The First 60 Days: American Public Health Agencies' Social Media Strategies in the Emerging COVID-19 Pandemic

Jeannette Sutton, Scott L. Renshaw, and Carter T. Butts

Abstract: In this paper, we capture, identify, and describe the patterns of longitudinal risk communication from public health communicating agencies on Twitter during the first 60 days of the response to the novel coronavirus disease 2019 (COVID-19) pandemic. We collected 138,546 tweets from 696 targeted accounts from February 1 to March 31, 2020, employing term frequency-inverse document frequency to identify keyword hashtags that were distinctive on each day. Our team conducted inductive content analysis to identify emergent themes that characterize shifts in public health risk communication efforts. As a result, we found 7 distinct periods of communication in the first 60 days of the pandemic, each characterized by a differing emphasis on communicating information, individual and collection action, sustaining motivation, and setting social norms. We found that longitudinal risk communication in response to the COVID-19 pandemic shifted as secondary threats arose, while continuing to promote pro-social activities to reduce impact on vulnerable populations. Identifying patterns of risk communication longitudinally allows public health communicators to observe changes in topics and priorities. Observations from the first 60 days of the COVID-19 pandemic prefigures ongoing messaging needs for this event and for future disease outbreaks.

Keywords: COVID-19, Social media, Risk communication, Epidemic management/response, Public health preparedness/response

INTRODUCTION

Pandemics can be viewed as prolonged risk incidents, marked by high initial uncertainty that decreases as cases accumulate and consensus about its modes of transmission, infection rate, and prognosis grows. As an episode progresses, communications will necessarily change to address developing events in response to emerging information needs[1] and to help individuals manage and reduce their own uncertainty about health and safety.[2,3] In the case of the novel coronavirus disease 2019 (COVID-19) pandemic, this initial uncertainty was been further exacerbated by the novelty of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus, political instability, and the unprecedented reliance on economically and socially disruptive measures as interventions to slow the spread of the disease. Here, we provide an overview of health responsive communicators' messaging over a prominent social media platform (Twitter) during the first 2 months of the pandemic's visible presence in the United States. As the most visible faces of health expertise to thegeneral public, health agencies and public safety organizations play a central role in alerting the public to emerging threats, providing guidance for protective action, motivating compliance with health directives, and combating misinformation. As official accounts representing expert knowledge and policy implementation, they have the potential to garner attention from members of the public seeking credible information about key protective actions that can be undertaken individually and broadly implemented. Understanding the patterns of agency communication during the first months of the pandemic may provide useful lessons for effective communication in the months ahead particularly in the event of a new surge of cases in areas previously considered ''safe'' or a ''second wave'' of infections if social distancing interventions are withdrawn. By identifying and characterizing frequently used keywords and phrases, health responsive communicators can plan for future campaigns to address emerging needs among the public. Beyond the COVID-19 pandemic, this episode provides a unique chance to examine the reaction of the public health messaging system to a rapidly emerging threat during a period of both public and expert uncertainty regarding both mitigation measures and potential societal impact. Characterization of the patterns of agency health communication during this critical period may also offer insight into the development of theory for longitudinal risk communication on social media with applicability to other emerging threat situations.

To capture broad patterns of communication during this period, we focus on the use of hashtags. Hashtags are a widely used device on platforms like Twitter for categorizing information, directing attention to topics or events,[4] and branding posts in an easily recognizable manner.[5] By adding a pound sign to a keyword (eg, #coronavirus), hashtagged words serve as both a symbol and an organizing mechanism for content. Agency hashtag use thus reveals the topics, campaigns, and ideas that the agencies themselves seek to promote when attempting to capture public attention. Prior research on hashtags for health communication has examined public health campaigns, for example, to measure organization–public engagement during a single campaign,[6] or to track sentiment, content, and networks that coalesce around a hashtag.[7] Here, in response to a longitudinal risk communication event, hashtags are an excellent tool to illustrate the shifting foci of public-facing communication by health-focused agencies as they respond to the emerging pandemic.

