media, celebrities) in order to amplify evidence-based messages that can better resonate with the public on the basis of trust in science. Medical endorsements were the most prevalent type of misinformation, indicating that people might respond more favorably towards credible *people* over credible *evidence*. Users who are scientists, researchers, or have the professional background, training or expertise to be a credible authority on an issue should consider making their credentials known when refuting misinformation or making an argument. This can also include discussing personal experiences that help inform one’s assessment. For example, instead of a statement such as “as a researcher, I believe this drug is ineffective” one could say “as a researcher who spent years running experiments on clinical drug trials, there is no strong evidence to support the use of this drug for COVID-19.” Importantly, our study found that the needed coordination and messaging among medical and public health professionals to amplify evidence-based information and combat misinformation introduced about hydroxychloroquine were largely absent following President Trump’s endorsement of misinformation.

Conclusion

The results from this study not only show that misinformation is prevalent in Twitter online discourse concerning treatment for COVID-19, but also that scientific authority in the form of promoting medical authority or questionable scientific support is evoked often to substantiate these false claims. Even with a scientific background, it can be difficult to assess the credibility of a study or statistic on social media platforms, especially if it is based in a field outside one’s domain of expertise. For the lay public that do not have scientific training, the difficulty to assess false claims made by credentialed “experts” from partisan media sources or political figures is amplified, particularly when generated from the highest places of power and influence. Future research should focus on better understanding the complex dynamics of misinformation propagation via social media networks that occurs during health crises, with particular focus on how scientific credibility acts as a basis for these claims. Combating future “viral” misinformation events will need to identify specific strategies that recognize the unique power of credibility and authority to

be misappropriated for political and non-scientific gains.

Authors' contributions

JL collected the data, all authors designed the study, conducted the data analyses, wrote the manuscript, and approved the final version of the manuscript.

Materials and data sharing agreement

The authors have published the supplementary material at the following link: https://figshare.com/projects/Scientific_Authority-related_Misinformation/101735

Ethics approval and consent to participate

Not applicable/Not required for this study. All information collected from this study was from the public domain and the study did not involve any interaction with users. User identifiable information was removed from the study results.

Declaration of conflicting interests

TKM and JL are employees of the startup company S-3 Research LLC. S-3 Research is a startup funded and currently supported by the National Institutes of Health – National Institute on Drug Abuse through a Small Business Innovation and Research contract for opioid-related social media research and technology commercialization. Author reports no other conflicts of interest associated with this manuscript and have not been asked by any organization to be named on or to submit this manuscript.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Michael Robert Haupt  <https://orcid.org/0000-0003-4985-7796>

Jiawei Li  <https://orcid.org/0000-0001-9801-4715>

Tim K Mackey  <https://orcid.org/0000-0002-2191-7833>

Supplemental Material

Supplementary material for this article is available online.

