2022

Reddit as a Source of COVID-19 Information: A Content Analysis of r/coronavirus during the Early Pandemic

Brent Hale
University of Southern Mississippi, brent.hale@usm.edu

Mark Alberta
Indiana University - Bloomington, malberta@iu.edu

Seung Woo Chae
Indiana University - Bloomington, seuchae@iu.edu

Recommended Citation

Hale, B. J., Alberta, M., & Seung, C. W. (2022). Reddit as a source of COVID-19 information: A content analysis of r/coronavirus during the early pandemic. Journal of Communication Technology, 5(1), 26-57. doi.org/10.51548/joctec-2022-002

JoCTEC Journal of Communication Technology

Published by the Communication Technology Division
of the Association for Education in Journalism and Mass Communication

Reddit as a Source of COVID-19 Information: A Content Analysis of r/coronavirus During the Early Pandemic

Brent J. Hale[a], Mark Alberta[b], and Seung Woo Chae[b]

[a]University of Southern Mississippi, Hattiesburg, Mississippi, USA; [b]Indiana University, Bloomington, Indiana, USA
Correspondence: brent.hale@usm.edu

Abstract

Emerging research has begun examining the utility of social media platforms for information dissemination during the COVID-19 pandemic. Following this developing thread, this work examines discourse within r/coronavirus, a Reddit forum (i.e., subreddit) developed to curate COVID-19 information that burgeoned during the early months of the pandemic. Through a content analysis of 226 posts and 2260 corresponding comments generated between February and May, 2020, this study investigated early-pandemic communication patterns in this platform, including what information was deemed important and how users framed causes and solutions. Overall, findings indicate that users of r/coronavirus prioritized information about COVID-19 spread, public health information (e.g., mask mandates), political and economic implications of COVID-19, and the experiences of medical workers, while also shaming people who failed to follow public health guidelines. Discourse was collectivistically oriented and negatively valenced, and engagement with the subreddit decreased over time, suggesting COVID-19 fatigue among r/coronavirus users. Taken together, this study provides a window into early pandemic discourse, furthering our understanding of COVID-19 communication in early 2020 and revealing informational needs that could be targeted during future health emergencies. Furthermore, progressive fatigue in r/coronavirus (i.e., decreasing curational motivation) showcases a difficulty facing public health communicators during a pandemic's emergence.

Keywords: COVID-19, social media, health communication, pandemic communication, Reddit

Introduction

The disruption of the COVID-19 pandemic to global health and economic wellbeing, caused by the novel coronavirus strain SARS-CoV-2, should come as no surprise to contemporary readers. The first case of COVID-19 was documented in Gansu Province, China in December 2019 (Fan et al., 2020), and the coronavirus infection rapidly spread to affect every continent in 2020 (including Antarctica; Booker, 2020). Through the combined global efforts of medical research scientists and immunologists, several vaccines were developed, tested, and approved for emergency use by the end of 2020 – an unprecedented timeline for vaccine development (FDA, 2021). However, by August 2021, over 202 million people had tested positive for the infection, resulting in nearly 4.3 million deaths globally (Bloomberg, 2021).

Although many sources of COVID-19 information are currently available to individuals interested in tracking developments (from government agencies, news sources, and research establishments), several factors inhibited information curation (and accuracy) in the early days of the pandemic. First, the disease spread quickly, initially infecting individuals in China, and then rapidly expanding to Italy, Germany, and the United States (Fan et al., 2020). Additionally, the nature of the disease made tracking spread difficult, as symptoms of COVID-19 take up to two weeks to manifest in those infected (which was initially unknown; CDC, 2020a), and some infected individuals may never exhibit any symptoms (CDC, 2020b). Perhaps due to the rapid spread and potential for asymptomatic transmission, many governments were slow to respond to the pandemic (Abutaleb et al., 2020). Consequently, agencies like the Centers for Disease Control and Prevention (CDC) in the United States were slow to curate information and provide guidance to American citizens, and in some cases provided directions that were later retracted (Elfrink et al., 2020; Dwyer & Aubrey, 2020). Furthermore, recent years have been characterized by a prevalence of misinformation, especially through social media platforms (Allcott et al., 2019; Tasnim, Hossain, & Mazumder, 2020; Valenzuela et al., 2019), resulting in what some have called an “infodemic” in response to COVID-19 (Brennen, Simon, Howard, & Nielsen, 2020; Brennen, Simon, & Nielsen, 2020).

Misinformation (and disinformation) issues notwithstanding, *Reddit* (reddit.com) boasted particular credibility during the early pandemic, partly due to the platform’s popularity, affordances, and usage. First, *Reddit* has emerged as a prominent social media platform, serving approximately 430 million monthly global users, though this userbase is notably smaller than the that of *Facebook* (2.74 billion), *YouTube* (2.29 billion), *WhatsApp* (2 billion), and *Instagram* (1.22

billion) (Statista, 2021). The relatively smaller size of *Reddit* may have protected the site from larger disinformation campaigns interested in maximizing views (which frequently targeted large platforms like *Facebook*, *WhatsApp*, and *Twitter*; Bruno Kessler Foundation; 2020). In addition to the platform's size, the bidirectional voting system (using upvotes and downvotes) provided *Reddit* an advantage over social media systems that employ unidirectional voting (e.g., likes or favorites), as the ability to downvote information afforded greater information curation – misinformation could be fact-checked and downvoted out of view, while accurate information could be upvoted to appear higher in post/comment rankings. Moreover, *Reddit*'s topic-focused organization allowed information centralization within COVID-19 subreddits, most notably r/coronavirus, which gained 2.4 million members by January 2021 (r/coronavirus, 2021) and has recently begun receiving scholarly interest (Gaur et al., 2021; Gozzi et al., 2020; Lai et al., 2020; Wu et al., 2021).

The r/coronavirus subreddit emerged as a preeminent source of COVID-19 information during the early pandemic, partly due to AMAs ("ask me anything") by individuals such as Bill Gates (Reddit, 2020a) and reporting from medical experts in South Korea (Reddit, 2020b), the United Kingdom (Reddit, 2020c; 2020d), Switzerland (Reddit, 2020e), and the United States (Reddit, 2020f; 2020g). Additionally, the subreddit has been moderated by a team of infectious disease researchers, virologists, computer scientists, doctors, and nurses (Solon & Glaser, 2020). Unsurprisingly, the subreddit grew rapidly, becoming an important source of current information about COVID-19. However, despite the popularity of r/coronavirus and the emergence of this subreddit amidst an early-pandemic information vacuum, little is known about what information was disseminated, and how users of this platform described the pandemic (for exception see Gozzi et al., 2020). This is an unfortunate limitation, as the emergence of r/coronavirus during the early pandemic combined with the unresponsiveness of governmental organizations (e.g., the CDC) provides a unique window to identify what information is prioritized by the public during a public health emergency.

Therefore, through a content analysis of posts and comments submitted to r/coronavirus, this study seeks to identify patterns of communication in this platform, including what information was deemed important by r/coronavirus users and how they framed causes and solutions. Accordingly, this study provides a window into early pandemic discourse, furthering our understanding of COVID-19 communication in early 2020 and revealing informational needs that could be targeted during future health emergencies. Moreover,

considered alongside other emerging research examining r/coronavirus (e.g., Gozzi et al., 2020), this work provides evidence for the utility of social media platforms like *Reddit* for public health communication.

Literature Review

Information Curation and Risk Assessment

Proponents of the risk information seeking and processing model (RISP) contend that individuals facing a perceived hazard (e.g., a pandemic's emergence) seek information to reach desired informational goals (i.e., achieve information sufficiency), promoting informed behavioral decisions and mitigating anxiety (Dunwoody & Griffin, 2015; Griffin, Dunwoody, & Neuwirth, 1999). Previous research suggests that individuals facing a disease outbreak seek various types of information, including details about prevention and treatment, personal health risks, and vaccines (Henrich & Holmes, 2011; Springer et al., 2020; Wong & Sam, 2010). When seeking information, people often turn to media platforms (e.g., television, radio, newspapers, and social media), which fulfil a central role in contemporary information dissemination, increasing perceived knowledge and understanding of personal susceptibility and risk, and promoting appropriate behavioral responses (Henrich & Holmes, 2011; Strekalova, 2017; Wong & Sam, 2010). According to RISP, individuals select these media sources strategically, making a judgment about the channel's "quality" (considering channel bias and perceived special interests) (Dunwoody & Griffin, 2015), which is perhaps of particular contemporary relevance considering issues with misinformation (and disinformation) and the COVID-19 "infodemic" (Allcott et al., 2019; Brennen et al., 2020; Tasnim et al., 2020).

As noted previously, r/coronavirus emerged as a potentially "high-quality" channel and a popular source of information during the COVID-19 pandemic (Gozzi et al., 2020), especially considering the lackluster response by government agencies like the CDC during the pandemic's infancy. This perceived credibility stemmed partially from public health experts and medical professionals appearing on the subreddit to share information with the public (frequently through AMAs), promoting dialogue between stakeholders interested in encouraging public health behaviors (e.g., Bill Gates; Reddit, 2020a), experts actively combatting the disease (e.g., practitioners; Reddit, 2020b), and *Reddit* users interested in COVID-19 developments and preventative health behaviors (Lai et al., 2020). In addition, the moderation strategy employed by the subreddit, which included moderation by infectious disease researchers, virologists, and doctors (Solon & Glaser, 2020), bolstered the perceived credibility

and utility of r/coronavirus.