To reveal changing patterns of emphasis across the more than 10,000 hashtags used by health responsive communicators during the study period, we employed text analysis techniques to identify distinctive hashtags that are characteristic of activity during particular periods—as opposed to being either idiosyncratic or completely ubiquitous during the full 2 months. These allow us to identify emergent communication strategies among organizations that arise in response to the unfolding pandemic and their subsequent decline. Using this approach, we address the question of how the

landscape of COVID-19 information communicated by public agencies online evolved in the early days of the pandemic response.

METHODS

Drawing from existing lists and manual collection, we identified 742 accounts involved in public health messaging within the United States. Our final data set consists of 138,564 English-language tweets from 696 of those accounts tweeting during the study period. Identified accounts include public health agencies at multiple levels of government, as well as emergency management agencies, mayoral accounts, and state governor accounts active in the response. A total of 414 Public health accounts (local n = 298; state n = 50; federal n = 65; international [World Health Organization] n = 1) were identified by drawing on and confirming accounts from publicly available lists8 and subsequent projects on social media risk messaging.9,10 Local mayor and local emergency management agency accounts were identified for the top 100 largest cities in the United States by searching their websites for associated social media accounts. This resulted in accounts from all but 5 states: Delaware, Maine, Vermont, West Virginia, and Wyoming. We also identified social media accounts from the largest city in each of those 5 states. In some cases, mayors (n = 7) and city emergency management agencies (n = 3) did not have active Twitter accounts; we then collected the account associated with the city, county, or affiliated agency linked to their website. In total, we identified accounts for 98 mayors and 102 local emergency management agencies. We also identified and collected the organizational accounts for 50 state emergency management agencies and the individual accounts for every state governor and territory available, including Guam, Puerto Rico, and the US Virgin Islands. In some cases, governors also held secondary accounts (eg, GovRonDeSantis and RonDeSantisFL), resulting in a total of 78 accounts for state-elected officials. While no list of Twitter accounts is comprehensive (indeed, choice of accounts is based upon individual investigator decisions, and, furthermore, there is no directory for account lookup), these targeted accounts represent official organizations with clear roles in communicating about public health and policy decisions at the state, local, and national levels.

We collected messages posted between February 1 and March 31, 2020 on Twitter using the REST API.11 Data collection was performed using the rtweet library12 library for the R statistical computing system.13

We conducted analyses in 3 stages. First, we used automated analyses to identify all of the hashtags in the dataset (n = 10,072). Second, term frequency-inverse document frequency (tf-idf) was used to evaluate the relevance of a hashtag to the collection of tweets. Third, manual coding of hashtags used in COVID-19-related tweets posted by the 696 targeted accounts was conducted to identify the thematic topics that characterized periods of time in the first 60 days.

Textual Analysis

To identify distinctive hashtags obtaining high levels of activity within constrained periods, we employed tf-idf weighting, a well-known technique from the literature on information retrieval.14 For purposes of analysis, all posts were grouped by day of posting, with weights computed for each hashtag on each day: high weights indicate hashtags that are frequently used on the day in question (a property captured by a quantity known as term frequency) while also being selective (a property captured by a quantity known as the inverse document frequency). These 2 elements are defined as follows. Let f_{ij} be the fraction of all hashtags appearing on day j that are copies of hashtag i; this is the term frequency of hashtag i on day j. Similarly, let d_i be the fraction of all days in the sample in which hashtag i appears; this is the document frequency of hashtag i. To obtain the tf-idf weight of hashtag i on day j, we

multiplied the term frequency of the log of the reciprocal of the document frequency, yielding the final weight wij = -fij log di.

The top 25 tf-idf weighted tags on each respective day were selected for subsequent content analysis. Tags with high tf-idf weights are both historically selective (ie, they are not a constant background feature of the data set) and prominent when present (ie, they are not idiosyncratic or infrequently used when they appear).

