



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Examining thematic and emotional differences across Twitter, Reddit, and YouTube: The case of COVID-19 vaccine side effects

Soyeon Kwon^a, Albert Park^{b,*}

^a Department of Management Information System, College of Business, Dongguk University, 30, Pildong-ro 1gil, Jung-gu, Seoul, 04620, Republic of Korea

^b Department of Software and Information Systems, College of Computing and Informatics, UNC Charlotte, Woodward 310H, 9201 University City Blvd, Charlotte, NC, 28223, USA

ARTICLE INFO

Handling Editor: Marianna Sigala

Keywords:

Social media
Consumer health information
Schema theory
Unsupervised machine learning
Social network analysis

ABSTRACT

Social media discourse has become a key data source for understanding the public's perception of, and sentiments during a public health crisis. However, given the different niches which platforms occupy in terms of information exchange, reliance on a single platform would provide an incomplete picture of public opinions. Based on the schema theory, this study suggests a 'social media platform schema' to indicate users' different expectations based on previous usages of platform and argues that a platform's distinct characteristics foster distinct platform schema and, in turn, distinct nature of information. We analyzed COVID-19 vaccine side effect-related discussions from Twitter, Reddit, and YouTube, each of which represents a different type of the platform, and found thematic and emotional differences across platforms. Thematic analysis using *k*-means clustering algorithm identified seven clusters in each platform. To computationally group and contrast thematic clusters across platforms, we employed modularity analysis using the Louvain algorithm to determine a semantic network structure based on themes. We also observed differences in emotional contexts across platforms. Theoretical and public health implications are then discussed.

1. Introduction

Social media has become an important source of health information with its ability to facilitate the exchange of social support (i.e., informational and emotional) among individuals (Fox & Duggan, 2013; White & Dorman, 2001). By overcoming barriers—geographical, temporal, and social—social media enables individuals to seek evidence-based information provided by health professionals (Metzger & Flanagin, 2011; Song et al., 2016) as well as experience-based information and emotional support from peers sharing similar health concerns (Park, Conway, & Chen, 2018; Song et al., 2016). The public's reliance on social media becomes more prominent during a health crisis as it aids them to cope with uncertainty caused by sudden outbreaks (Gui, Kou, Pine, & Chen, 2017). Different types of platforms can be leveraged to satisfy different information needs. For example, microblogs, such as Twitter, are used to develop collective sense-making of the emerging infectious diseases by sharing up-to-date information (Abd-Alrazaq, Alhuwail, Househ, Hamdi, & Shah, 2020; Lazard, Scheinfeld, Bernhardt, Wilcox, & Suran, 2015) and to combat misinformation (Song, Kwon, Lu, Fan, & Li, 2021). In online forums, such as

Reddit, people can share their true opinions and emotions about controversial issues (Wu, Lyu, & Luo, 2021) and develop collective solutions to conflicting information (Gui et al., 2017; Mamykina, Nakikj, & Elhadad, 2015).

Social media acts as a medium for exchanging information and opinions during a health crisis and has thus become a key data source that provides insights into the public's perceptions and sentiments. Unlike traditional methodologies such as surveys, which are accompanied by several limitations including delayed responses, analyzing social media discourse is proven to be an effective way to gain a comprehensive and timely understanding of public perceptions based on vast amounts of content (Chew & Eysenbach, 2010; Sinnenberg et al., 2017). Accordingly, social media data have been examined for various public health crises (e.g., H1N1, Zika) to understand public concerns (Abd-Alrazaq et al., 2020; Lazard et al., 2015; Signorini, Segre, & Polgreen, 2011), differences in perception across geographical regions (Hou et al., 2021; Hussain et al., 2021) and cultural backgrounds (Kwon & Park, 2022), spatiotemporal patterns to sentiments (Hu et al., 2021), and public opinions regarding certain health issues (Al-Ramahi, Elnoshokaty, El-Gayar, Nasralah, & Wahbeh, 2021; McNeill, Harris, & Briggs,

* Corresponding author.

E-mail addresses: miskwon@dongguk.edu (S. Kwon), apark11@uncc.edu (A. Park).

2016).

Despite the prevalence of research that examines social media discourse during a health crisis, what has been overlooked is the importance of the characteristics of social media platforms, which has been demonstrated in the communication and information science domains (Halpern & Gibbs, 2013; Hiaeshutter-Rice, Chinn, & Chen, 2021; Theocharis et al., 2021, pp. 1–26). Social media platforms with varying technological affordances can facilitate different forms of communication, norms, and cultures (Evans, Pearce, Vitak, & Treem, 2017; Vraga & Bode, 2018). This can cause users to have different expectations for each platform in terms of information exchange, which, in turn, can lead to different information exposures and user experiences in each platform (Hiaeshutter-Rice et al., 2021; Vraga & Bode, 2018). However, existing public health surveillance studies neglect the potential role of platforms in shaping the nature of information and are mostly based on a single platform (e.g., Abd-Alrazaq et al., 2020; Hou et al., 2021; Lazard et al., 2015; Wu et al., 2021), providing a partial and biased understanding of the public's various information needs during a health crisis. Examining public discourse across different types of social media platforms would thus be beneficial for theorizing informational variations across platforms as well as for public health practitioners who wish to have a complete understanding of public information needs during a health crisis.

This study focuses on COVID-19 vaccine side effects, which has received persistent attention on social media since the vaccine rollout (Huangfu, Mo, Zhang, Zeng, & He, 2022; Melton, Olusanya, Ammar, & Shaban-Nejad, 2021). Specifically, it seeks to explore COVID-19 vaccine side effect-related discussions across different types of platforms. Based on schema theory (Rumelhart, 1980), we suggest users have pre-defined expectations regarding platform usage in terms of information exchange, referred to as *social media platform schema* in this study. Within this theoretical framework, this study focuses on addressing whether differences in the nature of information exist across different types of platforms by examining emotional and thematic differences determined by key terminology. Using a typology of social media platforms developed based on its two dimensions (Zhu & Chen, 2015), we selected three platforms, that is, Twitter, Reddit, and YouTube, each of which represents a different type. We conducted thematic analysis to uncover the underlying themes of information discussed in each platform and then employed modularity analysis to group and contrast the nature of information across platforms. Emotional differences were further examined across platforms as well as platforms within the same module to understand the differences in emotional context from thematically similar information.

In the following section, a review of schema theory and the notion of social media platform schema is provided. We then discuss different types of social media platforms and expectations of their usage in terms of information exchange. Based on this, we determine thematic and emotional differences expected across platforms. Next, the procedure of data collection and analysis was provided. Results were then reported regarding salient themes in each platform and thematic and emotional differences across platforms. Finally, we conclude with a discussion of the theoretical and practical implications of our findings.

