"COVID19 is_": The Perpetuation of Coronavirus Conspiracy Theories via Google Autocomplete

Daniel Houli
Rutgers, the State University of
New Jersey, USA
daniel.houli@rutgers.edu

Marie L. Radford
Rutgers, the State University of
New Jersey, USA
mradford@rutgers.edu

Vivek K. Singh
Rutgers, the State University of
New Jersey, USA
v.singh@rutgers.edu

ABSTRACT

As the impact of the COVID-19 pandemic grew in 2020, uncertainty surrounding its origins and nature led to widespread conspiracy-related theories (CRT). Use of technological platforms enabled the rapid and exponential dissemination of COVID-19 CRT. This study applies social contagion theory to examine how Google Autocomplete (GA) propagates and perpetuates these CRT. An in-house software program, Autocomplete Search Logging Tool (ASLT) captured a snapshot of GA COVID-19 related searches early in the pandemic (from March to May 2020) across 76 randomly-selected countries to gain insight into search behaviors around the world. Analysis identified 15 keywords relating to COVID-19 CRT predictions and demonstrate how searches across different countries received varying degrees of GA predictions. When grouped with similar keywords, two major categories were identified "Man-Made Biological Weapon" (42%, n=2,111), and "Questioning Reality/Severity of COVID-19" (44%, n=2,224). This investigation is also among the first to apply social contagion theory to autocomplete applications and can be used in future research to explain and perhaps mitigate the spread of CRT.

KEYWORDS

COVID-19, coronavirus, Google Autocomplete, conspiracy theory, social contagion.

INTRODUCTION

The COVID-19 pandemic of 2020 rapidly changed the daily lives of people across the globe. It also spread fear. Differing stories about its origin, cures, and appropriate actions for COVID-19 dominated the news and social media platforms (Shahsavari, Holur, Tangherlini & Roychowdhury, 2020). Stemming from this uncertainty, COVID-19 conspiracy-related theories (CRT) arose. In general, a conspiracy theory helps explain major events involving a stealthy plot by a dominant or malicious group (Douglas, 2017). The goal of this research is to understand how Google Autocomplete (GA) is playing a role in propagating COVID-19 CRT through the lens of social contagion theory, taking a geospatial perspective across 76 countries.

It is important to study COVID-19 CRT because conspiracy-related storytelling can spread incorrect or misleading information that helps solidify others' beliefs and inspire individuals to take action (Shahsavari, et al., 2020). Once exposed to CRT, individual's lines of thought may be influenced, even though previously they would perhaps have never considered these ideas. Overall, promoting conspiracy theories can potentially induce long-lasting cynicism toward the government or ruling entities by engaging viewers in messaging and increasing the theory's credibility (Kim & Cao, 2016). Social media is also often used to discuss and promote conspiracy-beliefs (Goreis & Kothgassner, 2020). These technological platforms can potentially spread COVID-19 CRT throughout the population rapidly and exponentially as the popularity of trending posts is frequently more important than facts (Allem, 2020). For comparison, there have been over 52 million posts from conspiracy-related sites regarding COVID-19, while the World Health Organization (WHO) and Center for Disease Control (CDC) released several hundred thousand posts (Mian & Khan, 2020). Additionally, COVID-19 CRT tweets such as #FilmYourHospital gained popularity when going viral in late March (Gruzd & Mai, 2020). These COVID-19 related CRT have become so mainstream that they have been discussed and debated by the news media and politicians (Stephens, 2020).

Aside from research in social media platforms, less obvious ways COVID-19 CRT have been promoted have not yet been explored in-depth. This study seeks to gain insight into the propagation of COVID-19 CRT worldwide through examining GA. GA works by using results known as "predictions" to provide shortcuts for completing a search term (Roy & Ayalon, 2020). When a user begins typing a term in Google's search box, a list of GA predictions (search terms) appears below the box. These search terms can then be easily clicked on by the person initiating the search, although they may or may not have been what the individual intended prior to seeing the "prediction." According to Google's public liaison for search, GA results are called predictions rather than suggestions: "Autocomplete is designed to help people complete a search they were intending to do, not to suggest new types of searches to be performed" (Sullivan, 2018). It is difficult to separate GA's algorithms from the intention of the searcher. However,

according to Roy and Ayalon (2020) "Google autocomplete displays predictors based on actual searches carried out by a population, with some variations owing to a user's personal search history, current events, language, and geographical location. Therefore, Google is a mere reflection of what people wish to know in the first place (p. 1026)." GA thus captures, reflects, and amplifies a population's interests and may thus influence society's perceptions and beliefs (Roy & Ayalon, 2020). GA may act as a subtle yet useful looking glass reflecting trending phrases generated by members of society, thus allowing researchers to gain insight into individual or group perceptions and beliefs to understand what COVID-19 CRT exist and how they are spread.

These predictions may be intended to be a shortcut to complete a user's search. However, in making search predictions immediately visible, GA may influence a person's thought process, curiosity, or knowledge of a particular subject (Roy & Ayalon, 2020), leading the user down search paths that they may not have otherwise discovered. Roy and Ayalon (2020) assert that Google plays a "dual role," first by storing existing attitudes and beliefs, and then by acting as a reflection of this information, reinforcing these beliefs. To take it a step further, GA not only reinforces these beliefs but also spreads them to new individuals, perhaps unintentionally, as users focus on popularity instead of accuracy, as is the case with social media (Allem, 2020).

In addition, this study is significant as it is among the first to apply social contagion theory (Marsden, 1998) as a framework to examine COVID-19 CRT predictions. Social scientific research has shown that beliefs, attitudes, and behaviors can be spread through communities similar to the way an infectious disease passes from one person to another. In social contagion, users actively seek out and propagate the information (Hodas & Lerman, 2014). However, unlike social media, when using search engines users may not be aware of how information (or misinformation) is being propagated through autocomplete. According to Miller and Record (2017), an individual search can minimally influence the autocomplete algorithm, but large groups of individuals in a given region have the ability to trigger the algorithm, making the predictions trend. "Autocomplete inevitably and irreparably induces changes in users' epistemic actions, particularly their inquiry and belief formation. While these changes have potential epistemic benefits for searchers and society, they also have harms, such as generating false, biased, or skewed beliefs about individuals or members of disempowered groups" (Miller & Record, 2017, p. 1959). The authors acknowledge that we are considering GA as a social process, yet recognize that GA predictions are, of course, inherently tied to Google's algorithms.