Content Analysis

For each day of the study period, we identified the hashtags related to COVID-19 from within the 25 most prominent (tf-idf selected) tags included in the tweets of our 696 targeted accounts on that day. In making this determination, we considered hashtag words associated with pandemic related keywords and events, such as the name of the disease (eg, nCoV, coronavirus, COVID) protective-action recommendations for the virus (eg, handwashing) or key phrases used to explain the collective actions necessary to reduce the spread of the virus (eg, flattenthecurve, socialdistancing, stayhome). Hashtags associated with secondary keywords were also included such as key events (eg, the Princess cruise that was not allowed to dock due to coronavirus patients or the arrival of Navy medical ships in California and New York) or planned public events (eg, when Major League Baseball encouraged people to stream classic baseball games during lockdown using the hashtag #openingdayathome). Any remaining hashtags whose relevance was not immediately apparent was evaluated by the first author who manually searched Twitter for posts employing the hashtag. Tags found to be used for pandemic-related information or discussion were marked as related to COVID-19.

After determining which hashtags were topically related to the pandemic, either directly or as a secondary threat/impact, the first author identified conceptual categories that emerged from the data in order to describe how groups of hashtags functioned at different points in time (such as early onset) and to motivate action (such as hashtags that reference behavior change).15 The first author shared these conceptual categories and their related hashtags with the other 2 authors, and they collaboratively suggested alternative descriptors for the time periods and the functions that groups of hashtags appeared to fill. Final decisions for the conceptual categories and periods of time were also determined collectively. Hashtags were grouped according to the time period in which they emerged and persisted, providing a chronology of changing risk communication emphases in the early days of the pandemic.

RESULTS

We identified 7 distinct periods (Figure 1) of communication that emerged from the data from February 1 to March 31, 2020. These periods are described as ''epochs'' (see Table 1 for a list of related hashtags), each being characterized by a distinct pattern of communication emphasis, as revealed by distinctive hashtag use.

Epoch 1: February 1 to 19 – lead up
During this period, the coronavirus is becoming recognized as an emerging international threat, with some awareness of its coming impact to the United States. Early pandemic-related messages were tagged with the hashtag #nCoV19 (novel coronavirus) to describe the emerging threat. On February 12, the World Health Organization named the novel coronavirus disease ''COVID-19,'' which was quickly adopted as a signal word for the disease.

Epoch 2: February 20 to March 12 – acute onset

By late February, restrictions had been placed on international travel by the US Department of Homeland Security and the US Centers for Disease Control and Prevention (#CDCTravelNotice), and widespread travel recommendations had been made by the World Health Organization.16 Shortly thereafter, the virus was discovered among passengers onboard the Grand Princess cruise ship, which led to a ''coronavirus limbo'' and raised awareness of the coming impacts (#cruise #GrandPrincess).17 In quick succession, state and local organizations adopt geographically specific hashtags (#COVID19MD, #COVID19Colorado) in order to communicate local disease transmission and preparedness efforts.

Epoch 3: March 13 to 14 – individually focused action
A nationwide state of emergency was declared on Friday, March 13, 2020, initiating the use of hashtags to describe individual actions that could be taken to protect oneself from the virus (eg, #safehands, #selfquarantine, #protectyourself). Initially this was focused on minimally disruptive protective actions that individuals could take independently as part of their daily habits, such as hand-washing, and self-identification of symptoms that should lead individuals to self-quarantine if they were ill or suspected they had been exposed to someone infected with the virus.

Epoch 4: March 15 to 18 – community focused action
Measures to limit or inhibit transmission such as #socialdistancing were promoted as a type of hashtag campaign18 of simple slogans. Public agencies adopted the use of charts and graphs to visually show how our collective efforts could #flattenthecurve and #stopthespread of infection in order to reduce the coming strain on the healthcare system.

Epoch 5: March 13 to 21 – sustaining motivation
At the same time that action-oriented hashtags emerged, we also saw the rise of ''moral appeals,''19 calling upon the public to take action not just to protect themselves, but to also protect vulnerable populations (#protectthevulnerable). These motivational hashtags also called for unity of effort and action, such as #teameffort, noting that the end result will be victory (#wewillprevail).