References

- Bail CA, Argyle LP, Brown TW, et al. (2018) Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115(37): 9216–9221.
- Bail CA, Brown TW and Mann M (2017) Channeling hearts and minds: Advocacy organizations, cognitive-emotional currents, and public conversation. *American Sociological Review* 82(6): 1188–1213.
- Barberá P, Jost JT, Nagler J, et al. (2015) Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science* 26(10): 1531–1542.
- Belluz J and Hoffman SJ (2013) Dr. Oz's Miraculous Medical Advice. *Slate*, 1 January. Available at: www.slate.com/articles/health_and_science/medical_examiner/2013/01/can_you_trust_dr_oz_his_medical_advice_often_conflicts_with_the_best_science.html (accessed 25 February 2021).
- Bernays EL (1928) *Propaganda*. New York: H. Liveright.
- Bessi A and Ferrara E (2016) *Social Bots Distort the 2016 US Presidential Election* Online Discussion. ID 2982233, SSRN scholarly paper. Rochester, NY: Social Science Research Network. Available at: <https://papers.ssrn.com/abstract=2982233> (accessed 7 December 2020).
- Boldrini C, Toprak M, Conti M, et al. (2018) Twitter and the press: An ego-centred analysis. In: Companion proceedings of the web conference 2018, *WWW'18*, Republic and Canton of Geneva, CHE, 23 April 2018, pp.1471–1478. Geneva: *International World Wide Web Conferences Steering Committee*.
- Brenan M (2020) Americans' reported use of face masks surges in past week. Available at: <https://news.gallup.com/poll/308678/americans-reported-face-masks-surges-past-week.aspx> (accessed 10 December 2020).
- Brossard D and Nisbet MC (2007) Deference to scientific authority among a low information public: Understanding U.S. *International Journal of Public Opinion Research* 19(1): 24–52.
- Cathey L (2020) Timeline: Tracking trump alongside scientific developments on hydroxychloroquine. *ABC News*. Available at: <https://abcnews.go.com/Health/timeline-tracking-trump-alongside-scientific-developments-hydroxychloroquine/story?id=72170553> (accessed 9 December 2020).
- Celebrity Net Worth (2019) Dr Mehmet Oz net worth. Available at: www.celebritynetworth.com/richest-businessmen/dr-oz-net-worth/ (accessed 25 February 2021).
- CNBC (2020) Nearly 200 legislators have tested positive for coronavirus nationwide and four have died as GOP flouts rules. Available at: www.cnn.com/2020/11/21/many-gop-lawmakers-disobeying-mask-wearing-rules-in-their-own-capitols.html (accessed 10 December 2020).
- Dubois E and Gaffney D (2014) The multiple facets of influence: Identifying political influentials and opinion leaders on twitter. *American Behavioral Scientist* 58(10): 1260–1277.
- Einstein KL and Glick DM (2015) Do I Think BLS Data are BS? The Consequences of Conspiracy Theories. *Political Behavior* 37(3): 679–701. DOI: 10.1007/s11109-014-9287-z.
- Frederick S (2005) Cognitive reflection and decision making. *Journal of Economic Perspectives* 19(4): 25–42.
- Frenkel S and Alba D (2020) Misleading virus video, pushed by the Trumps, Spreads Online. *The New York Times*, 28 July. Available at: www.nytimes.com/2020/07/28/technology/virus-video-trump.html (accessed 9 December 2020).
- Hamby A, Ecker U and Brinberg D (2020) How stories in memory perpetuate the continued influence of false information. *Journal of Consumer Psychology* 30(2): 240–259.
- Hasher L, Goldstein D and Toppino T (1977) Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior* 16(1): 107–112.