LITERATURE REVIEW
Conspiracy-Related Theories and COVID-19
Joseph Uscinski of the University of Miami says that CRT "are generally characterized by acceptance of the following propositions: Our lives are controlled by plots hatched in secret places. Although we ostensibly live in a democracy, a small group of people run everything, but we don't know who they are. When big events occur – pandemics, recessions, wars, terrorist attacks – it is because that secretive group is working against the rest of us" (LaFrance, 2020, p. 37). CRT typically gain traction during times of societal crises (Bierwiaczonek, Kunst & Pich, 2020), and can be defined, measured, and observed, making them suitable for scientific research (Freeman & Bentall, 2017).

Over the past year, there has been a large increase of COVID-19 related CRT. Researchers have examined these theories from multiple, interrelated perspectives, including mental health (Bento, et.al., 2020; Sallam, et al., 2020); behavior and trust (Imhoff & Lamberty, 2020; Karić & Međedović, 2020; Freeman et al., 2020; Jovančević & Milićević, 2020) and social media (Ahmed, Vidal-Aladall, Downing, & Seguí, 2020; Allem, 2020; Bruns, Harrington, & Hurcombe, 2020; Gruzd & Mai, 2020; and Stephens, 2020). An underlying theme found in much of this research centers on misinformation and the inherent dangers of CRT (e.g., see Bierwaiczonek et al., 2020). Generally, conspiracy theories are excluded from the possibility of being a candidate of truth (Bjerg & Presskorn-Thygesen, 2017). Nevertheless, "...some conspiracy theories can in some case have a grain of truth to them, or be based on some facts, and then be wildly exaggerated" (Bartlett & Miller, 2010, p. 4). There may also be some truth specifically within COVID-19 CRT (Bernard, Akaito, Joseph, & David, 2020). Indeed, President Biden has ordered U.S. intelligence to investigate the possibility that COVID-19 was accidentally leaked, focusing particularly on laboratories located in Wuhan, China (Shear, Barnes, Zimmer & Mueller, 2021). However, to prevent the spread of misinformation, large technology companies have sought to remove or block COVID-19 CRT, including Google. A blog post regarding COVID-19 in March, from Google CEO Sundar Pichai (2020) states: "Our Trust and Safety team has been working around the clock and across the globe to safeguard our users from phishing, conspiracy theories, malware, and misinformation, and we are constantly on the lookout for new threats." Yet, as shown by the data analysis below, much proliferation occurred during the early months of the pandemic.

Social Contagion Theory
Originally, the term "contagion" was grounded in research to help understand the spread of epidemic diseases. Within the literature of social contagion theory, the term has been adapted to examine the phenomena of how

information can spread unintentionally through communities, more like an infectious disease than like a rational choice (Marsden, 1998). This theory has been applied in the study of various networked communities (Ugander, Backstrom, Marlow & Kleinberg, 2012; Scherer and Cho, 2003). The theory fundamentally rests on the interactions within these communities (Iacopini, Petri, Barrat & Latora, 2019), and has been applied in analysis of politics, fads, public opinion, and diffusion of new technologies (Ugander, et al., 2012). "Contagion theory suggests that those individuals who are most connected to each other through interpersonal contacts are also most likely to share similar information, attitudes, beliefs, and behaviors on controversial topics. Conversely, individuals who are not in frequent contact are less likely to have the same information about a topic, and are less likely to share similar attitudes and beliefs" (Scherer and Cho, 2003, p. 266). Although there have been multiple vehicles for observing the effects of social contagion, a popularly studied area is within social media.

Social media relies on the deliberate propagation of information content created by system users and consumed by other users (Hodas & Lerman, 2014). If the information is then amplified to many others, this spreading of information is known as social contagion. Research on social media platforms, such as Facebook, "provide natural sources of data for studying social contagion processes" (Ugander, et al., 2012, p. 5965). Studies of other social media sites such as Twitter have also demonstrated how users are exposed to and spread information through methods such as retweets (Lerman, Ghosh, & Surachawala, 2012). For example, if the information reaches a friend on social media who retweets or forwards to many friends, then it may rapidly spread with one consumer of information becoming many (Vishwanath, 2015). In the case of this GA research, the communication is mediated and amplified to become socially contagious through a search engine, as discussed below.

Google Autocomplete (GA)

The autocompletion feature is present in many different types of search engines and provides possible keywords or relevant sentence completion to help the user complete their task quicker. In 2004, Google released "Google Suggest," a form of autocompletion software (Ward, Hahn & Feist, 2012). It was initially created to assist those with disabilities by reducing keystrokes to complete a search (Popyer, 2016). Google soon realized its potential, and currently this feature, now called Google Autocomplete (GA), is available to all users (See Figure 1).
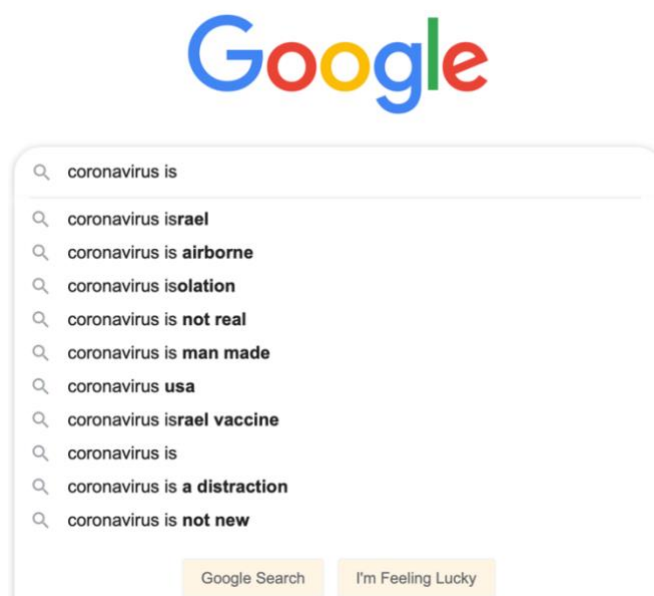

Figure 1. Google Autocomplete Predictions for "Coronavirus is…" (Retrieved on 3/30/20)