Notably, as r/coronavirus emerged as a platform for COVID-19 information, digital media use was increasing substantially, partially due to government-mandated lockdowns (Kemp, 2020). Indeed, as Russell and Powers (2020) note, in the era of COVID-19 "we're *living virtually*" (p. 1), resulting in especially prevalent traffic increases for social media platforms (Nguyen et al., 2020). Given this backdrop, with the subreddit receiving attention from influential sources (e.g., Bill Gates), the public's informational needs during the early pandemic, and the increase in digital media use, it is perhaps unsurprising that r/coronavirus gained 2.4 million members by January 2021 (r/coronavirus, 2021). However, despite the platform's emergent popularity and supposed informational utility, little research to date has examined what information was disseminated through r/coronavirus, or how users described the pandemic, although this content likely influenced how users calculated perceived risk (i.e., how dangerous is COVID-19?) and rendered behavioral decisions (e.g., should I wear a mask?).

Of particular interest for this study is the communication that emerged during the early pandemic (e.g., February-May 2020), as the information disseminated during this period offset the ambiguous (and sometimes counterintuitive) messages distributed by government agencies like the CDC in the United States. Moreover, r/coronavirus activity during this window provides unique insight into the public's information needs during a pandemic's emergence. Therefore, the first goal of this work is to categorize posts submitted to r/coronavirus thematically, including (but not limited to) themes related to disease spread, politics, vaccine development, and the experiences of medical professionals, to determine what information was important to users of the subreddit. Previous literature suggests that negatively-valenced information could be of particular priority, given a consistent prevalence of negative content in response to crisis events (e.g., Ferrara & Yang, 2015; King & Wang, 2021). Relatedly, this study also seeks to determine differences in user engagement (e.g., upvotes and comments) between thematic categories, as user engagement could function as an additional indicator of informational priority. Because posts and comments that receive high scores (i.e., many upvotes) and numerous response comments (from other r/coronavirus users) are prioritized by *Reddit*'s algorithm, engagement is a primary vehicle for *Reddit* users to influence content ranking. If particular thematic categories received consistently higher engagement, this would suggest that the r/coronavirus community prioritized this content, indicating perceived informational relevance.

RQ1a: What thematic categories appeared most frequently during the early pandemic?

RQ1b: What thematic categories received the most engagement (i.e., upvotes and comments)?

In addition to examining posts submitted to r/coronavirus, this study also seeks to identify the informational utility of comments generated in response to these posts. Various stakeholders have made efforts to combat the COVID-19 "infodemic" (Brennen et al., 2020), including governments (which have criminalized malicious coronavirus falsehoods in some cases), governmental agencies, and social media companies (Radu, 2020). *Reddit* took several steps to combat COVID-19 misinformation, including investigating coordinated misinformation campaigns, conducting experimental research, curating expert AMAs, and empowering moderators and users to report misinformation (Reddit, 2020h). Considering r/coronavirus specifically, the community description states that "this subreddit is for high-quality posts and discussion", and one of the subreddit rules – "keep information quality high" – asserts that while "there are many places online to discuss conspiracies and speculate, we ask you not to do so here" (r/coronavirus, 2021). Thus, several steps were taken both across *Reddit* and within r/coronavirus to combat misinformation, while this subreddit implemented additional moderation procedures designed to promote information accuracy.

However, given the large number of comments generated within r/coronavirus, it is unclear how many comments were indeed informational (and not conspiratorial or speculative). Therefore, the second goal of this study is to classify r/coronavirus comments, identifying the proportion that 1) provided information about COVID-19 and related issues (contributing to the satisfaction of users' informational needs), 2) included speculation about the disease response, including motivations behind decisions and potential outcomes (disregarding the subreddit's call to eliminate speculation), or 3) reacted to post content without providing any additional information or speculation. This evaluation will allow us to further diagnose r/coronavirus discourse, including the potential informational utility of comments, or conversely their digression from the high-quality information standards promoted by the subreddit's moderation team (ultimately inhibiting satisfaction of users' informational needs and risk assessments). Put differently, while the first objective of this study is to identify what types of information were prioritized by r/coronavirus during the early COVID-19 pandemic, our second aim is to identify the informational utility of this

subreddit's discourse for public health.

RQ2: To what extent were comments information, speculative, or reactive?

Individualism and Collectivism

In addition to satisfying informational needs, r/coronavirus discourse could also influence users' conceptual framework used to understand public health responsibility (thereby affecting behavioral decisions), including the role of individuals and/or collectives in facilitating beneficial public health outcomes. Emerging research suggests that individualistic/collectivistic orientation has influenced public health outcomes (including mental health) throughout the COVID-19 pandemic (Germani et al., 2020; Maaravi et al., 2021), and thus understanding the attributional frame used by r/coronavirus users could have implications beyond the prevalence of informational themes. Furthermore, prior scholarship has predominantly focused on framing in messages disseminated by professional communicators (e.g., public health organizations, politicians, and journalists) (Boukes et al., 2015; Kim et al., 2020). However, analysis of collectivistic/individualistic framing in r/coronavirus discourse would produce a unique view into how social media users (and perhaps the general public) understood the COVID-19 pandemic and the public's responsibility in combatting the disease.

Originating from cultural theories that place nations on an individualistic/collectivistic continuum (Hofstede, 2001; Rucinski, 1992; Triandis, 1995), framing scholarship commonly examines individualism and collectivism dichotomously (e.g., Boukes et al., 2015). According to Kim et al. (2020), communicators using the individualistic frame "lionize the agency of individual members of society and tend to explain (or blame) individuals as the source of problems", while those using the collectivistic frame "often frame social issues as the outcome of structural deficiencies" (p. 2). In other words, the individualistic frame is utilized when the causes of (or solutions to) a problem are understood as stemming from a particular person (in the COVID-19 pandemic, President Donald Trump could be an example), while the collectivistic frame instead suggests that systemic issues are responsible for problems and solutions (e.g., government deficiencies, coordinated public health campaigns).

Given the COVID-19 pandemic's evolution into a global issue, and the collectivistic responsibility implicit in public health messages (e.g., universal mask use), collectivistic attribution is perhaps likely to emerge in most r/coronavirus content. However, the centricity of the United States for the *Reddit* platform, and the heavy emphasis

on individualism in this country, could instead elicit individualistic framing. Given that collectivistic/individualistic framing has noteworthy implications for resulting beliefs, attitudes, and behaviors (Barrile, 1986; Iyengar, 1991; Maaravi et al., 2021; Viswanath et al., 2015), including affecting subjective norms that influence behavioral choices (Dunwoody & Griffin, 2015), documentation of these frames within r/coronavirus should prove elucidating. If r/coronavirus discourse leans collectivistically, this would indicate an understanding of shared responsibility (e.g., universal mask usage and social distancing), which is consistent with public health messaging during the COVID-19 pandemic. In contrast, a predominantly individualistic orientation among r/coronavirus users would suggest disagreement with public health messages, including the collective responsibility of the public to mitigate risk. Either way, the frame used by r/coronavirus users signifies attitudinal and behavioral expectations that may satisfy informational needs (e.g., "am I responsible for doing something?"). Therefore, the third goal of this work is to examine the attributional frame used by r/coronavirus members, which could have influenced public perception of the COVID-19 pandemic and public health responses (especially given the aforementioned importance of this subreddit for information curation).

RQ3: To what extent did users of r/coronavirus use individualistic or collectivistic framing?

COVID-19 Fatigue

The World Health Organization defines pandemic fatigue as "demotivation to follow recommended protective behaviours, emerging gradually over time and affected by a number of emotions, experiences and perceptions" (World Health Organization, 2020, p. 7), and considers pandemic fatigue as "an expected and natural reaction to the prolonged nature of this [COVID-19] crisis and the associated inconvenience and hardship" (p. 6). They contend that protective behaviors (e.g., mask wearing, social distancing, self-quarantining) are successfully adopted when individuals are capable of following guidelines, provided opportunities to comply, and motivated to follow health recommendations. However, individuals become decreasingly motivated to comply over a prolonged period of time, resulting in inattention to new information and resistance to protective behaviors. Moreover, concerns have been extended that incessant repetition of COVID-19 health messages could increase the risk of mental fatigue and desensitization (Koh et al., 2020), perhaps because of continued information provision after reaching desired informational goals (i.e., information sufficiency) (Dunwoody & Griffin, 2015).

While demotivation to follow protective behaviors and mental fatigue would be difficult to measure through *Reddit* content, fatigue could manifest differently in r/coronavirus posts and comments. Specifically, progressively decreasing user interaction (i.e., upvotes and comments) could suggest fatigue among users of r/coronavirus, as engaging with COVID-19 content progressively less over the course of the pandemic would indicate a decreasing willingness to curate information through this platform. Alternatively, this could indicate satisfaction of informational needs (i.e., users leaving the platform after achieving information sufficiency), though this seems implausible given the rapidly changing nature of public health guidelines, political argumentation, and public understanding of COVID-19 during the early pandemic (for a timeline see AJMC, 2021). Considering that r/coronavirus users were particularly committed to curating COVID-19 information during the early pandemic, evidence of fatigue during this window could indicate faltering motivation or perceived information satisfaction, which would have important implications for public health practitioners during future emergencies. In fact, early evidence suggests r/coronavirus users began to exhibit decreased interactions with the community in early 2020 (Gozzi et al., 2020), following similar information gathering patterns during previous health emergencies (often attributed to information saturation) (Pruss et al., 2019; Tausczik et al., 2012). This study seeks to further examine this possibility.

RQ4: Is there evidence of fatigue in r/coronavirus activity during the early pandemic?