2. Transition in COVID-19-related discussions on social media

Since the declaration of a global pandemic caused by COVID-19, recent studies have emerged that examine COVID-19 vaccine-related discussions on social media to understand the public reaction towards the new vaccines. General information about the vaccine was shared across social media, such as regarding the progress on its development and information on its accessibility (Lyu, Le Han, & Luli, 2021; Zhang, Wang, Shi, & Wang, 2021). Negative reactions, such as vaccine hesitancy have also been studied as negative information exposure on social media can contribute to the formation of negative attitudes toward the vaccine and the subsequent vaccination decisions (Dunn et al., 2017).

Accordingly, various causes for such hesitancy were found, such as safety concerns (Griffith, Marani, & Monkman, 2021; Hou et al., 2021), pre-existing anti-vaccination beliefs (Jamison et al., 2020), misinformation spread by anti-vaccination advocates (Calac, Haupt, Li, & Mackey, 2022), and other contextual factors including political orientation (Jiang et al., 2021) and mistrust in government or medical professionals (Griffith et al., 2021; Jiang et al., 2021). Since most previous studies analyzed posts created before the start of COVID-19 vaccine administration, public concerns were derived not directly from vaccine-related experiences but from pre-existing beliefs or contextual factors, which is in line with existing literature considering vaccine hesitancy as a multi-faceted notion (Poland & Brunson, 2015).

However, recent studies conducted after vaccine rollout have observed a shift in salient themes. Specifically, concerns about side effects based on direct or indirect vaccination experiences arose as a major theme (Huangfu et al., 2022; Melton et al., 2021). For example, Huangfu et al. (2022) found that concerns raised on Twitter have shifted from clinical trials or vaccine availability to side effects as vaccines become commonly available. Likewise, changes in discussion focus were also observed on Reddit, a primary concern being side effects rather than conspiracy theory-based misinformation (Melton et al., 2021). The change in public attention and online discussions can affect vaccination decisions (Babicki, Malchrzak, & Mastalerz-Migas, 2022). However, little research efforts have been directed toward examining online discourse reflecting the recent changes. Thus, this study focuses on COVID-19 vaccine side effects which are detected constantly as a major barrier to vaccination efforts so as to gain a better understanding of public concerns and develop communication strategies tailored to address those concerns.

3. Theoretical and conceptual background

3.1. Schema theory and social media platform schema

According to the schema theory, people use situation-specific cognitive structures stored in their memories to organize and process new information (Rumelhart, 1980). Individuals develop schemata or cognitive structures of knowledge about an object or a concept through experiences over time (Fiske & Taylor, 1991). When encountering new information, the associated perception is not external stimulus-driven but is created according to the pre-existing knowledge structures. Thus, the subsequent responses are shaped by the extent to which new information conforms to expectations derived from one's knowledge structure (Halkias & Kokkinaki, 2014). The role of schemata, as a guide to make sense of the world, enables individuals to efficiently organize, process, and evaluate a plethora of new information (Marshall, 1995). In the domain of consumer research, various forms of schema-based notions have been developed, such as product category schema (Halkias, 2015), brand schema (Halkias, 2015; Puligadda, Ross, & Grewal, 2012), and ad schema (Halkias & Kokkinaki, 2014), to account for how consumers process market-related stimuli. For example, consumers' attitude toward an ad can be formed according to their expectations or knowledge about how advertising is typically done, that is, ad schema (Halkias, 2015).

In this context, this study suggests a notion of *social media platform schema* to indicate users' pre-defined expectations or knowledge of social media platforms in terms of information exchange. Technological affordances that vary across social media platforms enable different forms of communication, norms, and cultures to be developed in the platform (Evans et al., 2017; Vraga & Bode, 2018). This allows users to develop different cognitive schema pertaining to platform usage, and the type of content that is likely to be created and shared corresponds to their schema-based expectations of the platform. Although not explicitly using the term, several empirical findings have supported the notion of social media platform schema (Reich & Pittman, 2020; Schulze, Schöler, & Skiera, 2014). For example, Schulze et al. (2014) showed that

messages for utilitarian purposes may not be effective in fun-oriented platforms (e.g., Facebook) where users are primed to primarily note hedonic messages. That is, users' platform schema can influence how they perceive information and what they share on the platform (Schulze et al., 2014; Theocharis et al., 2021, pp. 1–26; Vraga & Bode, 2018). Given the importance of social media platforms in shaping user expectations and the nature of shared information, reliance on a singular platform limits our understanding of public responses regarding COVID-19 vaccine side effects. Therefore, this study aims to enrich our understanding by examining online discourse in three social media platforms—Twitter, Reddit, and YouTube.

3.2. Different types of social media platform and information exchange

In this section, we identify the characteristics of platforms, which provides the basis for understanding informational differences across platforms. Zhu and Chen (2015) suggested two defining characteristics to develop a typology of social media platforms. First is the nature of relationships: profile-based versus content-based. Profile-based relationships indicate connections made based on acquaintances or users of interest (e.g., Twitter), while content-based connections are made around shared interests (e.g., Reddit and YouTube). Second is the level of customization of content: broadcasting versus customized content. Broadcasting content is intended for a broad audience (e.g., Twitter and YouTube), while customized content targets a specific audience and is tailored to their particular interests (e.g., Reddit).

Three social media platforms were selected in this study, each of which represents a different typology of the platform. Zhu and Chen's (2015) 2-by-2 typology leads to four categories of social media. However, a platform type which is profile-based and customized-content was excluded in this study. This is because Facebook, the best-known example of profile-based and customized-content, is limited in data accessibility as indicated in their terms of service.

3.2.1. Twitter, Reddit, and YouTube

The characteristics of the platform based on its two dimensions can shape the usage for information exchange. Twitter represents a platform type that is profile-based and broadcasting-content (Zhu & Chen, 2015). Twitter's profile-based connections entail asymmetrical followership wherein users can 'follow' accounts of interest, such as celebrities and organizations. Since asymmetric network structures are created without mutual agreement and, thus, operate on weak-tie networks, content is expected to be shared to a heterogeneous audience (Theocharis et al., 2021, pp. 1–26). In addition, interactive features, such as hashtags, retweets, and mentions, in conjunction with a character limit enable Twitter to serve as a prominent platform for situations that require rapid information exchange (Son, Lee, Jin, & Lee, 2019). These features facilitate the development of users' information sharing-oriented schemas, promoting information sharing-related content for disasters (Son et al., 2019), brands (Taecharungroj, 2017), and political communication (Stieglitz & Dang-Xuan, 2013). Likewise, in the context of a public health crisis, users' expectations of Twitter are to receive and share broadcasting messages from health professionals and government agencies in a timely manner (Neiger, Thackeray, Burton, Thackeray, & Reese, 2013; Vraga & Bode, 2018).