Although Google provides general information on how GA functions, details of its inner workings are proprietary. One factor influencing GA is popular, or "trending" searches by communities in certain areas, or world-wide at specific periods. As noted above, research indicates that GA captures, mirrors, and amplifies a population's interests, and via searching predictions, can influence society's perceptions and beliefs (Roy & Ayalon, 2020). Loh (2016) refers to GA as a "double-edged sword." On the one hand, it can benefit users by allowing them to find information quickly and with fewer keystrokes. Also, GA can provide predictions to a user that are relevant and targeted to their search needs (Popyer, 2016), which has been found to improve search quality, mainly when users have limited knowledge of a specific topic (Ward, Hahn & Feist, 2012), thereby enabling individuals to become self-informed through search. On the other hand, GA can also promote incorrect, misleading, or potentially harmful information (Loh, 2016). This contrast, coupled with the power of social influence that this information has, makes autocomplete a critically important area of study. Furthermore, GA has historically produced harmful results that have been seen

as racist, sexist, and/or homophobic (Baker & Potts, 2013). Noble (2018) has declared search engines to be "algorithms of oppression" and has advocated that these harmful Google results, including GA predictions, should provoke national discussion and action to reveal and ameliorate their effects. Google has been involved in lawsuits such as defamation of character-based content results related to GA (Popyer, 2016). Google has acknowledged that these particular results and predictions are a problem and allows users to report inappropriate GA predictions via their website. However, according to Google, these computer-generated results are not the company's fault (Noble, 2018). Therefore, Google has maintained that they should not be held liable as the results are based on algorithms using data created by users so that results are not created by the company itself (Morozov, 2011).

Though designed to predict users' search interest, based on the previous literature, GA may influence a user's thoughts, views, and beliefs (Roy & Ayalon, 2020). It can be argued that Google would otherwise not put such an effort into preventing certain GA predictions from being displayed. Furthermore, according to social contagion theory, a "simple exposure" can be enough for social transmission to occur (Marsden, 1998). Following the above literature review the following research questions were developed:

RQ1. How can GA be used as a methodological tool to understand the spread of COVID-19 CRT among google users during the pandemic?

RQ2. What are the variations in the amount of COVID-19 CRT predictions among Google users across the world?

RQ3. How does social contagion theory help to explain the spread of COVID-19 CRT through GA?

METHOD
To gather data for this study, the authors created Autocomplete Search Logging Tool (ASLT). During its design, the authors referred to Google's support page (See Figure 2) as this would provide the most accurate and up-to-date information regarding how GA functions.

Where search predictions come from

- Search terms you type.
- Relevant searches you've done in the past, if you're signed in to your Google Account and have Web & App Activity turned on. Learn how to see & control your Web & App Activity
- What other people are searching for, including trending searches. Trending searches are popular topics in your area that change during the day and aren't related to your search history. To get current trending searches, go to Google Trends ⌴ .

Search predictions aren't the answer to your search. They also aren't statements by other people or Google about your search terms.

Figure 2. Google Autocomplete Help Center
https://support.google.com/websearch/answer/106230 (Retrieved on 3/30/20)

From this description, the authors derived three main areas of focus: a) user's search terms, b) user's past searches (assuming the user is logged in), and c) similar or trending searches by other users. Additionally, the findings could be compared to results on Google Trends. Thus, ASLT was created with these three areas in mind and functions by taking input of a specific key term; for instance, "COVID-19 is." Then, to show variation across geo-location, ASLT accepts a specific country location code (i.e., "US"). Also, to capture variation across time, ASLT was used to collect sample data for 35 days, during the period from March 25 to May 14, 2020.

The queries spanned 76 randomly selected countries. ASLT collected the data in English, as this appeared to be a suitable language across multiple domains (Mazières & Huron, 2013). Lastly, to avoid the influence of researchers' "searches done in the past," ASLT was run without logging in to Google via a direct API (application programming interface) from the AWS cloud server.

ASLT was created in Python 3 and run on Amazon Web Services (AWS) SageMaker machine learning platform. ASLT iterated through each country and term, returning GA predictions along with the associated country (e.g., France (FR), see Figure 3). The results over the three months of data collection were stored in an AWS storage bucket for analysis.

```
'covid-19 is working from home really the new normal the economist FR',
'covid-19 is working from home really the new normal FR',
'covid-19 is there a disproportionate burden on working classes in france FR',
'covid 19 is caused by FR',
'covid 19 is it slowing down FR',
```

Figure 3. ASLT Output Example for France (FR) (Retrieved on 5/19/20)

DATA ANALYSIS

During the analysis phase, the data was retrieved from the AWS storage buckets and cleaned. Cleaning involved removing data that was not relevant so as to include only COVID-19 related statements, such as "coronavirus is…" and "COVID-19 is…" Next, the data was organized and separated into three columns: a) Phrase (autocomplete result), b) Country, and c) Date. The data was then ready for the authors to tag with the COVID-19 CRT keywords. In order to do so, "Conspiracy" keywords were defined by four characteristics which focus primarily on currently unproven ideas (Freeman & Bentall, 2017; Freeman, 2020). These four characteristics include 1) the world or an event is not as it seems; 2) there appears to be a cover-up by a powerful entity; the believer's explanation of the event is accepted only by a minority; or d) the majority of evidence provided does not support the explanation (Freeman, 2020). The COVID-19 CRT keywords were then used in conjunction with Python's Natural Language Toolkit (NLTK) to help identify conspiracy-related phrases. At the end of the three-month data collection period, analysis of keywords was conducted in a repeated and iterative fashion, until the authors were satisfied that all relevant keywords were included (Saunders et al., 2018).

Once an exhaustive list of COVID-19 CRT keywords was identified, like-words were grouped together that could fit into larger topic clusters. For instance, "weapon," "bioweapon," and "biological weapon" predictions were all grouped under the term "Weapon." This task was completed to reduce redundancy in meaning while, at the same time, identifying major CRT-related topics. A final list of COVID-19 CRT-related keywords of 15 terms was produced (See Table 1) and results were imported into Tableau Desktop for visualization and further analysis.

Man-Made	Over-Blown	Made in Lab
Not Bad	Lie	War
Weapon	Cover-Up	Political
Distraction	5G-related	Manipulation
Not Real	The End	Conspiracy

Table 1. COVID-19 CRT-Related Keywords List

RESULTS

Over the three months of data collection in which data for 35 days were sampled, 21,133 COVID-19-related results were collected in the AWS storage buckets from GA. Of these, 5,056 were COVID-19 CRT related, when users entered any of these three search terms: "coronavirus is…", "corona virus is…", and "COVID-19 is…" As shown in Table 2, the most frequent predictions returned by GA were assigned to the keywords "Man-Made," "Not Bad," "Weapon," "Distraction," and "Not Real."