Epoch 6: March 19 to 23 – setting social norms
Within 1 week of the national state of emergency, individual states begin to issue stay at home and shelter in place orders.20 Earlier recommendations were replaced by significantly disruptive directives that limited contact among people. Generic, collective motivational appeals soon gave way to state-based hashtags emphasizing local group membership. States adopt localized hashtags such as #togetherKY, #stayhomeforMN, #fightthespreadVA, and #inthistogetherOhio, encouraging populations to adopt new health and safety practices.

Epoch 7: March 24 to 31 – secondary impacts
As stay at home orders spread across the country, secondary impacts resulting from the spread of the virus and shutdown orders began to grow. New online campaigns emerged that addressed economic and social issues as they developed. For example, people were encouraged to support local businesses by participating in the #greatamericantakeout or to come together while apart by watching #openingdayathome, which consisted of re-releases of great games in Major League Baseball.

While several early hashtags persisted from the time of their emergence to the end of the study period (eg, #COVID-19; #stopthespread; #socialdistancing) new topics began to arise relative to unfolding economic and mental health stresses. This prefigures broader message fatigue from isolation due to social distancing and conversations about ''opening up'' the economy.

DISCUSSION

Prior research has characterized effective pandemic risk communication as messaging that instructs, informs, and motivates appropriate self-protective behavior.[1] If health-responsive organizations are actively monitoring the changing environment and informational needs of their constituents, it is apparent by the use of these changing hashtags that the organizations are responding to a perceived public demand for concrete, accurate, and consistent information about protective actions.[21] Research on pandemic risk communication has identified the need to deliver informative and instructive information, while also placing an emphasis on communicating risk to more vulnerable populations. While agencies clearly advocated ethical behavior through the use of hashtags emphasizing moral appeals, such as asking the public to protect those who are most vulnerable, there seemed to be an absence of hashtags directed specifically to high-risk populations.[22,23] For example, messages about the risks of COVID-19 could have been targeted to specific groups of people who have already been identified as key constituents of public health organizations, such as pregnant or nursing women, people with cancer, or individuals with chronic health conditions. Although there is limited evidence that elderly and institutionalized populations actively engage in the use of social media platforms, social media messaging directed to specific groups, such as racial and ethnic minorities, has a high potential to be seen.[24]

Initially, location-specific hashtags, such as #COVID19Seattle or #COVID19MD served to organize content that was specific to an individual's location; thus, hashtag use served as a strategy to direct persons to local updates about transmission, testing, and policy. Such strategies can be helpful for people looking for information specific to their jurisdiction. The use of such hashtags may also encourage others to contribute to local community reporting.[25] As states began to issue local shelter in place and stay at home orders, state-based hashtags also emphasized group membership such as #sayhomefortexas or #togetherKY. Notably, group membership is an important factor in developing social norms[26] and we often look to our neighbors to decide what actions we will take.[27] Under lockdown conditions, or when face-to-face interaction is limited, online sources are likely to take on a larger role in connecting individuals with others who are otherwise isolated—see, for example, the use of #hashtags to connect individuals during the Occupy movements in 2010-2011.[28] Motivational hashtags signifying unity, strength, and togetherness have the potential to make us digital neighbors in very difficult times.

PUBLIC HEALTH IMPLICATIONS

Research on pandemics has shown that effective health risk communication focuses on instructive, informative messages that motivate appropriate self-protective behavior and foster resiliency. We are learning that longitudinal risk communication cannot be focused entirely on emergency risk communication and instructive messages. It also, by necessity, must include sustained action across the broader population. Preparing motivational campaigns that reward people for their behavior may be necessary to instill ongoing action in the future. Considering how to recapture attention when it wanes, even as the pandemic continues, will also become important. Public health-oriented hashtag campaigns may help engage individuals to help them to feel part of a larger collective body and to participate locally by contributing information about their local context.