Reddit represents a content-based and customized-content platform (Zhu & Chen, 2015). Reddit has its own unique way to facilitate interactions among users in sub-forums called "subreddits" where users are gathered around topical interests (Anderson, 2015). Although an account is required to participate in the discussion, profiles of users are not focused unlike profile-based platforms. Rather, interactions are structured around the content and through its unique voting system: upvotes (downvotes) given by users to increase (decrease) the visibility of the content (Anderson, 2015). Despite its anonymous environment, the voting system and user contribution known as 'karma' allows participants to build credibility in the network and convey social cues to

others (Prakasam & Huxtable-Thomas, 2021). These features enable Reddit to harness collective intelligence to find solutions to questions, seek advice on difficult situations, or find agreement on debatable topics, and, thus, it is referred to as a 'collaboration' platform (Zhu & Chen, 2015).

YouTube is a prime example of a platform that is content-based and centers on broadcasting-content (Zhu & Chen, 2015). Users typically turn to YouTube primarily for audiovisual content consumption rather than to build relationships (Theocharis et al., 2021, pp. 1–26). While content consumption in profile-based platforms starts from a news feed that curates personalized content based on users' social connections, content on YouTube is available to the general public even without an account. That is, content consumption on YouTube is not guided by the users' friends on the network but by one's topical interests determined by features such as the search box, recommendation systems, and social cues (e.g., number of likes, comments). In addition, users can subscribe to channels of certain content creators or 'influencers', but they don't have to be subscribed back, resulting in asymmetric network structures like Twitter. That is, content is often provided by a YouTube celebrity in a broadcasting form targeting a broad audience sharing similar topical interests (Lange, 2007). In the context of health communication, public health organizations and professionals adopt YouTube to provide educational and prevention information targeting the general public (Bora, Das, Barman, & Borah, 2018). Accordingly, comments by individual users can provide real-time feedback, such as suggesting opinions and seeking additional information to content creators (Zheng, Xue, Sun, & Zhu, 2021).

3.3. Themes and emotions expressed on Twitter, Reddit, and YouTube

This study predicts that the differences in platform characteristics based on the nature of relationships and the level of content customization determine the differences in the nature of shared information, specifically, thematic and emotional differences across platforms.

3.3.1. Thematic differences across social media platforms

The first research question of this study is to examine themes salient in the platform. Since people expect to use Twitter to receive and share messages from credible sources in a timely manner (Shi, Rui, & Whinston, 2014), Twitter's thematic clusters would pertain to sharing public health guidance and up-to-date information regarding COVID-19 vaccines. On the other hand, on Reddit, themes regarding debatable issues around COVID-19 vaccines in addition to sharing anecdotal information are expected. The expectations of Reddit as a medium to share true opinions based on anonymity would enhance users' engagement to discuss their concerns about vaccine side effects (Wu et al., 2021). In particular, COVID-19 vaccine is susceptible to various concerns raised based on its unique characteristics, such as the rapid development and novel technologies (e.g., mRNA) (Dror et al., 2020; Gertz, Sewalk, & Brownstein, 2022; Wong et al., 2021). Thus, COVID-19 vaccine characteristics, in conjunction with the platform usage, would make Reddit a fertile environment for the development of various debates around its side effects. Regarding comments on YouTube, this study predicts discussion themes pertaining to seeking and sharing additional information not addressed in the video content. Khan (2017) examined the motives behind various forms of user engagement on YouTube and found that commenting in particular was related to seeking and giving information as well as social interaction. This is supported by Zheng et al.'s (2021) study that analyzed YouTube comments on Canadian Prime Minister's COVID-19-related briefings. The users used comments to provide feedback to the government and to seek customized information.

Given the various themes shaped by the users' different expectations of the platform usage, our second research question was concerned with further examining the extent to which these themes overlapped across platforms. Yoo, Paek, and Hove (2020) made a distinction between user-oriented and content-oriented platforms and examined their

differential effects on the public perceptions and subsequent behaviors. Compared to the content-oriented platform that affords the users with anonymity, the user-oriented platform—where the user's identity is likely to be visible in the content—would develop different communication norms and, thus, entail different consequences in terms of information exchange (Theocharis et al., 2021, pp. 1–26). In this context, it is plausible that discussion themes on Reddit would differ from those expected on Twitter. Specifically, debatable issues around the COVID-19 vaccine are expected to thrive on Reddit. In contrast, Twitter entails a user profile-oriented and asymmetric relationship network and, thus, is optimized to broadcast messages targeting a broad audience including those with cross-cutting viewpoints (Theocharis et al., 2021, pp. 1–26). Thus, incontrovertible or trustworthy information for heterogeneous audiences, such as factual information, is expected on Twitter (see Munger, 2017). Based on the other dimension of the platform typology, that is, broadcasting versus customized content, YouTube may share overlapping themes with Twitter. However, since the YouTube comments are the focus of this study, it is more plausible to expect themes that are expected on a content-oriented platform. For example, Halpern and Gibbs (2013) found different communication norms and cultures on YouTube compared to a user-oriented platform (e.g., Facebook), suggesting the importance of the affordance of anonymity in shaping the nature of information on the platform. Based on the discussions above, we address the following research questions.

RQ1. What major themes are salient on three social media platforms—Twitter, Reddit, and YouTube concerning COVID-19 vaccine side effects?

RQ2. To what extent do thematic differences exist across social media platforms?

3.3.2. Emotional differences across social media platforms

In addition to the two research questions, we examine whether emotional differences exist across three social media platforms. The overall polarity of sentiment is expected to reflect the prominent themes in the platform. That is, the expected themes on Twitter, such as public health guidance and up-to-date information, would lead to a high prevalence of neutral sentiments. On the other hand, we expect the salience of various negative emotions on Reddit. Reddit enables users, including those with cross-cutting viewpoints, to gather around shared interests and engage in long discussions and debates without character limits (Morales, Monti, & Starnini, 2021; Wu et al., 2021). As the discussion becomes longer, it can lead to comments characterized by greater negative sentiments (Zollo et al., 2015) as one seeks to reinforce his/her perspective (Weeks, 2015). Thus, we expect that various debates developed on Reddit would lead to the salience of various negative emotions. Regarding YouTube comments, this study predicts the prominence of neutral and negative sentiments. For instance, one expected theme regarding information requests by users should manifest in neutral sentiments. In addition, YouTube's relatively anonymous environment created by its content-based network is more likely to engender less polite and civil discussions than the network where the users are easily identifiable (Halpern & Gibbs, 2013). Thus, this study predicts that a more content-focused environment on YouTube could allow users to freely express their concerns about COVID-19 vaccine side effects, which would be detected by their negative sentiments. In sum, the following third research question is proposed.

RQ3. To what extent do emotional differences exist across social media platforms?