Keyword	Count & Percentage (N=5,056)
Man-Made	1,421 (28%) *
Not Bad	979 (19%)
Weapon	690 (14%)
Distraction	487 (10%)
Not Real	463 (9%)
Over-Blown	295 (6%)
Lie	215 (4%)
Cover-Up	124 (2%)
5G-related	102 (2%)
The End	94 (2%)
Made in Lab	71 (1%)
War	48 (<1%)
Political	32 (<1%)
Manipulation	20 (<1%)
Conspiracy	15 (<1%)

Table 2. Counts and Percentages of Keywords
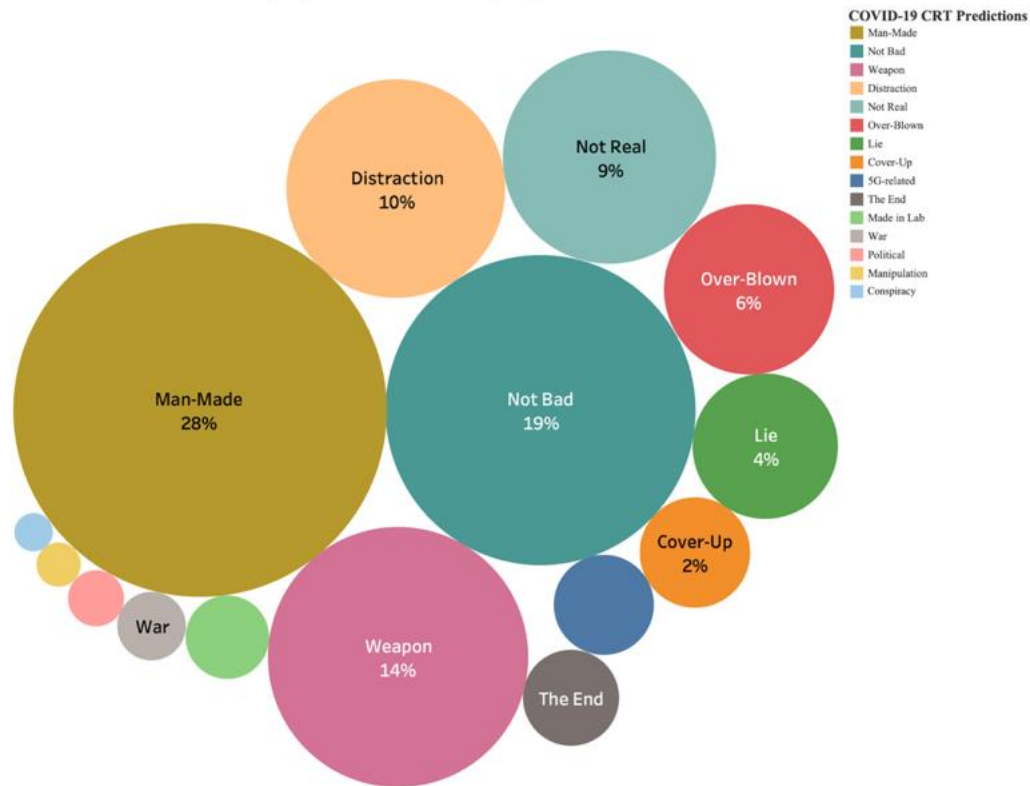*Note: Percentages are rounded up to the next whole number.

Figure 4. Relative Frequency of COVID-19 CRT-Related Keywords
Note: The smaller circles without percentages represent keywords below 2% of the total (see also Table 2).

These results indicate the most frequent among the 5,056 COVID-19 CRT-related keywords that were assigned to GA predictions across the 76 country codes. As shown in Table 2 and Figure 4, "Man-Made" was the keyword with the highest percentage (28%, n=1,421). If connected to the CRT belief that the virus was developed in a lab as a biological weapon, "Man-Made" could be combined with the "Weapon" (14 %, n=690), into a larger category "Man-Made Biological Weapon" which would account for a total of 42% (n=2,111). Additionally, four of the top five keywords related to questioning the reality and/or severity of the virus: "Not Bad" (19%, n=979), "Distraction" (10%, n=487), "Not Real" (9%, n=463), and "Over-Blown" (6%, n=295) could be combined into a second larger category "Questioning Reality/Severity of COVID-19," accounting for a total of 44% (n=2,224).

As graphically depicted in Figure 5, data analysis also revealed the variations of GA predictions across the seventy-six countries studied, indicating which areas, from a geospatial perspective, have higher COVID-19 CRT predictions than others. Figure 5 captures a snapshot of the worldwide dissemination of pandemic-related CRT predictions during the three-month data collection period. As noted above, the COVID-19 CRT-related keywords (N=5,056) had also been organized and separated by country, allowing the authors to determine the sum of these keywords for each country. Next, this number was divided by the sum of the total number of all predictions for each country.

The maximum percentage of COVID-19 CRT compared to all predictions per country was 38%, the minimum percentage was 0%, with a mean of approximately 17%. The value of n was different across each country, for instance in Australia the number of all COVID-19 predictions was 717, the number of COVID-19 keywords was 274 (38%). Whereas in Russia the number of all COVID-19 predictions was 459, the number of COVID-19 keywords were 95 (21%). Note that the 13 countries at the maximum value of 38% had the same number of total predictions (717) and COVID-19 keywords (274). It is difficult to interpret why these results were found. One possible explanation is that Google may be using the same parameters for their algorithm for those 13 countries. If this is the case, it is a limitation of this study, as we do not have access to Google's proprietary algorithms.
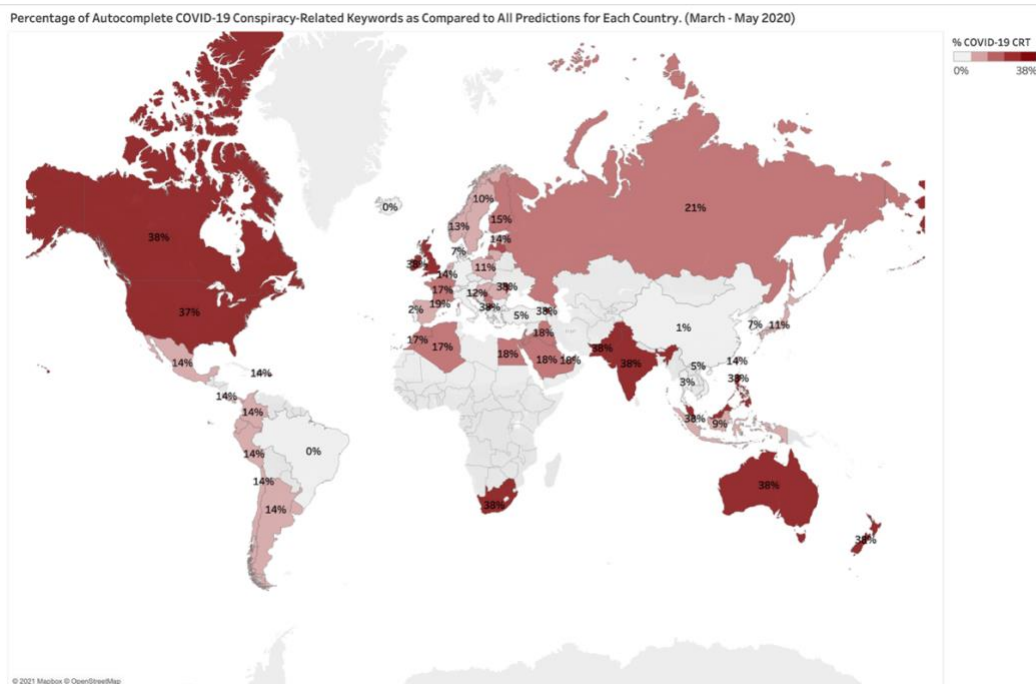
Figure 5. Percent of COVID-19 Conspiracy-Related Keywords Compared to All Predictions Per Country