Another central feature of the COVID-19 pandemic is the need to respond to rapidly changing circumstances, due to changes in the state of public health knowledge (eg, on the efficacy of masks for personal protection) or the evolving political and economic situation (eg, distancing regulations and resistance thereto). Pandemics are, ultimately, disasters, and the critical role of improvisation that is

central to effective disaster response29 is also inescapable here. It is unlikely that the COVID-19 pandemic will fully stabilize in the near future, therefore, public health agencies need to establish processes to continuously reassess and reevaluate messaging practices in light of changing events, and, where possible, to anticipate messaging that could be used if particular future conditions were to occur. It is becoming exceedingly apparent that risk communicators will need to prepare now for the future deployment of a vaccine, drawing from health communication campaigns and online social media interventions to aid in community-wide uptake.

We also observed that ''boosting'' a specific message to prominence often involved coordinated efforts by many parties within a given state or region to employ specific tags at specific times. This coordinated message discipline requires effort but can have an outsized effect on the ability of campaigns to spotlight particular content in the always tumultuous social media environment. While the COVID-19 pandemic is, in some respects, a unique and distinctive case, it also exemplifies the greatest challenges public communication may face with a rapidly emerging and highly uncertain threat. Future pandemics or sudden-onset health threats—such as the continent-wide radiation exposures experienced in Europe and portions of Asia following the 1986 Chernobyl disaster—will inevitably pose related challenges to public-facing agencies. Planning for rapid coordination, use of common language, and targeted boosting of critical communications against a threat that emerges quickly and is poorly understood will be important for successful response to future events.

ACKNOWLEDGMENTS

REFERENCES

1. Vaughan E, Tinker T. Effective health risk communication about pandemic influenza for vulnerable populations. Am J Public Health. 2009;99(suppl 2):S324-S332.
2. Lachlan KA, Spence PR, Seeger M. Terrorist attacks and uncertainty reduction: media use after September 11. Behav Sci Terrorism Polit Aggression. 2009;1(2):101-110.
3. Liu BF, Bartz L, Duke N. Communicating crisis uncertainty: a review of the knowledge gaps. Public Relations Rev. 2016; 42(3):479-487.
4. Bruns A, Burgess JE. The use of Twitter hashtags in the formation of ad hoc publics. Paper presented at: 6th European Consortium for Political Research General Conference; August 25-27, 2011; Reykjavik, Iceland.
5. Sutton J, Gibson CB, Phillips NE, et al. A cross-hazard analysis of terse message retransmission on Twitter. Proc Natl Acad Sci U S A. 2015;112(48):14793-14798.
6. Abroms LC. Public health in the era of social media. Am J Public Health. 2019;109(suppl 2):S130-S131.
7. Dumbrell D, Steele R. #worldhealthday 2014: the anatomy of a global public health Twitter campaign. Paper presented at: 48th Hawaii International Conference on System Sci-

ences; January 5-8, 2015; Kauai, HI.

8. Harris JK, Mueller NL, Snider D. Social media adoption in local health departments nationwide. Am J Public Health. 2013;103(9):1700-1707.

9. Vos SC, Buckner MM. Social media messages in an emerging health crisis: tweeting bird flu. J Health Commun. 2016;21(3):301-308.

10. Vos SC, Sutton J, Yu Y, et al. Retweeting risk communication: the role of threat and efficacy. Risk Anal. 2018;38(12): 2580-2598.

11. Twitter. Docs: guides and reference materials to help you get started, integrate, optimize, and troubleshoot your use of the Twitter Development Platform. Accessed September 29, 2020. https://developer.twitter.com/en/docs

12. Kearney MW. rtweet: collecting and analyzing Twitter data. J Open Source Software. 2019;4(42):1829.

13. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. Accessed September 29, 2020. https:// www.R-project.org

14. Manning CD, Schu¨tze H. Foundations of Statistical Natural Language Processing. Cambridge, MA: MIT Press; 1999.

15. Gibbs GR. Analyzing Qualitative Data. Thousand Oaks, CA: Sage Publications, Inc; 2007.

16. World Health Organization (WHO). Updated WHO recommendations for international traffic in relation to COVID-19 outbreak. Published February 29, 2020. Accessed September 29, 2020. https://www.who.int/news-room/articles-detail/updated-who-recommendations-for-international-traffic-in-relation-to-covid-19-outbreak

17. Chabria A, Fry H, Cosgrove J. What we know now that Grand Princess has docked and coronavirus quarantines have begun. Los Angeles Times. March 11, 2020. Accessed September 29, 2020. https://www.latimes.com/california/story/2020-03-09/what-we-know-about-grand-princess-docking-coronavirus-quarantine

18. Maibach EW, Parrott RL, eds. Designing Health Messages: Approaches from Communication Theory and Public Health Practice. Thousand Oaks, CA: Sage Publications, Inc.; 1995.