Methods

A content analysis was conducted to identify patterns of COVID-19 communication in the r/coronavirus subreddit during the early pandemic, including what information was deemed important and how users framed causes and solutions (i.e., individualistically or collectivistically). In total, 226 *Reddit* posts and 2260 comments were analyzed for thematic categorization (e.g., information about treatments, the economy, politics, or disease spread), user engagement (i.e., scores, number of comments), comment content (e.g., information, speculation, and reaction), and attributional frame (i.e., individualism or collectivism), providing insight into early informational discourse surrounding COVID-19.

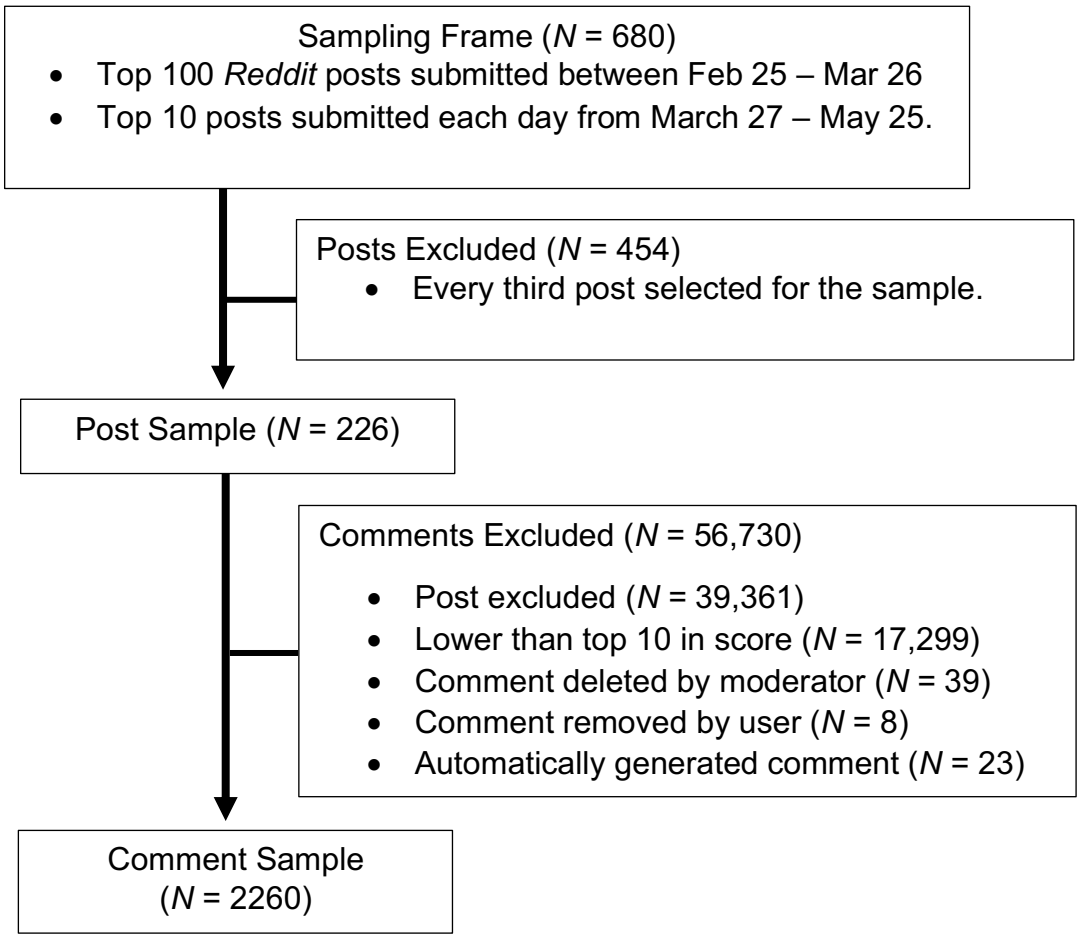
Sample

Using the *Reddit* API, a sampling frame was created by collecting the top ten posts (organized by score) submitted daily to r/coronavirus from March 27 – May 25, 2020, with a single retroactive collection on March 26 to capture the top 100 posts from the

preceding month (i.e., February 25 – March 26) (*N* = 680), a similar window to other early pandemic r/coronavirus research (Gaur et al., 2021; Gozzi et al., 2020). All user comments connected to these posts were also collected (*N* = 58,990). While it should be noted that this selective sampling procedure inhibits our ability to generalize findings to all r/coronavirus activity (i.e., all posts and comments), by collecting "top" posts (i.e., highly scored posts) this sampling frame included material deemed particularly noteworthy by the community (e.g., informative, controversial, etc.). Accordingly, this sampling procedure afforded an analysis of community-prioritized (and comparatively more influential) content. Put differently, while conducting an analysis of the entire r/coronavirus corpus could provide greater detail about average posting and commenting behavior, by sampling "top" posts and comments this work focused more directly on high-visibility content, which is perhaps more representative of community activity and sentiment.

After collecting the retroactive sample (March 26), data were collected using a staggered sampling strategy: posts and comments were collected at 8 a.m. on the first collection day (March 27), and collection was shifted an hour later each day until data were sampled at 8 p.m. (April 8). On the following day (April 9), the schedule began again at 8 a.m., and this process was repeated until May 25. This sampling strategy allowed us to capture content without introducing bias toward posts and comments submitted at a certain time of day or day of the week, following procedures used in previous *Reddit* research (Hale & Grabe, 2018). The sample was systematically produced from the sampling frame by including every third post (*N* = 226), reducing the time commitment required for human coding. In addition, the top ten comments for each post were included in the sample, after removing "deleted", "removed", or automated comments (*N* = 2260). Figure 1 visualizes this selection process.

Figure 1. A visualization of the sampling procedure used for posts and comments.



Several major events should be noted to situate this sampling frame within the pandemic's timeline. The retroactive collection began on February 25, when the Centers for Disease Control and Prevention (CDC) reported that COVID-19 was heading toward pandemic status (worldwide spread had not yet been achieved), and thus represents the initial period of understanding (and communicating) the pandemic's magnitude (see AJMC, 2021). Daily collection began on March 26, as this date marked several tragic landmarks: globally, the number of infections surpassed 500,000 and the death total reached 22,000, while the United States (a major pandemic epicenter) reported 10,000 new daily infections and 1,000 deaths, indicating that the disease was firmly entrenched. The sample was completed on May 25, three months after starting data collection and three days before the CDC reported 100,000 confirmed COVID-19 deaths in the United States. Beyond capturing the rapid expansion of the disease, this sample also includes numerous developments, including the World Health Organization formally declaring COVID-19 a global pandemic (March 11), United States President Trump declaring COVID-19 a national emergency (March 13), stay-at-home orders emerging throughout the United States (beginning March 19), the CARES Act becoming law (March 27), and pharmaceutical

companies receiving funds from the federal government to develop COVID-19 vaccines (starting around May 21) (AJMC, 2021). As noted previously, despite these rapid changes few resources were designated to curating COVID-19 information during this period, contributing to the expansion and perceived utility of the r/coronavirus subreddit.

Coding Procedures

A codebook was constructed to capture post and comment features, and was iteratively redesigned throughout coder training to improve accuracy and facilitate intercoder reliability. Three coders familiar with *Reddit* were trained using the codebook (ensuring familiarity with *Reddit*-specific language), and intercoder reliability was assessed using Krippendorff's alpha (range = 0.71-1.0). Overall, 50 posts (22.1% of the sample) and 515 comments (22.8%) were randomly selected and included in coder training, and 25 posts (11.1%) and 245 comments (10.8%) were used in calculating Krippendorff's alpha (overall $\alpha = 0.87$). The initial training subsample not used in calculating Krippendorff's alpha (i.e., the sample used to orient coders, refine content categories, discuss inconsistencies, and improve reliability) was recoded after completing the training set. After achieving acceptable intercoder reliability, the sample was equally shared between the three coders. During the coding process, *Reddit* posts were accessed via a stored hyperlink and categorized, while comments were accessed via a spreadsheet generated during data collection. All content categories were coded across text and visuals (if available). Problem cases were flagged and collectively discussed.

Measures

Post Categories

Two post factors were captured during data collection: *score* (i.e., the aggregated post score calculated by subtracting the number of downvotes from the number of upvotes) and *number of comments* (i.e., the number of comments that responded to the post). Six categories were human coded: *continent*, *scope*, *sentiment*, *collectivism*, *individualism*, and *thematic category*. The continental focus of each post was assessed with the *continent* category ($\alpha = 0.89$), which included each of the seven continents as possibilities, plus options for "global" (e.g., posts about multiple continents or global issues) and "none" (e.g., posts detailing COVID-19 symptoms). Geographic *scope* ($\alpha = 1.0$) was assessed using nine levels: individual (e.g., a story about a particular patient), local (e.g., church congregations, prisons, or schools), city (e.g., current events in New York City), county (e.g., information about infection rates in a particular county), state (e.g., statements by US governors), regional

(e.g., multistate agreements to coordinate response), national (e.g., China's response to COVID-19), continental (e.g., spread of COVID-19 across Europe), and global (e.g., international health efforts). These two categories were assessed to understand the geographic focus of r/coronavirus activity, providing descriptive nuance to thematic findings. *Sentiment* (α = 0.71) was categorized as "negative", "positive", or "neutral", and captured the overall tone of the post. For cases of equivalent positive and negative tone, the post was coded as neutral. *Collectivism* (α = 0.74) was considered present when a post framed COVID-19 and/or related issues as collectivistic (i.e., systemic) in nature, including causes of problems and solutions to them. Conversely, *individualism* (α = 0.73) was coded when the post framed COVID-19 and/or related issues individualistically (i.e., an individual is responsible). Collectivism and individualism were not mutually exclusive, and both were categorized using a dichotomous yes/no response.