- Haupt MR, Jinich-Diamant A, Li J, et al. (2021) Characterizing twitter user topics and communication network dynamics of the “liberate” movement during COVID-19 using unsupervised machine learning and social network analysis. *Online Social Networks and Media* 21: 100114.
- Henkel LA and Mattson ME (2011) From dreams to psychosis: A European science foundation exploratory workshop. *Reading is Believing: The Truth Effect and Source Credibility. Consciousness and Cognition* 20(4): 1705–1721.
- Islam MS, Sarkar T, Khan SH, et al. (2020) COVID-19-related infodemic and its impact on public health: A global social media analysis. *The American Journal of Tropical Medicine and Hygiene* 103(4): 1621–1629.
- Johnson HM and Seifert CM (1994) Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20(6): 1420–1436.
- Kahan DM (2017) *Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition*. ID 2973067, SSRN Scholarly Paper, 24 May. Rochester, NY: Social Science Research Network.
- Kalyanam J, Katsuki T, Lanckriet GRG, et al. (2017) Exploring trends of nonmedical use of prescription drugs and polydrug abuse in the Twittersphere using unsupervised machine learning. *Addictive Behaviors* 65: 289–295.
- Korownyk C, Kolber MR, McCormack J, et al. (2014) Televised medical talk shows – What they recommend and the evidence to support their recommendations: A prospective observational study. *BMJ* 349: g7346.
- Kuklinski JH, Quirk PJ, Jerit J, et al. (2000) Misinformation and the Currency of Democratic Citizenship. *Journal of Politics* 62(3): 790–816. DOI: <https://doi.org/10.1111/0022-3816.00033>.
- Lee C-J and Scheufele DA (2006) The influence of knowledge and deference toward scientific authority: A media effects model for public attitudes toward. *Journalism & Mass Communication Quarterly* 83(4): 819–834.
- Lewandowsky S, Cook J, Ecker UK, et al. (2020) Under the Hood of the Debunking Handbook 2020: A Consensus-Based Handbook of Recommendations for Correcting or Preventing Misinformation. <https://www.climatechange-communication.org/debunking-handbook-2020/>
- Lu J (2020) Themes and evolution of misinformation during the early phases of the COVID-19 outbreak in China – An application of the crisis and emergency risk communication model. *Frontiers in Communication* 5: 57.
- Mackey TK (2020) Risk of cyber superspreaders and influencers on product integrity: COVID-19 and beyond. Available at: <https://bpp.msu.edu/magazine/risk-of-superspreaders-and-influencers-on-product-integrity-covid19-and-beyond-december2020/> (accessed 1 March 2021).
- Mackey TK, Li J, Purushothaman V, et al. (2020a) Big data, natural language processing, and deep learning to detect and characterize illicit COVID-19 product sales: Infoveillance study on twitter and Instagram. *JMIR Public Health and Surveillance* 6(3): e20794.
- Mackey TK, Purushothaman V, Haupt M, et al. (2021) Application of unsupervised machine learning to identify and characterise hydroxychloroquine misinformation on twitter. *The Lancet Digital Health* 3(2): e72–e75.
- Mackey T, Purushothaman V, Li J, et al. (2020b) Machine learning to detect self-reporting of symptoms. *JMIR Public Health and Surveillance* 6(2): e19509.
- Mahase E (2020) Covid-19: WHO halts hydroxychloroquine trial to review links with increased mortality risk. *BMJ* 369.
- Miller Z (2020) Trump threatens to fire Fauci in rift with disease expert. Available at: <https://apnews.com/article/trump-threatens-fire-fauci-rift-disease-57c804db048aa7f1c99f227b495f52e6> (accessed 10 December 2020).
- Morrison S and Heilweil R (2020) How Trump and his son helped make a Covid-19 conspiracy theorist go viral in a matter of hours. *Vox*, 28 July. Available at: www.vox.com/recode/2020/7/28/21345008/donald-trump-jr-twitter-suspension-stella-immanuel-hydroxychloroquine-misinformation (accessed 9 December 2020).
- Mosleh M, Pennycook G, Arechar AA, et al. (2021) Cognitive reflection correlates with behavior on Twitter. *Nature Communications* 12(1): 921.
- National Institutes of Health (NIH) (2020) NIH halts clinical trial of hydroxychloroquine. Available at: www.nih.gov/news-events/news-releases/nih-halts-clinical-trial-hydroxychloroquine (accessed 10 December 2020).
- Niburski K and Niburski O (2020) Impact of trump’s promotion of unproven COVID-19 treatments and subsequent internet trends: Observational study. *Journal of Medical Internet Research* 22(11): e20044.
- O’Connor C and Weatherall JO (2019) The evangelization of peoples. In: *The Misinformation Age. How False Beliefs Spread*. London: Yale University Press, pp.93–146.
- Papakyriakopoulos O, Serrano JCM and Hegelich S (2020) Political communication on social media: A tale of hyperactive users and bias in recommender systems. *Online Social Networks and Media* 15: 100058.
- Pennycook G, McPhetres J, Zhang Y, et al. (2020) Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science* 31(7): 770–780.
- Pennycook G and Rand DG (2019) Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* 188: 39–50.
- Pennycook G and Rand DG (2020) Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality* 88(2): 185–200.
- Poland GA and Spier R (2010) Fear, misinformation, and innuendoes: How the Wakefield paper, the press, and advocacy groups damaged the public health. *Vaccine* 28 (12): 2361–2362. DOI: 10.1016/j.vaccine.2010.02.052.
- Self WH, Semler MW, Leither LM, et al. (2020) Effect of hydroxychloroquine on clinical status at 14 days in

- hospitalized patients with COVID-19: A randomized clinical trial. *JAMA* 324(21): 2165–2176.
- Shin J, Jian L, Driscoll K, et al. (2018) The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior* 83: 278–287.
- Stanford Cable TV News Analyzer (n.d.) Stanford Cable TV News Analyzer – about us. Available at: <https://tvnews.stanford.edu/about> (accessed 25 February 2021).
- Torabi Asr F and Taboada M (2019) Big data and quality data for fake news and misinformation detection. *Big Data & Society* 6(1): 2053951719843310.
- Vosoughi S, Roy D and Aral S (2018) The spread of true and false news online. *Science* 359(6380): 1146–1151.
- Walton D and Zhang N (2013) The epistemology of scientific evidence. *Artificial Intelligence and Law* 21(2): 173–219.
- Yan X, Guo J, Lan Y, et al. (2013) A biterm topic model for short texts. In: *Proceedings of the 22nd international conference on World Wide Web, WWW'13*. New York, NY, USA, 13 May 2013, pp.1445–1456. New York: Association for Computing Machinery.