19. Guttman N, Ressler WH. On being responsible: ethical issues in appeals to personal responsibility in health campaigns. J Health Commun. 2001;6(2):117-136.

20. Bizjak T, Yoon-Hendricks A, Breton M. Sacramento County issues formal order for residents to stay at home at coronavirus spreads. Sacramento Bee. March 19, 2020. Accessed September 29, 2020. https://www.sacbee.com/news/coronavirus/article241332696.html

21. Wray RJ, Becker SM, Henderson N, et al. Communicating with the public about emerging health threats: lessons from

the pre-event message development project. Am J Public Health. 2008;98(12):2214-2222.

22. Crouse Quinn S. Crisis and emergency risk communication in a pandemic: a model for building capacity and resilience of minority communities. Health Promot Pract. 2008;9(suppl 4):18S-25S.

23. Hutchins SS, Truman BI, Merlin TL, Redd SC. Protecting vulnerable populations from pandemic influenza in the United States: a strategic imperative. Am J Public Health. 2009;99(suppl 2):S243-S248.

24. Wojcik S, Hughes A. Sizing up Twitter users. Pew Research Center website. April 24, 2019. Accessed September 29, 2020. https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/

25. Starbird K, Palen L. (How) will the revolution be retweeted? Information diffusion and the 2011 Egyptian uprising. In: CSCW '12: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work. New York: Association for Computing Machinery; 2012:7-16.

26. Feldman DC. The development and enforcement of group norms. Acad Manage Rev. 1984;9(1):47-53.

27. Levine M, Prosser A, Evans D, Reicher S. Identity and emergency intervention: how social group membership and inclusiveness of group boundaries shape helping behavior. Personality Soc Psychol Bull. 2005;31(4):443-453.

28. Skinner J. Social media and revolution: the Arab Spring and the Occupy Movement as seen through three information studies paradigms. Working Pap Inf Syst. 2011; 11(169):2-26.