4. Methods

4.1. Social media platform and datasets

The data for this study is hosted in the popular social media

platforms, Reddit, Twitter, and Youtube. Reddit facilitates topically focused conversations in an anonymous environment. We collected Reddit posts using pushshift (Baumgartner, Zannettou, Squire, & Blackburn, 2020) from a subreddit called r/CovidVaccinated, which was one of the biggest subreddits for discussing COVID-19 vaccine and vaccination experience and had more than 46,000 members at the time of writing. We initially collected more than 165,000 posts. To select side effect related vaccination posts, we used keywords, 'covid', 'vaccin' and 'side effect' to verify whether a post would contain all the keywords in the same post. We used 'vaccin' to cover a number of different vaccine-related words, such as 'vaccinated'. This resulted in 2722 posts made from December 01, 2020 to August 04, 2021.

Twitter is a free-to-use, micro-blogging platform that can instantly broadcast short messages called tweets up to 280-characters long to the world with little content and frequency restrictions. We collected more than 1 million English tweets by streaming publicly available tweets related to coronavirus using the Python library called Tweepy (Roesslein, 2009) using the keywords from a previous study (Kwon & Park, 2022) from September 2022. Given the inequitable COVID-19 vaccine availability and uptake, we restricted tweets from the US using Twitter's profile information. Then, we filtered in 1567 tweets that had 'covid', 'vaccin', and 'side effect' keywords that were made from September 11, 2020 to June 15, 2021.

YouTube is a video sharing platform with a highly capable search engine. Thus, we used a search query "COVID vaccine side effect" to get 42 videos with comments and downloaded 3758 first level comments using YouTube API (version 3).

4.2. Thematic analysis

We computationally analyzed our datasets, due to the size of the dataset and range of topics, using an unsupervised algorithm called *k*-means clustering, which has been used to analyze online content (Chen, 2012; Park et al., 2018). We considered one post, one tweet, and one comment as a single document for Reddit, Twitter and YouTube respectively. First, we preprocessed the entire dataset to improve the analysis results. Less informative language, specifically URLs and HTML taggers within the text, were removed for our analysis. Stop words, punctuation, high- and low-frequency terms were moved and text was converted to lowercase and then tokenized using Python Natural Language Toolkit (NLTK) package (Bird, 2006). To remove corpus-specific stop words, terms that appear too frequently, we set *max_df* to 0.95. To remove terms that appear too infrequently, we set *min_df* to 10. We then represented our data in their respective vector spaces by generating term frequency matrices using weighted terms of uni-, bi-, and tri-grams. We used Scikit-learn (Pedregosa et al., 2011) to cluster our data using its default parameters for *k*-means clustering and estimated similarity with cosine similarity. After experimenting with varying numbers of clusters, we generated 7 clusters for each platform. Then, the research team collectively reviewed 50 most frequent terms and 50 randomly selected examples of each cluster to label and categorize them together to form our understanding of the discussions. For consistency, emerging themes were discussed among research team members in an open process, in which assumptions, preconceived notions, and different opinions could be challenged and reached to a consensus.

Following the thematic analysis, we used Gephi (Version 0.9.1) (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) with a Fruchterman Reingold graph layout algorithm (Fruchterman & Reingold, 1991) to generate an overview of network visualization of discussed themes on COVID-19 vaccine side effects. Clusters are presented with proportionally sized nodes and the thickness and weight of edges are determined by the Jaccard similarity score between each pair of clusters' 20 most frequently occurring words. Lastly, we applied the Louvain modularity algorithm available in Gephi to identify community structures in a network (Newman & Girvan, 2004) and to illustrate how clusters are semantically similar to one another. After experimenting with varying

Table 1
Topics, keywords, and example threads on Twitter.

Topic label	Topic size	Keywords	Example threads
#VaccinesWork	32.99%	people, getting, second, pfizer, safe, common, cdcgov, possible, effective, learn, dr, vaccinated, mild, vsafe, vaccineswork	I got my second shot of [vaccine brand] vaccine last weekend and experienced no side effects apart from the common minor soreness in my arm. I encourage everyone to get #COVID19 vaccine to save lives! #GetVaccinated #VaccinesWork
Symptoms	19.27%	vaccinated, shot, got, dose, pfizer, arm, second, moderna, people, im, getting, mild, day, sore, today, far, experience, hours	I got my second [vaccine brand] shot yesterday. Feeling no side effects, not even a sore arm. Just like the first shot.
Vaccine safety	18.38%	people, rate, dose, reports, vaers, arm, adults, young, data, cdcgov, hours, analyzed, teens, sore, myocarditis, unreported, report, risk	@[organization] has analyzed the Vaccine Adverse Event Reporting System data on #Covid vaccine in teens and young adults. It is terrible since most side effects go unreported even when they are serious ...
Information about COVID-19 vaccines	12.7%	people, dr, patients, learn, questions, adverse, women, common, cdc, safety, mma, lymph, nodes, unnecessary, cause, cancer, lesserknown, mistaken	@[organization] & @[organization] is hosting a live conversation about #COVID19 vaccine. We are discussing side effects, effectiveness, & site locations.
Vaccine myths	9.76%	arm, pfizer, hours, trial, new, sore, cdc, longterm, possible, long, women, death, concerned, dr, risk, cancer, mma, mammograms	Some women are experiencing irregular menstrual cycles after their #COVID19 vaccine. Is it another legit side effect? @[name] is verifying this in 11pm news tonight.
Information about side effects 1	4.34%	normal, getting, body, learn, signs, building, protection, away, days, ability, flu, daily, activities, fight, immune	Side effects of #COVID19 #vaccine are normal signs your body is building protection. Side effects may even feel like the flu and could affect your ability to do daily activities, but they should go away in a few days.
Information about side effects 2	2.55%	getting, vaccine, prepare, afterwards, effects, common, treat, article, facts, cdc, learn, ucirvine, aarp	Thinking about #Covid19 #vaccine? Learn how to get ready and manage afterwards. What to expect, and how can you treat your side effects? This article has the facts from [organizations]: [URL]

numbers of modules and parameters, we generated 5 modules with the following parameters. The edge weight between clusters were determined using Jaccard similarity score and the resolution was set to 0.9.

4.3. NRCLex

To compare the emotional differences among platforms (and topical clusters within the same module), we employed the NRCLex dictionary, which classifies sentiments according to their affective class (Mohammad & Turney, 2010, 2013). NRCLex has been verified and widely used for a wide range of applications, including in health (Lyu et al., 2021), brand (Mangiò, Pedeliento, & Andreini, 2021), and political communication (Hiaeshutter-Rice et al., 2021). Specifically, we classified sentiments according to their associated overall positive and negative emotion, and then further examined the micro levels of negative emotion, including fear, anger, sadness, and disgust, given that the content for this study is COVID-19 vaccine side effects. The six affective classes for each platform were presented in a spider graph then compared using the pairwise t-test. We performed the same procedure at the module level to understand how topically similar discussions (i.e., topical clusters within the same module) were expressed across platforms.