Overall, Great Britain, Ireland, Australia, the United States, India, and South Africa and some of the surrounding areas showed higher percentages of COVID-19 CRT predictions as compared to all predictions within these countries. Multiple countries in South America showed COVID-19 CRT predictions returned at 14%, also indicating that these countries returned the same predictions. Unlike other countries in South America, Brazil returned 0% of the predictions, as did Iceland, Germany, and Italy and several other countries (note that Figure 5 does not display all 0% values). Google Transparency Report does not give insight into why these countries returned 0% of predictions. One can get a glimpse of the challenges faced by the country officials, as seen in Figure 6, which displays one of Google's Transparency Reports for India (January 2020 - June 2020), sending multiple requests for the removal of content which includes conspiracy theories related to COVID-19 along with "news reports and criticism of the authorities' handling of the pandemic."
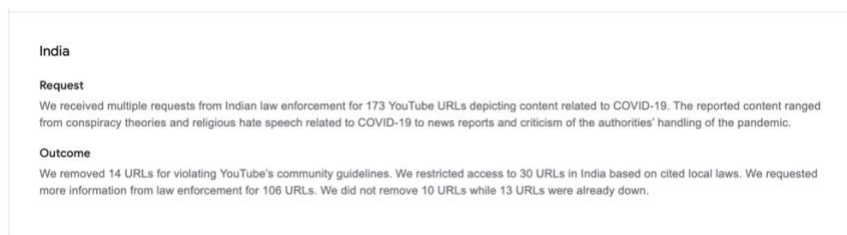

Figure 6. Jan 2020 - Jun 2020 Google Transparency Report India
https://transparencyreport.google.com/government-removals/by-
country/IN?hl=en&country_breakdown=period:Y2020H1;country:IN&lu=country_request_explore
& country_request_explore=period:;authority:IN;p:2 (Retrieved on 4/12/21)

As previously mentioned, Google refers users to consult Google Trends (see Figure 2, above) to understand what is trending or popular in a given location and period. The authors referenced Google Trends to validate a subset of the above findings. In Figure 7, Google Trends displays the top three COVID-19 CRT-related percentages of predictions found in this study; "COVID-19 is man-made," "COVID-19 is not bad," "COVID-19 is a weapon." Since multiple countries could not be specified, the scope of the Google Trends search was worldwide.
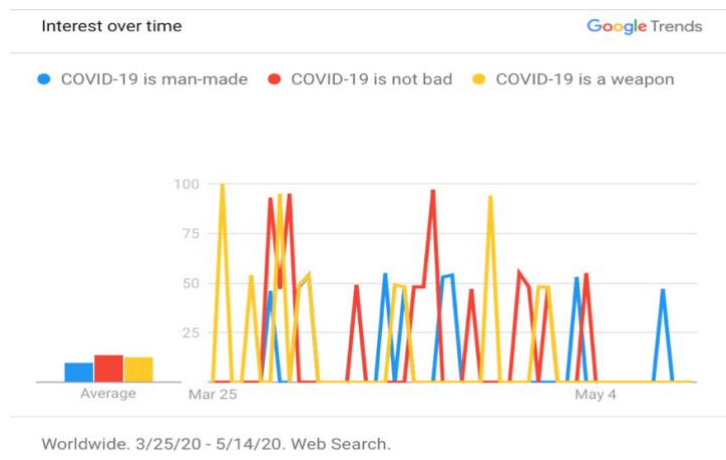
Figure 7. World-Wide Google Trends view from March – May 2020
https://trends.google.com/trends/explore?date=2020-03-25%202020-05-14&q=COVID-19%20is%20man-made,COVID-19%20is%20not%20bad,COVID-19%20is%20a%20weapon (Retrieved on 1/5/21)

The results displayed in Figure 7 for Google Trends differ slightly from what was reported in the authors' analysis. Google's scale represents the peak of each search term's popularity ranging from 0% to 100%. At different times "COVID-19 is a weapon" and "COVID-19 is not bad" approached that number. Interestingly, "COVID-19 is man-made" hovered around the 50% area. Results from ASLT showed "COVID-19 is man-made" was the highest average over the 76 countries. This difference can be due to Google Trends' result being worldwide and thus more encompassing than results from the 76 countries selected for this study. Alternatively, this can result from Google's statement that trending data is only one factor in considering GA prediction results. In any case, these results from Google Trends support this study's findings that these COVID-19 CRT-related phrases, and especially the two larger categories identified in the author's analysis, were popular or trending during this timeframe.

LIMITATIONS

This study has several limitations. As noted above, the authors have chosen to characterize GA as both a social process and an algorithm (Roy & Alayon, 2020) in applying social cognition theory to aid our understanding of results. Additionally, although the study does give perspective over multiple countries worldwide, studying 76 countries limits a full worldview. The frequency of collection may have potentially impacted the percentages as the data was collected over a three-month timeframe, sampling a total of 35 days, running at various times of the day and on different days of the week. However, the duration from the May to March timeframe appears appropriate as mentioned previously as after this period Google is actively seeking to "safeguard" from threats such as conspiracy theories (Pichai, 2020).