Thematic category was inductively developed through a systematic assessment of the sampling frame (*N* = 680) in which two authors exhaustively assessed captured r/coronavirus posts for distinct themes across three rounds of categorization. An initial thematic assessment resulted in ten categories, including themes related to government (local and national), business, essential and medical workers, and disease spread. A second round of categorization expanded to thirteen categories, separating private sector responses to COVID-19 (related to business operations) and economic implications (focused on state, national, and international economies), adding categories for positive news and shaming (initially included in an overused miscellaneous category), and introducing a category for public health messaging. The final assessment resulted in 12 distinct post themes that captured most content variability (see Table 1). Each thematic category was independently assessed using a dichotomous yes/no response, and thus were not mutually exclusive. Nearly all posts were categorized within at least one of the identified categories (*N* = 207; 91.6%). Exceptions included a post suggesting the importance of face-to-face schooling, another documenting monetary donations from Ireland to the Navajo nation, and a post suggesting that a Texas church cancelled services after a priest's death due to COVID-19.

Table 1. Thematic post categories.

Category	Description	α
Disease Spread	Information about the spread of the disease ("we're now at 60,000 infections in the US"), current/future mortality rates ("100,000 deaths are expected"), or particular deaths ("Elizabeth Warren's brother has died of COVID-19").	0.75
Treatments/Vaccines	Information about possible treatments (e.g., hydroxychloroquine) or upcoming vaccines (new vaccines, testing of vaccines, number of possible candidates, etc.).	1.00
Business	Information concerning movements in the private sector designed to combat COVID-19 (e.g., production of ventilators, vaccines, masks).	1.00
Economy	Information about the economic implications of COVID-19, including state ("Georgia to reopen economy"), national ("Germany is shut down"), and/or international ("Europe in disagreement about reopening trade") economies.	1.00
Politics	Information about politicians and the ramifications of COVID-19 for politics (e.g., campaigning, elections, approval ratings).	0.84
Social/Environmental	Information about social (e.g., racism) or environmental issues (e.g., coal production, environmental changes) stemming from COVID-19.	0.85
Medical Workers	Information about the experiences of medical workers, including the availability of personal protective equipment (PPE) and medical devices (e.g., ventilators).	0.89
Essential Workers	Information about the experiences of essential workers (e.g., grocery employees, delivery drivers, factory workers, etc.), including healthcare, sick leave, and safety protocols.	1.00

Public Health	Information about public health protocols (e.g., social distancing, wearing masks, quarantining, testing).	0.95
Misinformation	Corrections to misinformation, including information shared by government officials (e.g., President Trump), private corporations, and/or everyday people.	0.85
Positive News	Anecdotal news designed to uplift the community (e.g., "Elderly couple recover from COVID-19", "Child with cancer recovers from coronavirus").	0.90
Shaming	Instances of shaming for defying public health orders, including individuals (e.g., "man defies orders and infects 30 people"), groups (e.g., protesters, churches), and others.	1.00

Post Categories

Similarly to posts, comment *score* was captured during data collection. However, four categories were content analyzed: *sentiment*, *collectivism*, *individualism*, and *comment type*. Following the same procedure used for posts, comment *sentiment* (α = 0.82) was coded as "negative", "positive", or "neutral", *collectivism* (α = 0.81) was coded dichotomously for instances of collectivistic framing, and *individualism* (α = 0.76) was categorized when comments framed COVID-19 and/or related issues individualistically. Lastly, comments were categorized according to their *comment type* – *informational* (α = 0.83), *speculative* (α = 0.78), or *reactive* (α = 0.84). Comments were coded as *informational* when they shared or curated information (e.g., compiling related information, making connections between items, sharing news), or attempted to correct misinformation. Comments were categorized as *speculative* when they provided non-factual speculation about something (e.g., motivations behind decisions, potential outcomes). Finally, comments were coded as *reactive* when they responded to post content without providing additional information (e.g., "It's amazing how little we know", "how are they managing to screw this up so badly?"). These three categories were mutually exclusive.

Results

Descriptives

Overall, the top posts submitted to r/coronavirus between February 25 and May 25, 2020 received relatively high scores (aggregate scores subtracting "downvotes" from "upvotes") (M = 20,197.87),

though significant variance existed between posts (SD = 21,538.66; range = 1,290-111,742). Commenter engagement was also noteworthy, with posts receiving an average of 965.78 comments (SD = 1,094.04; range = 65-7,138). Comments received an average score of 636.83 and exhibited significant variation (SD = 1,397.41; range = 2-15,897). Posts and comments yielded a similar bias toward collectivistic framing, with 62.83% of posts (63.23% of comments) categorized as collectivistic, and 37.61% (26.33%) as individualistic. This suggests that users of r/coronavirus primarily framed COVID-19 and related issues collectivistically. Regarding sentiment, posts and comments were most frequently coded as negative (58.41% of posts; 60.8% of comments), though posts included more instances of positive sentiment (25.66%) than neutral (15.93%), while comments were more neutral (21.33%) than positive (17.88%). Thus, negativity was particularly pronounced in comments (average sentiment = -0.42; scores ranged from -1 (negative) to 1 (positive)).

Additionally, posts exhibited a clear North American focus (75.22%), which makes sense given the U.S. focus of the *Reddit* platform (Statista, 2021), with fewer posts targeting Europe (12.39%), global issues (5.31%), Asia (3.54%), Australia (1.77%), or no continental affiliation (1.77%). No posts in this sample addressed South America, Africa, or Antarctica. Posts also focused most prominently on national issues (41.45%), followed in scope by individual (23.89%), state (11.5%), local (9.73%), city (6.64%), global (5.31%), regional (1.33%), and county (0.44%). Taken together, posts submitted to r/coronavirus most commonly focused on national North American issues (typically the US), with a secondary focus on individuals (e.g., doctors, nurses, patients, and politicians). Comments were most frequently categorized as reactive (45.18%), followed by informational (30.8%) and speculative (23.98%), suggesting that commenters responded to post details more frequently than offering new information (or speculation).

Post Themes

Our first research question concerned what thematic categories appeared most frequently in r/coronavirus posts during the early pandemic, and which categories received the most engagement. Table 2 outlines the frequency of each theme, plus averaged values for engagement (i.e., score and number of comments), sentiment, and collectivism (posts and comments). A comparison of scores for posts categorized within the lower frequency categories (i.e., those falling below the median category frequency: social/environmental, treatment, misinformation, essential workers, positive news, and business) with those not classified within these categories was significant, $t(193.76) = 1.99$, $p < .05$. A secondary test comparing

posts of higher frequency categories (i.e., those falling above the median category frequency: economy, shaming, medical workers, politics, public health, and disease) to other posts was nearly significant, $t(115.93) = -1.81$, $p = .07$[1]. Additionally, low frequency posts received significantly fewer comments, $t(163.15) = 2.53$, $p = .01$, while high frequency posts garnered significantly more comments, $t(126.98) = -2.36$, $p = .02$. Regarding sentiment, low frequency posts were more positive at nearly significant levels, $t(138.11) = -1.86$, $p = .06$, and high frequency posts were significantly less positive, $t(87.04) = 4.75$, $p < .001$. Considered together, these findings suggest that users of r/coronavirus prioritized thematic categories that prompted negative responses, indicated by increased posting frequency, higher aggregate score, and a greater number of comments. Moreover, these findings indicate that r/coronavirus users were particularly interested in information about disease spread, public health recommendations, politics (typically in the United States), the experiences of medical workers, and economic implications of COVID-19. Users also showed an interest in shaming individuals who failed (or refused) to follow public health guidelines (e.g., mask weaking and social distancing).

Because posts could be categorized within multiple themes (i.e., themes were not mutually exclusive), two separate analyses were conducted to evaluate low and high frequency categories. This allowed posts included in both low and high frequency categories (e.g., a post that simultaneously addresses treatment information and politics) to be evaluated in each analysis independently.

Table 2. Post and Comment Features by Thematic Category (N = 226)

Category	*N*	%	Score	# of Comments	Sentiment	Post Collect %	Collect %
Social/Environ.	4	1.77	16169.75	269.50	**-0.40**	**100.00**	**70.00**
Treatment	6	2.65	5236.50	675.67	-0.55	50.00	**65.00**
Misinformation	8	3.54	**23717.67**	884.22	**-0.18**	**88.89**	**71.11**
Essential Work	14	6.12	**22657.43**	**1037.64**	-0.46	**64.29**	**67.14**
Positive News	20	8.85	14319.70	358.05	**0.12**	20.00	31.00
Business	23	10.18	18368.74	759.00	-0.50	**65.22**	**70.87**
Low Freq.	69	30.53	16567.75	715.16	**-0.34**	53.62	57.83
Economy	24	10.62	**25652.50**	**1168.75**	-0.59	**91.67**	**77.08**
Shaming	30	13.27	19546.93	921.00	-0.80	43.33	**63.67**
Medical Work	32	14.16	**21834.91**	747.75	**-0.40**	40.63	52.81
Politics	39	17.26	**22967.08**	**1248.08**	-0.51	46.15	58.21
Public Health	53	23.45	**27258.36**	**1466.55**	-0.53	**75.47**	**67.17**
Disease Spread	54	23.89	19301.20	911.28	-0.55	**70.37**	**69.63**
High Freq.	166	73.45	**21669.63**	**1058.81**	-0.53	**65.06**	**65.18**
Not Categorized	19	7.14	15844.53	572.00	**0.13**	**68.42**	62.63

Note. Bolded figures fall above dataset means: score (20197.87), number of comments (965.78), comment sentiment (-0.43), post collectivism (62.83%), and comment collectivism (63.23%). Low Freq. = aggregated figures for all posts categorized within low frequency categories (i.e., those falling below the median category frequency). High freq. = figures for all posts categorized within high frequency categories (i.e., those falling above the median category frequency). Not categorized = posts not categorized within any specified category.