4.4. User privacy

We only analyzed publicly available data in this paper, which is normally granted exemption from review by Institutional Review Boards. However, we still removed any user identifiable information (e.g. usernames) and slightly modified user quotations in the manuscript to protect user privacy.

5. Results

5.1. Thematic clusters on Twitter, Reddit, and YouTube

RQ1 aims to examine thematic clusters regarding COVID-19 vaccine side effects salient on three platforms.

5.1.1. Twitter

A total of seven clusters were generated using tweets. Most prominent types of tweets were created by individual users to share their vaccination experiences: ‘#VaccinesWork’ (32.99%) and ‘symptoms’ (19.27%). The cluster ‘symptoms’ concerned sharing personal stories of side effects, while ‘#VaccinesWork’ focused on positive anecdotal evidence that can be used as persuasion tactics to influence others’ vaccination decisions (e.g., *I had the [Vaccine Name] #COVID19 #Vaccine and I had no side effects other than being a little tired for about a day! If you had the #CovidVaccine with little to no side effects then #SpeakUp!!! Let people hear about it!*). In addition, tweets created by health professionals and organizations were sharing credible information: ‘Information about COVID-19 vaccines’ (12.7%), ‘Information about side effects 1’ (4.34%) and ‘Information about side effects 2’ (2.55%). ‘Information about COVID-19 vaccines’ shared tweets with links to webinars or external resources (e.g., *Learn more: [URL]*) to encourage people to learn about the COVID-19 vaccines such as common side effects, vaccine efficacy, and mRNA technology. ‘Information about side effects 1 and 2’ were clusters of retweets that included links to external resources, such as health organizations, to share information, including expected side effects and how to manage them. The remaining two clusters were ‘vaccine safety’ (18.38%) and ‘vaccine myths’ (9.76%). ‘Vaccine safety’ included both anti- and pro-vaccination tweets. A prominent example of anti-vaccination tweets concerned vaccine safety based on beliefs that severe cases were often underreported (e.g., *most side effects go unreported even when they are serious[...]*). Tweets supporting vaccine safety were also found in the form of positive personal experiences. ‘Vaccine myths’ included misinformation as well as correction messages aimed at

counteracting prevalent misperceptions and misinformation (e.g., *MYTH: [...]. FACT: [...]. Bust the #COVID19 myths, get #TheFacts*). Table 1 summarizes the overview of seven clusters found on Twitter.

5.1.2. Reddit

Among a total of seven clusters generated using Reddit comments, the largest proportion of posts were related to debates on issues around COVID-19 vaccines: 'Comparing risks of COVID-19 to vaccine side effects' (16.86%), 'debate on vaccine safety' (14.7%), and 'vaccination debate' (13.63%). The cluster 'comparing risks of COVID-19 to vaccine side effects' included posts that justify one's vaccine hesitancy by arguing that vaccination may cause more harm than COVID-19. This was counter-argued by vaccine supporters who argued such misperceptions were based on false premises, such as comparing two incomparable rates (e.g., side effects rate and fatality rate). 'Debate on vaccine safety' showed posts dedicated to safety-related debates caused by the characteristics of COVID-19 vaccine, such as quick development and mRNA technology. In the cluster 'vaccination debate', posts regarding one's decision to get the vaccine were found. For example, those who were against or hesitant about the COVID-19 vaccines often showed their safety concerns and claimed one's right to choose, while proponents highlighted the importance of vaccination and supported vaccine safety based on data reporting extremely rare cases of severe side effects.

The second largest proportion were related to personal anecdotal experiences regarding COVID-19 vaccine side effects: 'Record of vaccine experiences' (17.74%) and 'symptoms' (10.14%). 'Symptoms' focused on personally experienced side effects while 'record of vaccine experiences' illustrated detailed vaccination experiences documented in an hourly or daily manner, ranging from preparation before vaccination to symptoms and treatment after vaccination.

The rest of the two clusters were 'need for information' (14.47%) and 'negative emotions around vaccination' (12.45%). 'Need for information' revealed the need for complete information on vaccine side effects, which were based on the personal belief that many cases could be underreported and much is unknown, including long-term effects. In 'negative emotions around vaccination', many posts contained negative emotions, such as anxiety, worry, and fear, induced by negative firsthand experiences (e.g., *I'm hoping I can get some reassurance that this is a normal effect and I don't have any cause to be worried*) and social media posts (e.g., *some of the stories online and in this sub really scare me*). Table 2 summarizes the overview of clusters found on Reddit.

5.1.3. YouTube

As with previous platforms, we generated seven clusters using YouTube comments. The largest proportion of comments had to do with public mistrust regarding vaccine safety and conspiracy theories: 'Mistrust of vaccine safety 1' (49.02%), 'mistrust of vaccine safety 2' (15.59%), 'mistrust of vaccine safety 3' (7.16%), and 'conspiracy theory' (8.33%). The mistrust of vaccine safety was mainly due to distrust in health professionals, some cases of severe side effects ('mistrust of vaccine safety 1'), a lack of legal liability from pharmaceutical companies ('mistrust of vaccine safety 2'), and a lack of information on long-term side effects ('mistrust of vaccine safety 3'). A set of hypothesized explanations including Plandemic propaganda and DNA conspiracy (e.g., genetic modification) were observed in 'conspiracy theory'. The second largest proportion of comments were about symptoms that users have experienced ('symptoms' (7.18%)) and vaccination experiences including tips to share, such as how to manage side effects ('vaccine experiences' (6.76%)). A small proportion of comments indicated one's hesitancy in accepting COVID-19 vaccines and willingness to wait ('vaccine hesitancy' (5.96%)). Concerns about one's vaccine hesitancy which may dissuade others from getting vaccines were also raised. Table 3 summarizes the overview of clusters found on YouTube comments.

5.2. Semantic overlap among clusters across platforms

To address RQ2 aiming at examining whether differences exist among thematic clusters across platforms, we conducted a modularity analysis. Modularity based on the semantics of a thematic network structure is shown in Fig. 1. Overlapping keywords reported below are aside from query-related keywords (e.g., covid, covid19, coronavirus, effect, side effects).