For transparency, ASLT collected data from the same AWS (server) location (US East 1/N. Virginia). However, this server was not logged into a google account. The API used by ASLT, which allows a point-to-point connection from AWS to Google, searched for dynamic content (GA predictions). Additionally, the URL specified the country code (i.e., FR = France), thereby "spoofing" the search location, as it was not possible to travel to different locations to conduct the searches during the pandemic. For these reasons, to the authors' knowledge, client-side caching, remembering previous searches for search predictions, was not a factor. As noted above, there is also the limitation that results for 13 countries had the same number of predictions and keywords (38%) which cannot be attributed to mere chance, so that interpreting the data for these countries is limited and Google has not provided full information about their algorithms.

As Mazieres and Huron (2013) noted, although English appears to be a prevalent and fitting language to use while studying GA, this limits data collected from non-western, non-English speaking countries. Although Google is the most used search engine globally, it has below 40% market share in China, Japan, and Russia (Mazieres & Huron, 2013). Therefore, it is recognized that Google search is limited in these countries and does not entirely reflect all of the population's interests.

DISCUSSION

The analysis of keywords from ASLT regarding predictions identified two main categories. One of these, Man-Made Biological Weapon (42%, n=2,111) resonates with Stephen's (2020) finding that one third of Americans believe some variant of the theory that COVID-19 is constructed as a bioweapon. Similarly, in Great Britain, almost half of surveyed individuals "endorsed" a statement that the coronavirus was a bioweapon developed by China to abolish the West (Freeman, et al.,2020). Results regarding GA predictions also coincide with popular stories

indicating that the virus was a side-effect of the 5G network (Shahsavari, et al., 2020), building upon the narrative regarding the perceived negative environmental and health effects resulting from the 5G rollout (Bruns, Harrington, & Hurcombe, 2020). The second major category, Questioning Reality/Severity of COVID-19 (44%, n=2,224) is in line with Fuchs (2020), who asserts that a downplay or denial of the seriousness of COVID-19 is an ideological dimension of these CRT. During analysis the authors observed some similarities, differences, and contradictions in the COVID-19 CRT GA predictions. Even if one believed that COVID-19 CRT had some factual components, a level of misinformation within these predictions was evident. For instance, COVID-19 cannot be a "bio-weapon," yet at the same time, "not real." Consequently, being presented with misinformation from GA predictions can threaten public health by confusing individuals about factual information (Roozenbeek et al., 2020), thereby further dividing society members, especially in today's political climate (Mian & Khan, 2020). Belief in pandemic denial CRT have potentially harmful consequences that may result in spreading the disease (e.g., reduction of social distancing, see Bierwiaczonek, et al., 2020).

Regarding RQ1, this study explored how GA can be a useful methodological tool to trace searching behavior and to mirror societal perceptions and beliefs on a particular subject as asserted by Roy and Ayalon (2020). Knowledge of these beliefs and perceptions is critical to understanding what COVID-19 CRT exist worldwide and how they are spread. As previously mentioned, studying social media, such as Twitter or Facebook, provides venues for researchers to examine CRT. However, research using GA can also be beneficial in at least two different ways. The first is that GA is a more subtle, less obvious way to reflect users' search behavior, especially regarding what is classified as misinformation or "fake news." For the mainstream public COVID-19 CRT would fall within this category. Social media companies, such as Facebook, use their News Feed Algorithm to counter what they determine to be fake news (Andersen & Søe, 2020). This is different than GA, which Google has affirmed is user generated and as such, it is the user's responsibility to report inappropriate content, not Google's. The second is that using a software program such as ASLT combined with data analysis software such as Tableau enables researchers to get a geospatial perspective of GA predictions during a particular point in time. One may consider using Google Trends for this type of analysis, although the user must specify the keyword (i.e., "COVID-19 is not real"), so existing GA predictions are not always revealed to the researcher. Additionally, RQ 1 focused attention on the ways that GA predictions contribute to the spread of COVID-19 CRT among Google users.

Related to RQ2, results indicated that individuals from many of the 76 countries did receive GA COVID-19 CRT predictions when searching Google during these three months with many similar keyword themes across the multiple countries. The distribution was demonstrated in Figure 5. Although COVID-19 CRT must have originated from one or more specific locations, as indicated by the spread of this information worldwide, it is difficult to determine where these CRT originated. Nevertheless, as the data shows, the highest prediction percentages are in countries where there may be less data suppression (e.g., The United States, England, and Australia), similar to Stephens' (2020) CRT Twitter analysis findings. If there is less data suppression, then the data can be perpetuated by users clicking through GA predictions, thereby making them trend, which these results suggest.

RQ3 asked how the application of social contagion theory could help to explain the spread of COVID-19 CRT through GA. Previous research focused on social contagion being extensively studied across social media, centering on spreading messages such as misinformation and "fake news" (Alshaabi et al., 2020) and on networked communities (e.g., Scherer and Cho, 2003). To understand how social contagion theory can help to explain the behavior of GA, the authors examined how it was used in a somewhat similar domain, such as social media. Social contagion typically rests on the behavior or property of influential relationships at the micro-social level affecting the macro-social level (Berndt, Rodermund, & Timm, 2018). For instance, an example of social contagion as bonding can be 'sharing' or 'liking' a post on Facebook (Harrigan, Achananuparp, & Lim, 2012). Further, influential figures on social media can act as a "hub of social contagion," allowing researchers to study the diffusion of information across networks (Alshaabi et al., 2020). Social contagion theory is appropriate to apply to the current study of GA, since, according to Scherer and Cho (2003) "The idea of social contagion suggests that individuals adopt the attitudes or behaviors of others in the social network with whom they communicate. The theory does not require that there is intent to influence, or even an awareness of influence, only that communication takes place" (p. 262). So, although the direct nature of social interaction is not present in GA, the dynamics of social contagion can be shown to be present when one considers GA as a type of network setting. Although the user interacts with a "machine" to perform a search query, the GA predictions contain trending data that ties the user to a network of popular beliefs within the user's community or region and directly helps to amplify these beliefs. For even the most skeptical Google users, this search engine is still influential and its algorithms can therefore act as a hub for social contagion in accordance with the theory.