The relationship between collectivism and category frequency is somewhat less straightforward than sentiment. No significant findings emerged for high frequency categories for post collectivism ($t(100.58) = -1.18$, $p = .26$) or comment collectivism ($t(90.44) = -1.62$, $p = .11$). Low frequency categories included less collectivism at both the post level ($t(122.97) = 1.86$, $p < .07$) and comment level ($t(127.76) = 1.94$, $p < .06$), though these findings are nearly significant. Of particular note, however, is that the findings for low frequency categories are strongly influenced by the lack of collectivism (and increased individualism; $M = 43\%$ of comments) in the positive news category. If removed, the remaining low frequency categories yield no significant findings ($p > .26$ for post and comment collectivism). Instead, collectivism may be better understood by comparing individual categories, with some categories producing higher levels of collectivistic discourse (e.g., economy, misinformation, and public health) and others facilitating individualism (e.g., positive news, medical workers, and politics).

Extending this analysis to the relationship between posts and comments, the presence of a collectivistic frame in posts significantly correlated with collectivism in comments ($r(224) = .54$, $p < .001$),

suggesting that posts set the stage for comment discourse, which agrees with previous scholarship (e.g., Hale, 2017). Further bolstering this interpretation is a similar correlation between post and comment sentiment ($r(224) = .61$, $p < .001$). Therefore, while collectivistic attribution was generally widespread in r/coronavirus discourse, thematic categorization seemed to affect the uniformity of collectivistic (or individualistic) orientation.

Comment Category

Concerning our second research question, which asked about the informational utility of r/coronavirus discourse, comments were categorized as informational, speculative, or reactive (see Table 3). As previously noted, comments were most frequently categorized as reactive (45.18%), followed by informational (30.8%) and speculative (23.98%). This indicates that commenters responded to post content more frequently than providing new information, though a noteworthy number of comments were informational. Moreover, although nearly 24% of comments contained some speculative element, this was the least common category, suggesting that the moderation policies of r/coronavirus were at least somewhat effective.

To assess whether comment categorization was associated with score (an indication of community support), we conducted a one-way ANOVA. The result indicated that this categorization significantly predicted comment score ($F(2, 2257) = 5.76$, $p < .01$), and post-hoc tests revealed that informational comments scored significantly higher ($M = 786.07$) than speculative ($M = 568.61$, $p < .02$) and reactive ($M = 571.38$, $p < .01$) comments, while the speculative and reactive categories were not significantly different from one another ($p = .99$). Therefore, although informational comments were less frequent than reactive, they received higher scores. Perhaps interestingly, the post categories receiving the lowest percentages of collectivism (i.e., positive news, shaming, treatment, medical workers, and politics; see Table 2), elicited the fewest informational comments, but the highest percentage of reactive comments. Conversely, the post categories coded most frequently as collectivism (i.e., public health, business, economy, social/environmental, disease spread, misinformation, and essential workers; see Table 2) received the highest percentage of informational comments, but the lowest number of reactive comments. Indeed, post collectivism was correlated with comment information ($r(224) = .14$, $p < .03$), and negatively correlated with comment reaction ($r(224) = -.22$, $p < .001$). Perhaps collectivistic orientation empowered r/coronavirus users to provide information about COVID-19 to the community, while an individualistic framework hindered information generation and promoted reactive discourse.

Table 3. Comment Category by Thematic Category

	Comment Category		
Post Category	Informational	Speculative	Reactive
Positive News	22.50	18.50	**59.00**
Shaming	26.00	**27.00**	**47.00**
Treatment	26.67	**25.00**	**48.33**
Medical Workers	29.06	21.25	**49.69**
Politics	29.23	**25.64**	44.87
Public Health	**31.51**	**25.09**	43.21
Business	**35.22**	19.57	**45.22**
Economy	**34.58**	**32.92**	32.50
Social/Environmental	**35.00**	22.50	42.50
Disease Spread	**36.48**	23.89	39.63
Misinformation	**36.67**	**30.00**	33.33
Essential Workers	**41.43**	20.71	37.86

Note. All values represent the percentage of comments within each category. Bolded figures fall above dataset means: informational (30.8%), speculative (23.98%), and reactive (45.18%).

Our third research question asked about the prevalence of individualistic and collectivistic framing in r/coronavirus comments, and has been addressed throughout the previous results. To summarize, collectivism emerged in a majority of posts (62.83%) and comments (63.23%) and was inversely related to individualism (37.61% and 26.33%, respectively). Average comment collectivism dropped below 50% only in response to the positive news category, and average post collectivism for four categories (shaming, medical workers, politics, and positive news). Therefore, while some categorical differences emerged, the majority of r/coronavirus discourse generated during the early pandemic was collectivistically oriented, indicating reinforcement of public health messaging during the early COVID-19 pandemic (e.g., universal mask use, social distancing, and self-quarantining).

COVID-19 Fatigue

Our final research question asked if evidence of fatigue (i.e., decreasing user interaction) can be found in r/coronavirus activity during the early pandemic. To test this possibility, two linear regressions were constructed to determine if time significantly predicted post scores and/or number of comments. The regressions for both score, $F(1, 122) = 13.94$, $p < .001$, $R^2 = .10$, and number of comments, $F(1, 122) = 13.94$, $p < .01$, $R^2 = .06$, were significant. Accordingly, posts received lower scores over time ($t = -3.73$, $p < .001$) and fewer comments ($t = -3.02$, $p < .01$), indicating progressively decreasing interactions with r/coronavirus posts. This

indeed suggests that COVID-19 fatigue had begun to emerge between February and May, 2020 in the r/coronavirus subreddit, which reinforces previous findings (Gozzi et al., 2020). Considering that r/coronavirus users were particularly committed to curating COVID-19 information during the early pandemic, evidence of fatigue during this window is troublesome, as this suggests faltering motivation to remain engaged even during the early stages of the pandemic.

Figure 2. Post score over time.

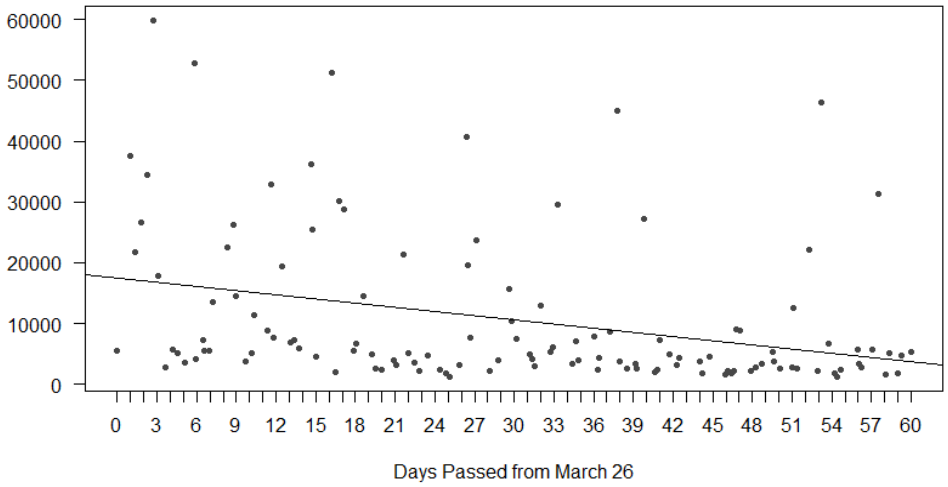
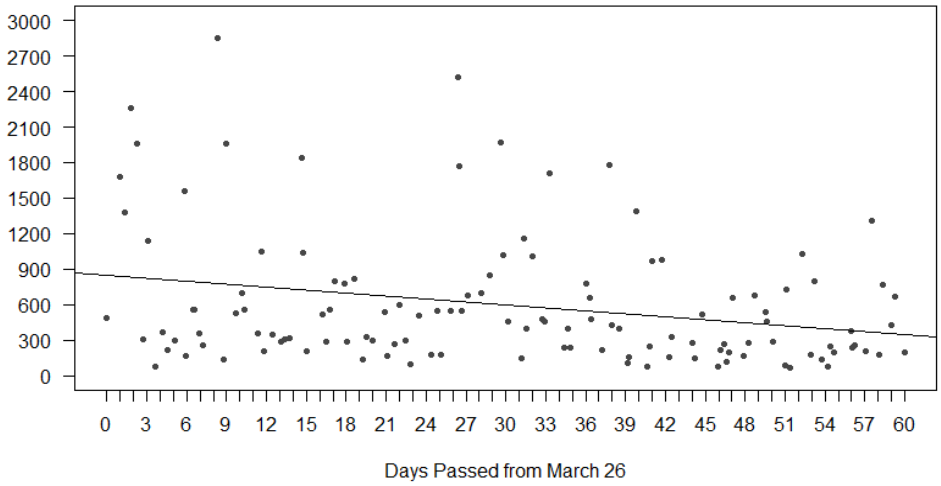


Figure 3. Number of comments over time.