The module with the largest size (38.1%) included eight clusters from two platforms: Four clusters, 'vaccination debate', 'comparing risks of COVID-19 to vaccine side effects', 'debate on vaccine safety', and 'need for information', came from Reddit and the other four, 'mistrust of vaccine safety 1, 2, 3', and 'conspiracy theory', came from YouTube. These clusters semantically overlapped and together represented concerns about vaccine safety. Keywords shared among clusters from YouTube (e.g., death, die, experimental, heart, companies, severe, risk, trust) indicated concerns about severe side effects, including deaths, which are extremely rare but possible. These keywords were also observed in the clusters from Reddit and semantically linked to debates on various issues around COVID-19 vaccines (i.e., 'comparing risks of COVID-19 to vaccine side effects', 'debate on vaccine safety', and 'vaccination debate'). This module also included 'need for information' in which users called for more information especially on severe side effects.

The module with the second largest size (23.81%) covered five clusters from Twitter: '#VaccinesWork', 'symptoms', 'vaccine safety', 'vaccine myths', and 'information about COVID-19 vaccines'. The overlapping keywords concerned vaccine experiences or symptom-related words (e.g., people, get, shot, common, arm, cdc, pfizer, moderna, second, feel, know, like, dr, receive, sore, hours). Our qualitative analysis revealed that in addition to '#VaccinesWork', 'symptoms', and 'information about COVID-19 vaccines', tweets regarding social media myths about COVID-19 vaccine risks ('vaccine myths') and safety concerns ('vaccine safety') were also associated with symptoms.

The third largest module (19.05%) covered four clusters from two platforms: 'Symptoms' and 'record of vaccine experiences' from Reddit, and 'symptoms' and 'vaccine experiences' from YouTube. Keywords among clusters from Reddit (e.g., sore, symptoms, fever, headache, chills, fatigue) were semantically overlapping with those from YouTube (e.g., second, fever, headache, chills, symptoms), which together illustrated vaccine experiences. During the manual verification of the clusters and module, we discovered that a wide range of symptoms were expressed on Reddit and YouTube from mild to highly negative and severe experiences.

The fourth module (9.52%) included two clusters from two platforms: 'Negative emotions around vaccination' and 'vaccine hesitancy' from Reddit and YouTube, respectively. Keywords shared by these two clusters (e.g., get, risk, think, people, take, wait) indicated the users' vaccination decisions: most remained hesitant. Specifically, 'negative emotions around vaccination' from Reddit involved posts containing negative emotions about COVID-19 vaccine side effects (differentiating keywords: feel, anxiety, worried, scared, fear), which were semantically linked to keywords surrounding vaccine hesitancy (e.g., *I do want to get vaccinated, but I'm honestly pretty fearful of potential side effects. I'd feel a lot more comfortable getting it somewhere down the road when more information is available.*).

The fifth module (9.52%) covered two clusters of tweets sharing information on COVID-19 vaccine side effects: 'Information about side effects 1' and 'Information about side effects 2'. These two were clusters of retweets sharing information provided by health professionals and organizations and, thus, were grouped together apart from the rest of the clusters of Twitter.

Table 2
Topics, keywords, and example threads on Reddit.

Topic label	Topic size	Keywords	Example threads
Record of vaccine experiences	17.74%	shot, arm, day, got, felt, hours, sore, second, days, fever, symptoms, headache, dose, experience, fine, chills, aches, mild, sleep, injection, weeks, anxiety, fatigue, Tylenol, home	Some background notes: I was positive for Covid on [Date]. I was miserable for [number] days. Although I did not require hospitalization, I did end up completing 2 rounds of antibiotics after fighting COVID19 for [illness]. DAY 0: [Date] I received my first dose [vaccine brand] at [time] on through my local health department in a drive thru clinic. I did not feel anything at the time of injection. I felt a little tingle in my arm and neck after approximately [number] minutes but that resolved before I left the parking lot. On my drive home, I felt a little brain fog, but I honestly think it was more emotional than an actual side effect ... I slept terribly that night because my arm hurt pretty bad and I was so warm/feverish. DAY 1: [Date] [vaccine experience description] DAY 2: [Date] [vaccine experience description] DAY 3: [vaccine experience description] DAY 4: [Date] [vaccine experience description]. Overall, I still would do this for a couple days rather than get COVID again.
Comparing risks of COVID-19 to vaccine side effects	16.86%	covid, vaccine, effects, long, term, symptoms, risk, worse, likely, better, chance, far, chances, higher	You should compare the side effects from COVID-19 to the side effects of the vaccine, or the fatality rate of COVID-19 to the fatality rate of the vaccine. The vaccine is far safer on both counts ... Take your pick.
Debate on vaccine safety	14.7%	long, immune, mrna, term, anxiety, doctors, safety, research, study, clinical, fda, approved, trial, efficacy, women, period, trust, quickly, myocarditis, menstrual	To quickly developed covid vaccine in response to a humanitarian crisis was nothing short of a miracle. Anti-vaccine people like to use this as a reason to show their doubt about the vaccine, but they should know that the scientists used a ton of research done in the early 2000s when the SARS-CoV-1 virus came out as their basis for this vaccine. Mistakes of the past and decades of research on mRNA technology came together in what will later be seen as one of the greatest medical breakthroughs of all time.
Need for information	14.47%	effects, long, term, people, mrna, dont, know, risk, data, years, reported, rare, doctor, report, longterm, potential, known	Some people who have been vaccinated and experienced troubling side effects are admittedly reluctant to report them, officially or unofficially, for fear of dissuading others from getting vaccinated. I understand this but to sweep things under the rug for the greater good by under-reporting or selectively reporting side effects, is dishonest. This is why people are suspicious of these vaccines that have been necessarily fast tracked.
Vaccination debate	13.63%	people, dont, know, getting, im, think, want, better, believe, right, choice, decision, worse, risk, issues	There's a difference between encouraging and being pushy. Deciding to help protect everyone by vaccinating is definitely a personal choice. The side effects that range from mildly unpleasant to downright uncomfortable and scary are real. Choose what's right for you, but please know that your choice isn't solely about your own life, but about the lives of the people around you as well.
Negative emotions around vaccination	12.45%	im, getting, ive, got, feel, vaccinated, people, anxiety, long, effect, symptoms, think, worried, scared, experience, post, read	Some online stories and in this sub really scare me. Some people are very gung-ho about the vaccine and but others regret it and really worried about long term effect. I do want to get vaccine shots, but I'm honestly afraid of potential side effects ...
Symptoms	10.14%	dose, effects, second, got, arm, days, hours, sore, symptoms, fever, headache, felt, mild, experience, fatigue, chills, aches, weeks, ago	I'm [age and gender], don't think I had covid and never smoked and I am documenting my experience for people who might be nervous. First dose: about [number] weeks ago at a mass vaccination clinic. It was a quick process. I only had a sore arm and felt sleepy for a couple days. Second dose: [number] hours ago. My arm is much more sore and very tired with mild headache and had a low grade fever about [number] hours after the vaccine. Then I had wild fever dreams and night sweats last night. Overall it feels like a mild flu, took the day off to rest, sucks but not awful. Although this is just an anecdote, I hope it help someone with anxiety!