In utilizing ASLT, results suggest that the predictions presented by GA unobtrusively at the time of search helped to spread COVID-19 CRT information. As consistent with social contagion theory, these CRT represent a nonpermanent event that develops over time and space through social networking dynamics (Scherr, 2020). Social

contagion theory assists in our understanding of the phenomena occurring in GA and how they influence the search behavior of Google users. These users are actively seeking information (in this case, regarding COVID-19) but then through GAs are "discovering" CRT-related predictions and further propagating these when they click on the prediction (including amplification through Google Trends). As a result, the user is unconsciously, yet actively, contributing to a social contagion hub perpetuating COVID-19 CRT information, misinformation, and disinformation or "fake news." Rather than depending on deliberate human intention, the algorithm-driven spread of information found in GA (Hodas & Lerman, 2014) may be more determinantal as users who may not believe or support the information are unknowingly amplifying it (see also Allum, 2020).

CONCLUSION

Knowledge of how stories (such as COVID-19 CRT) spread and persist within groups is fundamental to understanding social phenomena (Alshaabi et al., 2020). This study explored using ASLT to determine and capture COVID-19 conspiracy-related predictions during the early stages of the worldwide pandemic. This software application enabled the authors to describe and visualize these predictions worldwide. Although more conventional avenues for researching the digital spread of COVID-19 CRT exist (e.g., Twitter, Facebook, YouTube, and Reddit), autocomplete is a less obvious, yet equally or potentially more valuable data source. For academics, a large part of digital research relies on sources provided by a monopoly composed of a few Big Tech companies who may keep critically important proprietary user data confidential. Newer, unconventional methods and data sources must be explored as possible sites to gain previously unknowable understanding of both user behavior and proprietary system algorithms.

The authors found that tracking GA predictions provided a snapshot across three months during the pandemic to discover insights into the spread of COVID-19 CRT information. However, unfortunately, for COVID-19 CRT this door is closing, as Google is actively removing pandemic CRT-related data. A recent preliminary test of running ASLT after this study was conducted indicated that GA returned no COVID-19 CRT across the 76 countries. The authors are in agreement with Miller and Record (2017), "First, speech should be protected on the Internet. Censorship is bad. We agree. But omitting autosuggestions is not censoring search results. The distinction is between finding information one is seeking and being involuntarily informed that certain information exists" (p. 1957). However, as one door closes another one opens and there are ample and newly forming CRT-related topics of current interest that still could be explored using GA, including those related to vaccine-related queries, QAnon (LaFrance, 2020) and more. The authors also look forward to seeing GA as a more normative exploration tool when exploring and capturing societal data and search behavior.

This research is among the first to apply social contagion theory to a study of autocomplete applications, and results demonstrate that this framework can be successfully extended to consideration of GA results world-wide or focused on particular countries. This theoretical framework could also be used for future research seeking to explain (and perhaps mitigate) the exponential spread of CRT which may well be dangerous during a pandemic and other times of global crisis. Although Google has taken action to stop the flow of COVID-19 -related CRT, this research shows that before these actions were taken, CRTs had already spread rapidly and dramatically around the globe, growing in contagion, partially due to GA, alongside of the pandemic itself.

ACKNOWLEDGMENTS

REFERENCES

Ahmed, W., Vidal-Alaball, J., Downing, J., & Seguí, F. L. (2020). COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. *Journal of Medical Internet Research, 22*(5), e19458.

Allem, J. P. (2020). Social media fuels wave of coronavirus misinformation as users focus on popularity not accuracy. Recuperado el, 8.

Alshaabi, T., Dewhurst, D. R., Minot, J. R., Arnold, M. V., Adams, J. L., Danforth, C. M., & Sheridan Dodds, P. (2020). The growing amplification of social media: Measuring temporal and social contagion dynamics for over 150 languages on Twitter for 2009--2020. *arXiv e-prints*, arXiv-2003.

Andersen, J., & Søe, S. (2020). Communicative actions we live by: The problem with fact-checking, tagging or flagging fake news – the case of Facebook. *European Journal of Communication* (London), *35*(2), 126–139. https://doi.org/10.1177/0267323119894489

Baker, P., & Potts, A. (2013). 'Why do white people have thin lips?' Google and the perpetuation of stereotypes via auto-complete search forms. *Critical Discourse Studies*, *10*(2), 187-204.

Bartlett, J., & Miller, C. (2010). *The power of unreason: Conspiracy theories, extremism and counter-terrorism*. London: Demos.

Bento, A. I., Nguyen, T., Wing, C., Lozano-Rojas, F., Ahn, Y. Y., & Simon, K. (2020). *Evidence from internet search data shows information-seeking responses to news of local COVID-19 cases.* Proceedings of the National Academy of Sciences, *117*(21), 11220-11222.

Bernard, F. O., Akaito, J. A., Joseph, I., & David, K. B. (2020). COVID-19: The trends of conspiracy theories vs. facts. *The Pan African Medical Journal, 35*(Suppl 2).

Berndt, J. O., Rodermund, S. C., & Timm, I. J. (2018, December). Social contagion of fertility: An agent-based simulation study. In *2018 Winter Simulation Conference (WSC)* (pp. 953-964). IEEE.

Bierwiaczonek, K., Kunst, J.R. and Pich, O. (2020). Belief in COVID-19 Conspiracy Theories Reduces Social Distancing over Time. *Applied Psychology: Health and Well-Being*, *12*(4), 1270-1285. https://doi.org/10.1111/aphw.12223

Bjerg, O., & Presskorn-Thygesen, T. (2017). Conspiracy Theory: Truth Claim or Language Game? *Theory, Culture & Society*, *34*(1), 137–159. https://doi.org/10.1177/0263276416657880

Bruns, A., Harrington, S., & Hurcombe, E. (2020). <? covid19?>'Corona? 5G? or both?': the dynamics of COVID-19/5G conspiracy theories on Facebook. *Media International Australia*, *177*(1), 12-29.

Douglas, S. (2017). The Psychology of Conspiracy Theories. *Current Directions in Psychological Science: a Journal of the American Psychological Society, 26*(6), 538–542. https://doi.org/10.1177/0963721417718261

Freeman, D., & Bentall, R. P. (2017). The concomitants of conspiracy concerns. *Social psychiatry and psychiatric epidemiology*, *52*(5), 595-604.

Freeman, D., Waite, F., Rosebrock, L., Petit, A., Causier, C., East, A., Jenner, L., Teale, A. L., Carr, L., Mulhall, S., Bold, E., & Lambe, S. (2020). Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychological medicine*, 1–13. Advance online publication. https://doi.org/10.1017/S0033291720001890

Fuchs, C. (2020). Everyday life and everyday communication in coronavirus capitalism. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, *18*(1), 375-399.