Discussion

When confronting a disease outbreak, people seek potentially life-saving information, including details about prevention and treatment, personal health risks, and vaccines (Henrich & Holmes, 2011;

Springer et al., 2020; Wong & Sam, 2010). According to RISP (Dunwoody & Griffin, 2015; Griffin et al., 1999), individuals seek this information to achieve desired informational goals (e.g., what is COVID-19?), assess risk (e.g., how dangerous is COVID-19?), and render behavioral decisions (e.g., should I wear a mask?), and purposefully seek out "high quality" information channels. Previous research suggests that media platforms commonly emerge as important sources of information **(**Henrich & Holmes, 2011; Strekalova, 2017; Wong & Sam, 2010), and social media platforms like r/coronavirus became particularly important during the early months of the COVID-19 pandemic. The emergence of r/coronavirus was partly indebted to *Reddit*'s popularity during this period, the affordances of the platform (e.g., bidirectional voting which allows content filtering), and appearances by trusted public health authorities (e.g., AMAs with Bill Gates and numerous global medical experts). However, despite the popularity of this subreddit and its purported informational utility, little research has systematically examined what information was disseminated and how users of this platform described the pandemic (for exceptions see Gozzi et al., 2020; Lai et al., 2020). This is an important gap for public health researchers, as the emergence of r/coronavirus during the early pandemic combined with the unresponsiveness of governmental organizations (e.g., the CDC) provides a unique window to examine pandemic communication amidst an information vacuum.

Accordingly, through a content analysis of 226 posts and 2260 corresponding comments, this work identified patterns of communication in r/coronavirus during the early pandemic, including what information was deemed important and how users framed causes and solutions. Overall, the findings of this study suggest that users of r/coronavirus prioritized information about COVID-19 spread, public health information (e.g., mask mandates and social distancing recommendations), political and economic implications of COVID-19, and the experiences of medical workers (all signifiers of risk), while also shaming people who failed to follow public health guidelines (perhaps establishing norms and motivating behavioral conformity). Discourse was collectivistically oriented and negatively valenced, and engagement with the subreddit decreased over time, suggesting COVID-19 fatigue among r/coronavirus users. Taken together, this study provides a window into early pandemic discourse, furthering our understanding of COVID-19 communication in early 2020 and revealing informational needs that could be targeted during future health emergencies. Moreover, these findings elucidate *Reddit*'s role in information curation and dissemination during the early COVID-19 pandemic, contributing to a growing body of literature illustrating the potential utility of social

media platforms during public health emergencies.

The first goal of this work was to categorize posts submitted to r/coronavirus thematically (see Table 1) to determine what information was important to users during the early months of the COVID-19 pandemic (e.g., February – May, 2020). Relatedly, this study also sought to determine differences in user engagement (i.e., upvotes and comments) between thematic categories. Our findings indicate that users prioritized information related to six categories: disease spread, public health recommendations, economic and political activity, experiences of medical professionals (e.g., doctors and nurses), and shaming of individuals who failed to conform to public health measures (e.g., not wearing a mask, social distancing, or self-quarantining). Compared to the remaining thematic categories (i.e., business activities, positive anecdotal news, experiences of essential workers, corrections of misinformation, treatments, and social/environmental issues), posts within these themes appeared more frequently, received higher overall scores, and received more user comments. Moreover, high frequency categories scored lower in sentiment, meaning that r/coronavirus users prioritized negatively valenced information (e.g., disease spread, economic hardship, exhausted medical workers).

These findings, in which post category frequency, score, and number of comments exhibited complementary patterns, suggest that certain forms of negative information are particularly important for the public during a health emergency's infancy, and are consistent with previous scholarship. Research examining crisis events indicates that negatively valenced information is more quickly diffused than positive or neutral information on social media (Ferrara & Yang, 2015; King & Wang, 2021), which could indicate a need to gather information related to risk (Dunwoody & Griffin, 2015). While the underlying mechanisms of this tendency toward negative (and risk-signifying) themes cannot be found within the current research design, prior literature in human information processing research indicates that "the human brain has an *automatic* attentive response to negatively compelling stimuli" (Grabe & Kamhawi, 2006, p. 347), which is unvarying regardless of the media type (Blake et al., 2001; Lang et al., 1996; Reeves & Nass, 1996). Therefore, public health experts may want to consider the public's particular attention to negative information, strategically targeting information relevant to related themes (e.g., disease spread, experiences of medical personnel) during a pandemic's emergence, affording greater control over emerging discourse (e.g., public health narratives) and encouraging protective behaviors.

Another goal of this work was to examine the prevalence of

collectivistic and individualistic framing in r/coronavirus discourse, as previous scholarship indicates that framing can exhibit significant influence on beliefs, attitudes, and behaviors (Barrile, 1986; Iyengar, 1991; Maaravi et al., 2021; Viswanath et al., 2015). Moreover, framing can be used to satisfy informational needs about behavioral responsibility, especially during periods of uncertainty (e.g., "am I responsible for doing something?" and "should I wear a mask?"). Given the COVID-19 pandemic's evolution into a global issue, and the collectivistic responsibility implicit in public health messages (e.g., universal mask use), collectivistic attribution was perhaps likely. However, the centricity of the United States for the *Reddit* platform (a highly individualistic society) could have instead elicited individualistic framing. In fact, posts and comments yielded a similar bias toward collectivism, with a majority of posts and comments (62.83% and 63.23%, respectively) categorized as collectivistic, and a minority (37.61% and 26.33%, respectively) as individualistic. This indicates that users of r/coronavirus primarily framed COVID-19 and related issues collectivistically, suggesting the emergence of collective behavioral norms (e.g., shared responsibility for mask wearing). However, some differences emerged between post themes, with seven categories yielding heightened levels of collectivism: disease spread, public health recommendations, economic implications of COVID-19, business activities, experiences of essential workers, corrections to misinformation, and social/environmental issues. Therefore, pandemic messaging may generally benefit by continuing to employ a collectivistic frame (matching an existing orientation in the population), although communicators interested in targeting some informational categories (e.g., treatment information, positive news, experiences of medical workers, and political ramifications) might consider incorporating elements of individualistic framing.

In addition to categorizing r/coronavirus comments as collectivistic/individualistic, we also classified comments in three groups: those that 1) provided information about COVID-19, 2) speculated about the disease response or related issues (e.g., motivations behind decisions), or 3) reacted to post content without providing any additional information or speculation. Given that the subreddit's rules encouraged information curation and discouraged speculative discussion (the community description states that "this subreddit is for high-quality posts and discussion"), and the need of individuals to collect information when facing a perceived hazard (Dunwoody & Griffin, 2015), a bias toward informational comments could be expected. However, reactive comments emerged most frequently (45.18%), followed by informational (30.8%) and speculative (23.98%), suggesting that commenters responded to

post details more frequently than offering new information (or speculation). While this might indicate limited informational utility of the subreddit, reactive comments may have served an important role alongside post and comment information. While reactive comments – defined here as responses to post content without providing additional information or speculation – might not provide information directly, they frequently bolstered post information by vocalizing distress (e.g., "we're so f'ed", "this is scary"), frustration (e.g., "you can't make this up", "I guess Trump did not contain the virus just yet"), and outrage (e.g., "there are zero excuses to let this continue", "WTF is wrong with people?"). In other words, many comments supported post information through reaction, potentially reinforcing the legitimacy of information and promoting acceptance (i.e., bolstering the perceived "quality" of the r/coronavirus channel). To provide a further example, one commenter responded to a post about the United States' pandemic response, stating "When I die, I want [Vice President] Mike Pence to bury me at my funeral. So that he can let me down one last time." While potentially introducing humor into the discourse (and receiving many upvotes), this comment also reinforces Mike Pence's culpability for the lackluster US response specified by the post.

Of additional note, the proportion of informational, speculative, and reactive categories differed by post theme, with high-collectivism themes (i.e., positive news, shaming, treatment, medical workers, and politics), eliciting the highest percentage of informational comments and the fewest reactive comments, and low-collectivism themes (i.e., public health, business, economy, social/environmental, disease spread, misinformation, and essential workers) receiving the lowest percentage of informational, but the highest number of reactive comments. One interpretation for this finding is that collectivistic orientation empowered users to provide information (due to a shared responsibility), while an individualistic framework instead promoted reactions (e.g., declarations of frustration). This interpretation agrees with previous scholarship, as collectivism has been found to facilitate knowledge sharing (Bao, Zhang, & Chen, 2015; Pian, Jin, & Li, 2019). Therefore, the generally collectivistic orientation of r/coronavirus may have facilitated information sharing, serving the curational responsibility assigned by the moderation team and reinforcing the informational utility of this subreddit during the early pandemic.

Furthermore, the data reported here supports the emergence of COVID-19 fatigue in r/coronavirus, supporting other previous early-pandemic research (Gozzi et al., 2020), as posts received progressively lower scores and fewer comments (i.e., decreasing interactions) from February to May, 2020. Given that r/coronavirus

"seeks to monitor the spread of the disease" (r/coronavirus, 2021), and became the primary source of COVID-19 information on *Reddit* (and a resource for users beyond the platform; Solon & Glaser, 2020), evidence of fatigue in this subreddit is troubling. This highly motivated community struggled to maintain engagement, following the same fatigue patterns documented elsewhere (Koh et al., 2020; World Health Organization, 2020), including during previous health emergencies (Pruss et al., 2019; Tausczik et al., 2012), suggesting that continuous attention to pandemic information is difficult, even when motivated (perhaps due to information saturation or overload). One possibility is that users progressively surpassed their information sufficiency threshold (Dunwoody & Griffin, 2015), determining that they had enough information to make informed behavioral decisions. Therefore, in the case of a prolonged pandemic, public health communicators will need to identify ways to continuously motivate listeners (e.g., demonstrating a need for additional information). Strategically targeting the informational needs identified here could help, but additional research is needed to determine what strategies might directly moderate fatigue.