Table 3
Topics, keywords, and example threads on YouTube.

Topic label	Topic size	Keywords	Example threads
Mistrust of vaccine safety 1	49.02%	people, heart, death, blood, immune, information, rare, doctor, trust, safe, experimental, died, lol, fda, truth, government, children, Johnson, cdc, companies, poison, inflammation, kill	If your physician tells you that the benefit outweighs the risk, ask them "since the risk to my quality of life is so low, are you willing to legally accept all liability for financial damages resulting from adverse reactions to the low-risk immunization recommended by you???" If the vaccines are so safe, why do the manufacturers have the full legal immunity from damages that the vaccine found to induce?
Mistrust of vaccine safety 2	15.59%	pfizer, Johnson, government, die, death, doctor, J&J, experimental, sue, liability, safe, companies, trust	The two questions you should be asking: 1. If I have severe negative effects from this vaccine, which pharmaceutical companies, media outlets, and government agencies can my family sue and who goes to prison for it? 2. If I have to sign a waiver to get the vaccine giving up my right to sue for damages, then how much confidence do you have that it won't cause long-term health problems?
Conspiracy theory	8.33%	people, dont, know, books, died, die, bible, god, death, government, media, older	They're not vaccinating people, they're genetic modifying the workings of your immune system.
Symptoms	7.18%	dose, got, second, arm, days, feel, fever, yesterday, pain, pfizer, headache, sore, symptoms, chills, hours, body, today, sick, aches, fatigue, night	Took mine [time], my side effects include a sore arm and fingers started to hurt for [time]. Then I woke up with severe symptoms, including chills, nausea, [number] fever, a really bad headache, and body aches. Only lasted for about [time] and now I'm back to 90% except a slight headache.
Mistrust of vaccine safety 3	7.16%	side, effects, long, term, rare, dont, know, years, unknown, longterm, shot, death, trial, wont	That's Unbelievable. The long-term effects of a covid vaccine are UNKNOWN!?! Crazy.
Vaccine experiences	6.76%	shot, got, second, first, 2nd, days, like, pfizer, fever, felt, hours, Moderna, flu, arm, sore, flu, normal, headache, chills, fine, symptoms	I was miserable within [number] hours after my shot number 1. Now I feel like having the flu, and cannot work. Side effects include fever, chills, aches, headaches, shakes, shivering, funny smell, dry mouth. People need to prepare for the vaccine: hydration, vitamin c, rest and two days for recovery. Oh, and I am a fit [gender] athlete over [age] with no known conditions. Good Luck All
Vaccine hesitancy	5.96%	im, getting, ive, got, dont, think, going, people, taking, risk, until, wait, future, waiting	Since the lawsuit protection ends until [year], so I'm going to wait until then to see what really is going on.

5.3. Emotional differences

5.3.1. Emotional differences across platforms

RQ3 aims to examine whether and how emotions expressed in posts differ across platforms. Table 4 shows the proportion of emotions salient on each platform and pairwise comparisons across three platforms. For Reddit, fear (11.41%) was the most prominent emotion, followed by sadness (9.99%), anger (6.32%), and disgust (4.31%). A similar pattern was found for other platforms. Fear (8.64%) was the most salient emotion on Twitter, followed by sadness (6.66%), anger (4.65%), and disgust (2.78%). Similarly, YouTube showed fear (8.50%) as the most prominent emotion, followed by sadness (6.75%), anger (5.21%), and disgust (3.10%) (Table 4). Fig. 2 illustrates the proportion of emotions on each platform.

Figs. 3–5 indicate the pairwise comparison across platforms. When comparing Reddit and Twitter, we found a significantly higher proportion of positive sentiments on Twitter (35.00%) over Reddit (22.93%) ($t = -16.87, p < 0.001$) and negative sentiments on Reddit (18.30%) over Twitter (12.52%) ($t = 16.71, p < 0.001$). In addition, a higher proportion of negative emotions was found on Reddit over Twitter for fear ($t = 10.75, p < 0.001$), anger ($t = 8.31, p < 0.001$), sadness ($t = 14.26, p < 0.001$), and disgust ($t = 10.00, p < 0.001$) (Fig. 3). When comparing Reddit and YouTube, we observed a higher proportion of negative sentiments on Reddit (18.30%) over YouTube (14.40%) ($t = 10.20, p < 0.001$), but no significant difference was found for positive sentiments ($t = 1.38, p > 0.05$). A higher proportion of negative emotions was found on Reddit over YouTube for fear ($t = 10.74, p < 0.001$), anger ($t = 5.30, p < 0.001$), sadness ($t = 15.08, p < 0.001$), and disgust ($t = 8.26, p < 0.001$) (Fig. 4). When comparing Twitter and YouTube, we found a higher proportion of positive sentiments on Twitter (35.00%) over YouTube (22.09%) ($t = 14.75, p < 0.001$) and negative sentiments on YouTube (14.40%) over Twitter (12.52%) ($t = -3.65, p < 0.001$). Among negative emotions, a significantly higher proportion of anger was found on YouTube (5.21%) over Twitter (4.65%) ($t = -1.97, p < 0.05$). A marginal difference was found for disgust between YouTube (3.10%) and Twitter (2.78%) ($t = -1.72, p = 0.08$) (Fig. 5).

5.3.2. Emotional differences within the same module across platforms

We further examined emotional differences across platforms within the same module. For the module with the largest size (38.1%), denoted as module1, which covered clusters from Reddit and YouTube regarding concerns about vaccine safety, Reddit showed a significantly higher proportion of fear ($t = 10.06, p < 0.001$), anger ($t = 4.27, p < 0.001$), sadness ($t = 12.41, p < 0.001$), and disgust ($t = 6.90, p < 0.001$) over YouTube. Similarly, in the fourth largest module (i.e., module4) which included 'Negative emotions around vaccination' and 'vaccine hesitancy' from Reddit and YouTube, respectively, Reddit contained a higher proportion of fear ($t = 4.08, p < 0.001$) and sadness ($t = 5.18, p < 0.001$) words over YouTube. However, the third largest module (i.e., module3), where thematic clusters regarding sharing vaccine experiences overlapped, showed an opposite direction. Contrary to the module 1 and 4, in which Reddit showed higher proportion of negative sentiment, YouTube demonstrated a higher proportion in the use of fear ($t = -4.39, p < 0.001$), anger ($t = -6.06, p < 0.001$), and sadness ($t = -4.31, p < 0.001$) words compared to Reddit (Table 5).