29. Butts CT, Petrescu-Prahova M, Remy Cross B. Responder communication networks in the World Trade Center disaster: implications for modeling of communication within emergency settings. J Math Sociol. 2007;31(2):121-147.
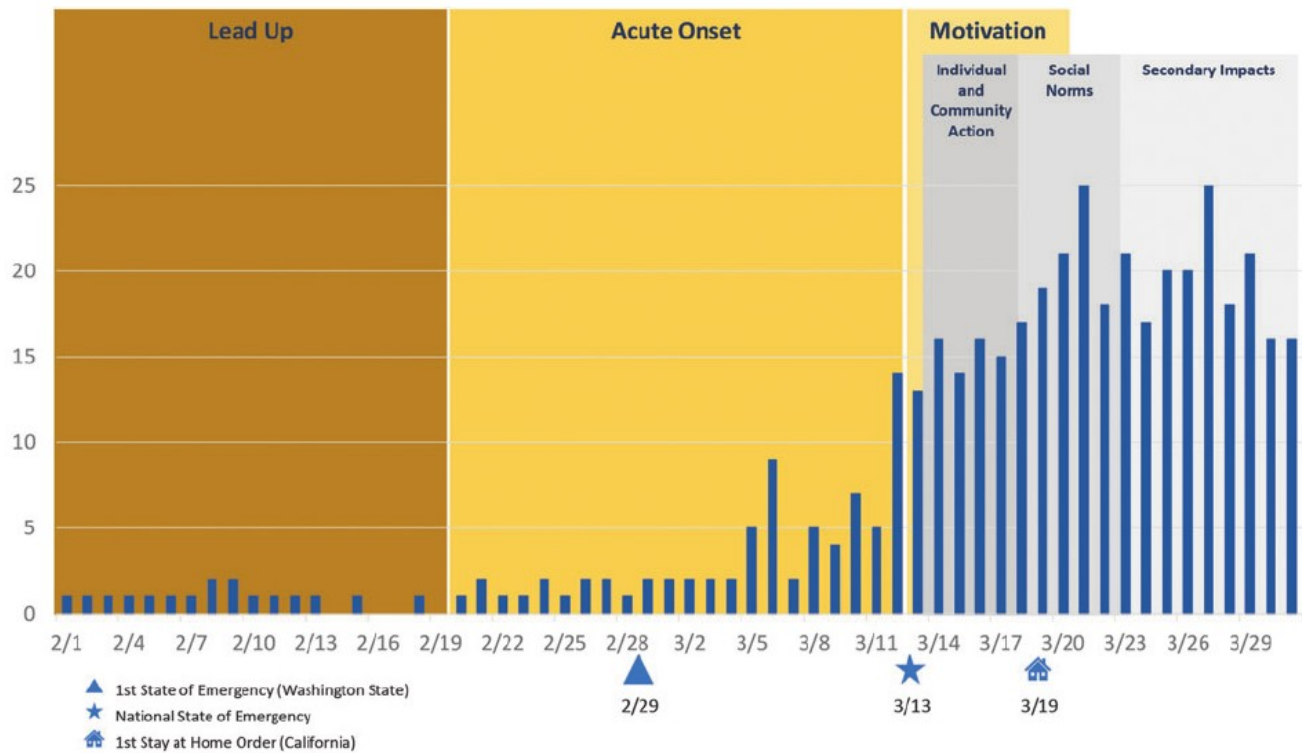
Figure 1: COVID-19 hashtag frequencies, February 1 to March 31, 2020. The figure shows the frequency of the top 25 hashtags used by the targeted accounts during this period. The blue bars along the X axis represent the number of COVID-19-related hashtags in the top 25 hashtags per day. In the lead up to the declaration of the National State of Emergency, few of the top 25 hashtags from targeted accounts were related to COVID-19. On March 21 and 28, all of the top 25 hashtags used by target accounts were related to COVID-19.

Epoch	#Hashtags that Emerge in Each Epoch
Epoch 1: February 1 to 19	
February 1 to 11	#2019nCoV
February 12	#COVID19[a]
Epoch 2: February 20 to March 12	
February 20 to 29	#CDCTravelNotice
March 3 to 8	#cruise #GrandPrincess
March 5 to 12	#COVID19OhioReady[a] #sgchdCOVID10 #COVID19MD #COVID19NC #Covidwa #COVID19Seattle #COVID19Colorado[a] #COVID19Va
Epoch 3: March 13 to 14	#safehands #selfquarantine #protectyourself
Epoch 4: March 15 to 18	#flattenthecurve[a] #staytheFhome #socialdistancing[a] #sharefactsnotgerms #Stopthespread #fightthespread #protectthevulnerable #stayhomesavelives
Epoch 5: March 13 to 21	#togetherusall #protectyourcommunity #healthytogether #allhandsondeck #doingmypart #teameffort #wewillprevail #inthistogether #saferathome[a]
Epoch 6: March 19 to 23	#stayhomefornevada #togetherKY #stayhomeMN #estevirusloparamosunidos #stayhometexas #fightthespreadVA #marylandunites #alltogetherNM #inthistogetherOhio[a] #stocktonstrong #cheyennestrong #newyorkstatestronger #stayhomeWV #stopCVMKE #stayhomeOhio[a]
Epoch 7: March 24 to 30	#greatamericantakeout #clinicalresearch #beactive #openingdayathome #usnsmercy #thankyouphilly #chalkoutcovidok #usnscomfort

[a] #hashtags that persist from the date of emergence to end of study period.

Table 1: Representative COVID-19 #hashtags for the Study Period February 1 to March 31, 2020