Limitations and Future Directions

Although this study provided an overview of r/coronavirus activity during the early COVID-19 pandemic, furthering our understanding of COVID-19 communication and informational needs, a few limitations and recommendations for future research should be noted. First, the sample collected for this study only allowed examination of r/coronavirus activity between February 25 and May 25, 2020. While we believe this timeframe is important for understanding informational needs during a crucial moment in the disease's emergence (e.g., new daily infections in the United States surpassed 10,000 and 1,000 deaths were reported), this sample does not cover the earliest days of the pandemic (e.g., initial diagnosis in China) or the months following May 25 (e.g., spikes in cases during July and November). Additional work could examine activity beyond the timeframe analyzed here, allowing for longitudinal analyses of information needs and COVID-19 fatigue.

Second, some of this study's findings regarding post sentiment, individualism, and collectivism may need to be evaluated cautiously. Reliability for these categories was somewhat lower than the 0.8 alpha threshold recommended by Krippendorff (2013) (α = 0.71, 0.73, and 0.74, respectively). However, it may be noted that comment sentiment and collectivism were both reliable (both α = 0.81), indicating that coders were capable of assessing these categories, though slightly less reliably within posts. Intercoder reliability for comment individualism still fell below the recommended threshold (α = 0.76). To counteract this reduced reliability, coders

were instructed to note problem cases (i.e., posts and comments for which they were not certain about their categorization), and these cases were collectively discussed and categorized. Despite this additional measure, Krippendorff (2013) suggests that categories scored between 0.667 and 0.80 may be used for drawing tentative conclusions (other scholars have instead suggested 0.7; Lacy et al., 2015), and thus findings related to these categories should be considered carefully.

Additionally, this content analysis focused on posts and root comments (i.e., comments that responded directly to posts), excluding response comments (i.e., comments that respond to root comments). While this decision focused our analysis on comments directly related to posts, future research could examine dialogue within comment threads, providing additional nuance to our understanding of r/coronavirus discourse. Finally, our findings indicate that individualistic frames emerged in approximately 38% of posts and 26% comments. While significantly less frequent than collectivism, future analyses of individualistic frames could prove elucidating. Specifically, further work could identify specific targets of blame (e.g., President Donald Trump), allowing for a more nuanced understanding of causal attribution among r/coronavirus users. Overall, more work is needed to unravel communication in r/coronavirus, furthering our understanding of user-curated pandemic information and social media's role during the COVID-19 era.

Brent J. Hale (PhD, Indiana University) is an Assistant Professor of Communication Studies at the University of Southern Mississippi. His research focus includes interlocutory and psychological responses to social media content through commenting, including messages related to health and politics.

Mark Alberta (MA, DePaul University) is a PhD candidate in the Media School at Indiana University. His research generally focuses on media effects and psychology, including specific interests in morality, emotions, and video games.

Seung Woo Chae (MBA, Korea Advanced Institute of Science and Technology) is a Ph.D. candidate at Indiana University and former television producer at Korean Broadcasting System (KBS). His research focuses on computer-mediated communication in online comments, especially on online video platforms including Twitch and YouTube.

References

Abutaleb, Y., Dawsey, J., Nakashima, E., & Miller, G. (2020, April). The U.S. was beset by denial and dysfunction as the coronavirus raged. *The Washington Post*. https://www.washingtonpost.com/national-security/2020/04/04/coronavirus-government-dysfunction/?arc404=true

American Journal of Managed Care (AJMC) (2021, January). *A timeline of COVID-19 developments in 2020*. https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020

Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, *6*(2), 2053168019848554. https://doi.org/10.1177/2053168019848554

Bao, G., Zhang, Z., & Chen, J. (2015). The mediation of in-group identification between collectivism and knowledge sharing. *Innovation*, *17*(3), 341-363. https://doi.org/10.1080/14479338.2015.1054605

Barrile, L. (1986). Television's "bogeyclass"? Status, motives, and violence in crime drama characters. *Sociological Viewpoints, 2*(1), 39–56.

Blake, T., Varnhagen, C., & Parent, M. (2001). Emotionally arousing pictures increase blood glucose levels and enhance recall. *Neurobiology of Learning and Memory*, *75*, 262-273. https://doi.org/10.1006/nlme.2000.3973

Bloomberg (2021, January). *India promises free vaccines; Sydney orders masks: Virus update.* https://www.bloomberg.com/news/articles/2020-12-31/n-y-and-california-end-year-shattering-records-virus-update

Bloomberg (2021, January). *Mapping the coronavirus outbreak across the world.* https://www.bloomberg.com/graphics/2020-coronavirus-cases-world-map/

Booker, B. (2020, December). *The coronavirus has reached Antarctica. Now it's on every continent.* NPR. https://www.npr.org/sections/coronavirus-live-updates/2020/12/23/949552848/the-coronavirus-has-reached-every-continent-after-positive-cases-inantarctica

Boukes, M., Boomgaarden, H. G., Moorman, M., & de Vreese, C. H. (2015). Political news with a personal touch: How human interest framing indirectly affects policy attitudes. *Journalism & Mass Communication Quarterly*, *92*(1), 121–141. https://doi.org/10.1177/1077699014558554

Brennen, J. S., Simon, F., Howard, P. N., & Nielsen, R. K. (2020). Types, sources, and claims of COVID-19 misinformation. *Reuters Institute*, *7*, 3-1.

Brennen, J. S., Simon, F. M., & Nielsen, R. K. (2020). Beyond (mis) representation: Visuals in COVID-19 Misinformation. *The International Journal of Press/Politics*. https://doi.org/ 10.1177/1940161220964780

Bruno Kessler Foundation (2020). *Fake news in the time of C-19*. https://members.tortoisemedia.com/2020/03/23/the-infodemic-fake-news-coronavirus/content.html

CDC (2020a, December). *COVID-19: Symptoms*. https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html

CDC (2020b, October). *Overview of testing for SARS-CoV-2 (COVID-19).* https://www.cdc.gov/coronavirus/2019-ncov/hcp/testing-overview.html

CDC (2021, January). *CDC COVID data tracker*. https://covid.cdc.gov/covid-data-tracker/#vaccinations

Dunwoody, S., & Griffin, R. J. (2015). Risk information seeking and processing model. In H. Cho, T. Reimer, & K. A. McComas (Eds.), *The SAGE handbook of risk communication* (pp. 102-116). Thousand Oaks, CA: Sage.

Dwyer, C., & Aubrey, A. (2020, April). CDC now recommends Americans consider wearing cloth face coverings in public. *NPR*. https://www.npr.org/sections/coronavirus-live-updates/2020/04/03/826219824/president-trump-says-cdc-now-recommends-americans-wear-cloth-masks-in-public

Elfrink, T., Guarino, B., & Mooney, C. (2020, September). CDC reverses itself and says guidelines it posted on coronavirus airborne transmission were wrong. *The Washington Post*. https://www.washingtonpost.com/nation/2020/09/21/cdc-covid-aerosols-airborne-guidelines/

Fan, J., Liu, X., Pan, W., Douglas, M. W., & Bao, S. (2020). Epidemiology of coronavirus disease in Gansu Province, China, 2020. *CDC Research*, *26*(6). https://stacks.cdc.gov/view/cdc/88897

FDA (2021, January). *COVID-19 vaccines*. https://www.fda.gov/emergency-preparedness-and-response/coronavirus-disease-2019-covid-19/covid-19-vaccines

Ferrara, E., & Yang, Z. (2015). Quantifying the effect of sentiment on information diffusion in social media. *PeerJ Computer Science*, *1*, e26. https://doi.org/10.7717/peerj-cs.26

Gaur, M., Roy, K., Sharma, A., Srivastava, B., & Sheth, A. (2021). " Who can help me?": Knowledge Infused Matching of Support Seekers and Support Providers during COVID-19 on Reddit. *arXiv:2105.06398*.

Germani, A., Buratta, L., Delvecchio, E., & Mazzeschi, C. (2020). Emerging adults and COVID-19: The role of individualism-collectivism on perceived risks and psychological maladjustment. *International journal of environmental research and public health*, *17*(10), 3497.

Gozzi, N., Tizzani, M., Starnini, M., Ciulla, F., Paolotti, D., Panisson, A., & Perra, N. (2020). Collective Response to Media Coverage of the COVID-19 Pandemic on Reddit and Wikipedia: Mixed-Methods Analysis. *Journal of medical Internet research, 22*(10), e21597.

Grabe, M. E., & Kamhawi, R. (2006). Hard wired for negative news? Gender differences in processing broadcast news. *Communication Research*, *33*(5), 346-369. https://doi.org/10.1177/0093650206291479

Griffin, R. J., Dunwoody, S., & Neuwirth, K. (1999). Proposed model of the relationship of

information seeking and processing to the development of preventive behaviours, *Environmental Research*, *80*, S230-S245.

Hale, B. J. (2017). "+ 1 for Imgur": A content analysis of SIDE theory and common voice effects on a hierarchical bidirectionally-voted commenting system. *Computers in Human Behavior*, *77*, 220-229. https://doi.org/10.1016/j.chb.2017.09.003

Hale, B. J., & Grabe, M. E. (2018). Visual war: A content analysis of Clinton and Trump subreddits during the 2016 campaign. *Journalism & Mass Communication Quarterly*, *95*(2), 449-470. https://doi.org/10.1177/1077699018766501

Henrich, N., & Holmes, B. (2011). Communicating during a pandemic: information the public wants about the disease and new vaccines and drugs. *Health promotion practice*, *12*(4), 610-619. https://doi.org/10.1177/1524839910363536

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage.