Goreis, A., & Kothgassner, O. D. (2020). Social Media as Vehicle for Conspiracy Beliefs on COVID-19. *Digital Psychology*, *1*(2), 36-39.

Gruzd, A., & Mai, P. (2020). Going viral: How a single tweet spawned a COVID-19 conspiracy theory on Twitter. *Big Data & Society, 7*(2), 2053951720938405.

Harrigan, N., Achananuparp, P., & Lim, E. P. (2012). Influentials, novelty, and social contagion: The viral power of average friends, close communities, and old news. *Social Networks*, *34*(4), 470-480.

Hodas, N. O., & Lerman, K. (2014). The simple rules of social contagion. *Scientific Reports, 4*, 4343.

Iacopini, I., Petri, G., Barrat, A., & Latora, V. (2019). Simplicial models of social contagion. *Nature Communications, 10*(1), 1-9.

Imhoff, R., & Lamberty, P. (2020). A Bioweapon or a Hoax? The Link Between Distinct Conspiracy Beliefs About the Coronavirus Disease (COVID-19) Outbreak and Pandemic Behavior. Social Psychological and Personality Science, 11(8), 1110–1118. https://doi.org/10.1177/1948550620934692

Jovančević, A., & Milićević, N. (2020). Optimism-pessimism, conspiracy theories and general trust as factors contributing to COVID-19 related behavior – A cross-cultural study. *Personality & Individual Differences, 167*, N.PAG. https://doi-org.proxy.libraries.rutgers.edu/10.1016/j.paid.2020.110216

Karić, T., & Međedović, J. (2021). Covid-19 conspiracy beliefs and containment-related behaviour: The role of political trust. *Personality & Individual Differences, 175*, N.PAG. https://doi-org.proxy.libraries.rutgers.edu/10.1016/j.paid.2021.110697.

Kim, M., & Cao, X. (2016). The impact of exposure to media messages promoting government conspiracy theories on distrust in the government: Evidence from a two-stage randomized experiment. *International Journal of Communication*, *10*, 20.

LaFrance, A. (2020). Nothing Can Stop What Is Coming. *Atlantic, 325*(5), 26–38.

Loh, L. C. (2016). Autocomplete: Dr Google's "helpful" assistant? *Canadian Family Physician, 62*(8), 622-623.

Lerman, K., Ghosh, R., & Surachawala, T. (2012). Social contagion: An empirical study of information spread on Digg and Twitter follower graphs. *arXiv preprint arXiv:1202.3162*.

Locher, D. A. (2002). *Collective Behavior*. Upper Saddle River, NJ: Prentice Hall

Marsden, P. (1998). Memetics and social contagion: Two sides of the same coin. *Journal of Memetics-Evolutionary Models of Information Transmission*, *2*(2), 171-185.

Mazieres, A., & Huron, S. (2013, May). Toward Google Borders. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 244-247).

Mian, A., & Khan, S. (2020). Coronavirus: The spread of misinformation. *BMC Medicine*, *18*(1), 1-2.

Miller, J. M. (2020). Do COVID-19 conspiracy theory beliefs form a monological belief system? *Canadian Journal of Political Science/Revue Canadienne de Science Politique, 53*(2), 319-326.

Miller, B., & Record, I. (2017). Responsible epistemic technologies: A social-epistemological analysis of autocompleted web search. New Media & Society, 19(12), 1945–1963. https://doi.org/. 10.1177/1461444816644805

Morozov, E. (2011). Don't be evil. *The New Republic*, *242*(11),18-24.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

Pichai, S. (2020, March 06). Coronavirus: How we're helping. Retrieved December 22, 2020, from https://blog.google/inside-google/company-announcements/coronavirus-covid19-response/

Popyer, K. (2016). Cache-22: The Fine Line between Information and Defamation in Google's Autocomplete Function. *Cardozo Arts & Ent. LJ*, *34*, 835.

Roozenbeek, J., Schneider, C., Dryhurst, S., Kerr, J., Freeman, A., Recchia, G., van der Bles, A., & van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 7(10), 201199–201199. https://doi.org/10.1098/rsos.201199

Roy, S., & Ayalon, L. (2020). Age and gender stereotypes reflected in Google's "autocomplete" function: The portrayal and possible spread of societal stereotypes. *The Gerontologist*, *60*(6), 1020-1028.

Sallam, M., Dababseh, D., Yaseen, A., Al-Haidar, A., Ababneh, N. A., Bakri, F. G., & Mahafzah, A. (2020). Conspiracy beliefs are associated with lower knowledge and higher anxiety levels regarding COVID-19 among students at the University of Jordan. *International Journal of Environmental Research and Public Health*, *17*(14), 4915.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Scherer, C. W. & Cho, H. (2003). A social network contagion theory of risk perception. *Risk Analysis. 23*(2), 261-267.

Scherr, S. (2020). Measuring Contagion on Social Media. *The International Encyclopedia of Media Psychology*, 1- 6.

Shahsavari, S., Holur, P., Tangherlini, T. R., & Roychowdhury, V. (2020). Conspiracy in the time of corona: Automatic detection of covid-19 conspiracy theories in social media and the news. *arXiv preprint arXiv:2004.13783*.

Shear, M. D., Barnes, J. E., Zimmer, C., & Mueller, B. (2021, May 27). Biden orders intelligence inquiry into origins of virus. *The New York Times*. https://www.nytimes.com/2021/05/26/us/politics/biden-coronavirus-origins.html

Stephens, M. (2020). A geospatial infodemic: Mapping Twitter conspiracy theories of COVID-19. *Dialogues in Human Geography, 10*(2), 276-281.

Sullivan, D. (2018, April 20). How Google autocomplete works in search. https://blog.google/products/search/how- google-autocomplete-works-search/.

Ugander, J., Backstrom, L., Marlow, C., & Kleinberg, J. (2012). Structural diversity in social contagion. *Proceedings of the National Academy of Sciences*, *109(16),* 5962-5966.

Vishwanath, V. (2015). Diffusion of deception in social media: Social contagion effects and its antecedents. *Information Systems Frontiers, 17*(6), 1353–1367. https://doi.org/10.1007/s10796-014-9509-2

Ward, D., Hahn, J., & Feist, K. (2012). Autocomplete as research tool: A study on providing search suggestions. *Information Technology and Libraries*, *31*(4), 6-19.