Iyengar, S. (1991). *Is anyone responsible? How television frames political issues*. Chicago, IL: University of Chicago Press.

Johns Hopkins University (2021, Janurary). *United States Overview*. https://coronavirus.jhu.edu/region/united-states

Kemp, S. (2020, April 24). Report: Most important data on digital audiences during coronavirus. *Growth Quarters—The Next Web*. https://thenextweb.com/growth-quarters/2020/04/24/report-most-important-data-on-digital-audiences-during-coronavirus/

Kim, M., Hale, B. J., Grabe, M. E., & Bas, O. (2020). Who is responsible? The impact of emotional personalization on explaining the origins of social problems. *Atlantic Journal of Communication*, 1-14. https://doi.org/10.1080/15456870.2020.1752693

King, K. K., & Wang, B. (2021). Diffusion of real versus misinformation during a crisis event: A big data-driven approach. *International Journal of Information Management*, 102390. https://doi.org/10.1016/j.ijinfomgt.2021.102390

Koh, P. K. K., Chan, L. L., & Tan, E. K. (2020). Messaging fatigue and desensitisation to information during pandemic. *Archives of Medical Research*, *51*, 716-717. https://doi.org/10.1016/j.arcmed.2020.06.014

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd edition). Thousand Oaks, CA: SAGE Publications.

Lacy, S., Watson, B. R., Riffe, D., & Lovejoy, J. (2015). Issues and best practices in content analysis. *Journalism & Mass Communication Quarterly*, 92(4), 791-811.

Lai, D., Wang, D., Calvano, J., Raja, A. S., & He, S. (2020). Addressing immediate public coronavirus (COVID-19) concerns through social media: Utilizing Reddit's AMA as a framework for Public Engagement with Science. *PloS one*, *15*(10), e0240326.

Lang, A., Newhagen, J., & Reeves, B. (1996). Negative video as structure: Emotion, attention, capacity, and memory. *Journal of Broadcasting & Electronic Media*, *40*, 460-477. https://doi.org/10.1080/08838159609364369

Maaravi, Y., Levy, A., Gur, T., Confino, D., & Segal, S. (2021). "The tragedy of the commons": How individualism and collectivism affected the spread of the COVID-19 pandemic. *Frontiers in Public Health*, *9*, 37.

Nguyen, M. H., Gruber, J., Fuchs, J., Marler, W., Hunsaker, A., & Hargittai, E. (2020). Changes in digital communication during the COVID-19 global pandemic: Implications for digital inequality and future research. *Social Media+ Society*, *6*(3). https://doi.org/10.1177/2056305120948255

Pian, Q. Y., Jin, H., & Li, H. (2019). Linking knowledge sharing to innovative behavior: the moderating role of collectivism. *Journal of Knowledge Management*, *23*(8), 1652-1672.

Pruss, D., Fujinuma, Y., Daughton, A. R., Paul, M. J., Arnot, B., Albers Szafir, D., & Boyd-Graber, J. (2019). Zika discourse in the Americas: A multilingual topic analysis of Twitter. PloS one, 14(5), e0216922.

r/coronavirus (2021, January). *Reddit*. https://www.reddit.com/r/Coronavirus/

Radu, R. (2020). Fighting the 'infodemic': Legal responses to COVID-19 disinformation. *Social Media+ Society*, *6*(3). https://doi.org/10.1177/2056305120948190

Reddit (2020a, March). *I'm Bill Gates, co-chair of the Bill & Melinda Gates Foundation. AMA about COVID-19*. https://www.reddit.com/r/Coronavirus/comments/fksnbf/im_bill_gates_cochair_of_the_bill_melinda_gates/

Reddit (2020b, March). *I'm Joong Sik Eom MD, a physician treating COVID-19 patients at a hospital near Seoul. I serve on a team advising the South Korean government's pandemic response. AMA.* https://www.reddit.com/r/Coronavirus/comments/fo4pj4/im_joong_sik_eom_md_a_physician_treating_covid19/

Reddit (2020c, March). *I'm a critical care doctor working in a UK high consequence infectious diseases centre. Many units are totally full, and we are scrambling to create more capacity. The initial UK government approach has been a total failure. Ask me anything.* https://www.reddit.com/r/Coronavirus/comments/fnl0n6/im_a_critical_care_doctor_working_in_a_uk_high/

Reddit (2020d, March). *I'm a critical care doctor working in a UK HCID (high consequence infectious diseases) unit. Things have accelerated significantly in the past week. Ask me anything.* https://www.reddit.com/r/Coronavirus/comments/fgfspi/im_a_critical_care_doctor_working_in_a_uk_hcid/

Reddit (2020e, March). *We are four Swiss scientists studying COVID-19/SARS-CoV-2 - AMA!* https://www.reddit.com/r/Coronavirus/comments/fi14db/we_are_four_swiss_scientists_studying/

Reddit (2020f, March). *I'm Ali Raja, MD and Shuhan He, MD emergency physicians from Mass General Hospital/Harvard Medical School. We're back to report from the front lines of COVID-19. Let's talk PPE, new updates & science, testing, quarantine and more. AMA.* https://www.reddit.com/r/Coronavirus/comments/flvu7k/im_ali_raja_md_and_shuhan_he_md_emergency/

Reddit (2020g, March). *I'm Dr. John Torres, medical correspondent for NBC News who practices emergency medicine, and I'm joined by Dr. Joseph Fair, a virologist, epidemiologist and global outbreak responder. We're here to discuss the new COVID-19 coronavirus outbreak. AMA.* https://www.reddit.com/r/Coronavirus/comments/fefj9n/im_dr_john_torres_medical_correspondent_for_nbc/

Reddit (2020h, April). *Misinformation and COVID-19: What Reddit is doing.* https://www.reddit.com/r/ModSupport/comments/g21ub7/misinformation_and_covid19_what_reddit_is_doing/

Reeves, B., & Nass, S. (1996). *The media equation: How people treat computers, television and new media like people and places*. Stanford, CA: CSLI Publications.

Rucinski, D. (1992). Personalized bias in news: The potency of the particular? *Communication Research, 19*(1), 91–108. https://doi.org/10.1177/009365092019001004

Russell, A., & Powers, M. (2020). COVID19 and the media introductory note. *Social Media+ Society*, *6*(3). https://doi.org/10.1177/2056305120947663

Sattelberg, W. (2021, April). "The Demographics of Reddit: Who Uses the Site?" https://www.alphr.com/demographics-reddit/.

Solon, O., & Glaser, A. (2020, April). *How Reddit's coronavirus community became a destination*. NBC News. https://www.nbcnews.com/tech/tech-news/how-reddit-s-coronavirus-community-became-destination-n1162661

Springer, S. (2020). Google Trends provides a tool to monitor population concerns and information needs during COVID-19 pandemic. *Brain, Behavior, and Immunity*. https://doi.org/10.1016/j.bbi.2020.04.073

Statista (2021, January). *Most popular social networks worldwide as of January 2021, ranked by number of active users.* https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

Strekalova, Y. A. (2017). Health risk information engagement and amplification on social media: News about an emerging pandemic on Facebook. *Health Education & Behavior*, *44*(2), 332-339. https://doi.org/10.1177/1090198116660310

Tasnim, S., Hossain, M. M., & Mazumder, H. (2020). Impact of rumors and misinformation on COVID-19 in social media. *Journal of preventive medicine and public health*, *53*(3), 171-174. https://doi.org/10.3961/jpmph.20.094

Tausczik, Y., Faasse, K., Pennebaker, J. W., & Petrie, K. J. (2012). Public anxiety and information seeking following the H1N1 outbreak: blogs, newspaper articles, and Wikipedia visits. Health communication, 27(2), 179-185.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.

Valenzuela, S., Halpern, D., Katz, J. E., & Miranda, J. P. (2019). The paradox of participation versus misinformation: Social media, political engagement, and the spread of misinformation. *Digital Journalism*, *7*(6), 802-823. https://doi.org/10.1080/21670811.2019.1623701

van Dijck, J., & Alinead, D. (2020). Social media and trust in scientific expertise: Debating the COVID-19 pandemic in the Netherlands. *Social Media+ Society*, *6*(4). https://doi.org/10.1177/2056305120981057

Viswanath, K., Finnegan, J., Gollust, S., Glanz, K., & Rimer, B. K. (2015). Communication and health behavior in a changing media environment. *Health Behavior: Theory, Research, and Practice, 5th Edition. San Francisco, CA: Jossey-Bass*, 327-348.

Wong, L. P., & Sam, I. C. (2010). Public sources of information and information needs for pandemic influenza A (H1N1). *Journal of community health*, *35*(6), 676-682. https:// doi.org/10.1007/s10900-010-9271-4

World Health Organization (2020). *Pandemic fatigue – reinvigorating the public to prevent COVID-19. Policy framework for supporting pandemic prevention and management.* Copenhagen: WHO Regional Office for Europe.

Wu, W., Lyu, H., & Luo, J. (2021). Characterizing Discourse about COVID-19 Vaccines: A Reddit Version of the Pandemic Story. *arXiv preprint arXiv:2101.06321*.

To cite this article:

Hale, B.J., Alberta, M., & Chae, S.W. (2022). Reddit as a source of COVID-19 information: A content analysis of r/coronavirus during the early pandemic. *Journal of Communication Technology, 5*(1), 26-57. DOI: 10.51548/joctec-2022-002.