6. Discussion and implications

6.1. Discussion

Social media serves as an important data source that provides insight into the public's perceptions during a health crisis (Lazard et al., 2015; Park, Bowling, Shaw, Li, & Chen, 2019; Signorini et al., 2011). However, reliance on a single platform provides only a partial picture of public concerns (Bode & Vraga, 2018; Theocharis et al., 2021, pp. 1–26)—for as we demonstrate, different aspects of public concerns can be captured

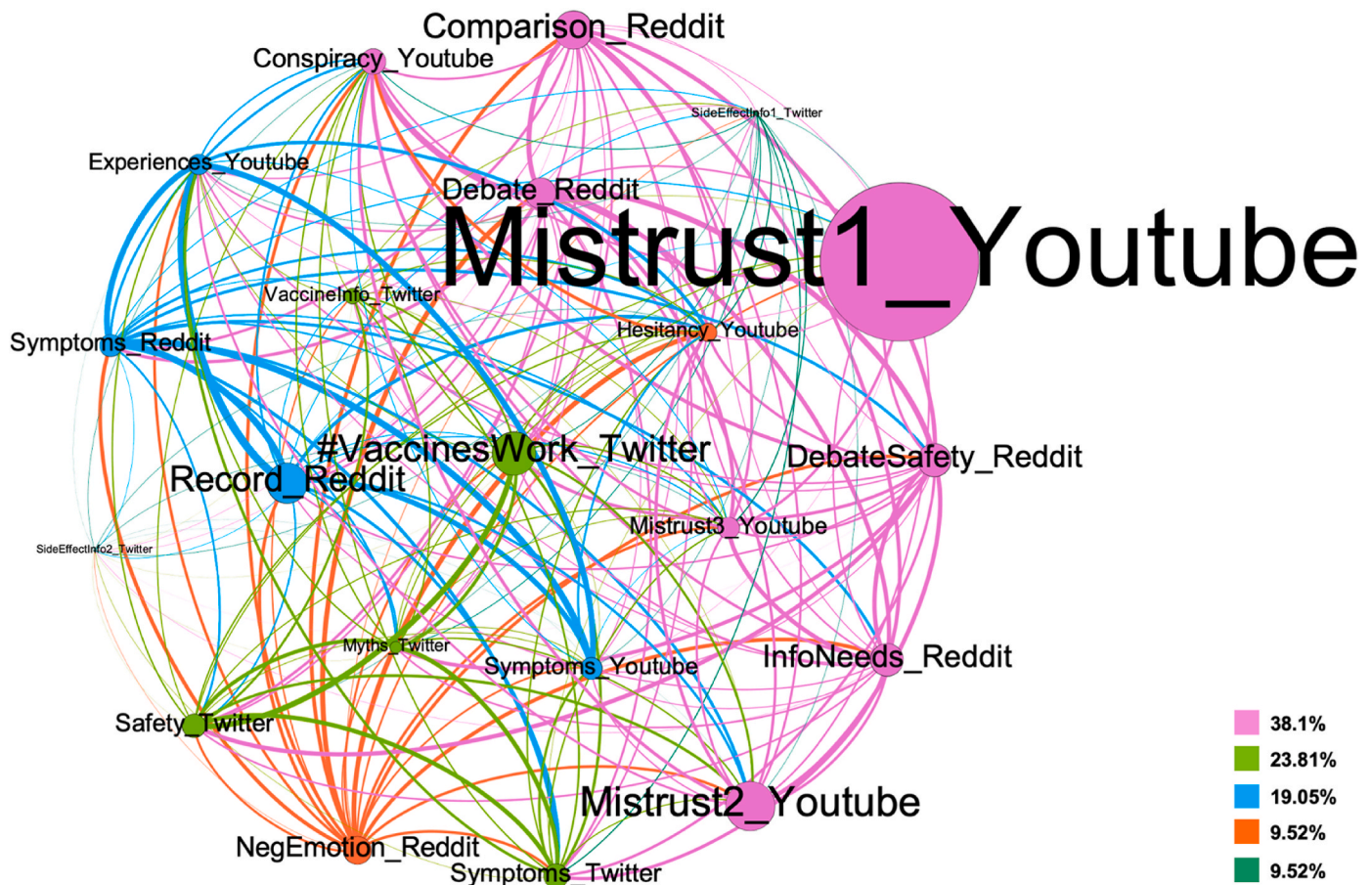


Fig. 1. Thematic network structure concerning Covid vaccine side effect discussions across multiple social media platforms. The topic labels for nodes have been shortened to fit into the visualization. The complete conversion information is available in Appendix A – The Complete Label Conversion Table for Fig. 1.

Table 4
Proportion of emotions and pairwise comparisons across Reddit, Twitter and YouTube.

	Proportion of emotions			Pairwise comparisons		
	Reddit	Twitter	YouTube	Reddit vs. Twitter	Reddit vs. YouTube	Twitter vs. YouTube
Positive	22.93	35.00	22.09	-16.87***	1.38	14.75***
Negative	18.30	12.52	14.40	16.71***	10.20***	-3.65***
Fear	11.41	8.64	8.50	10.75***	10.74***	0.39
Anger	6.32	4.65	5.21	8.31***	5.30***	-1.97*
Sadness	9.99	6.66	6.75	14.26***	15.08***	-0.33
Disgust	4.31	2.78	3.10	10.00***	8.26***	-1.72

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

from different types of platforms. Users’ different platform usage, shaped by the characteristics of the platform (i.e., the nature of the relationships within the network and the degree of content customization), was denoted as “social media platform schema”, and it was suggested as a driver for users to share distinct information. As predicted, variations in the prominent themes and emotional contexts around the same issue (i.e., COVID-19 vaccine side effects) were found across Twitter, Reddit, and YouTube, and the use of different platform schemas explains informational variations. Details of the findings are presented below.

The major finding of this study is the thematic variation across social media platforms. Information sharing-oriented schemas among Twitter users enabled by profile-based and broadcasting-content networks contributed to forming thematic clusters characterized by information-rich content (e.g., “Information about side effects”). Reddit, on the other hand, involved thematic clusters of various debates (e.g., “debate

on vaccine safety”), where the main platform usage can be identified as “collaboration”. That is, users with shared concerns about COVID-19 vaccine side effects collectively engaged to tackle difficult situations (Zhu & Chen, 2015). In response to the YouTube-video content, information on issues that were not addressed or insufficient in the video—in particular, extreme side-effect cases—was further requested by the public through comments. Considering that the video content featured in this study was mostly delivered by news agencies or health professionals, the public may attribute the lack of information on the possible risks of vaccination to the presenters’ intention to conceal such information (Griffith et al., 2021; Murphy et al., 2021), and this attribution may have contributed to forming thematic clusters regarding mistrust of vaccine safety and even conspiracy theories.

Modularity analysis enables us to further explore semantic overlapping among thematic clusters across platforms, which were predicted to be derived from the characteristics of the platform. We found greater

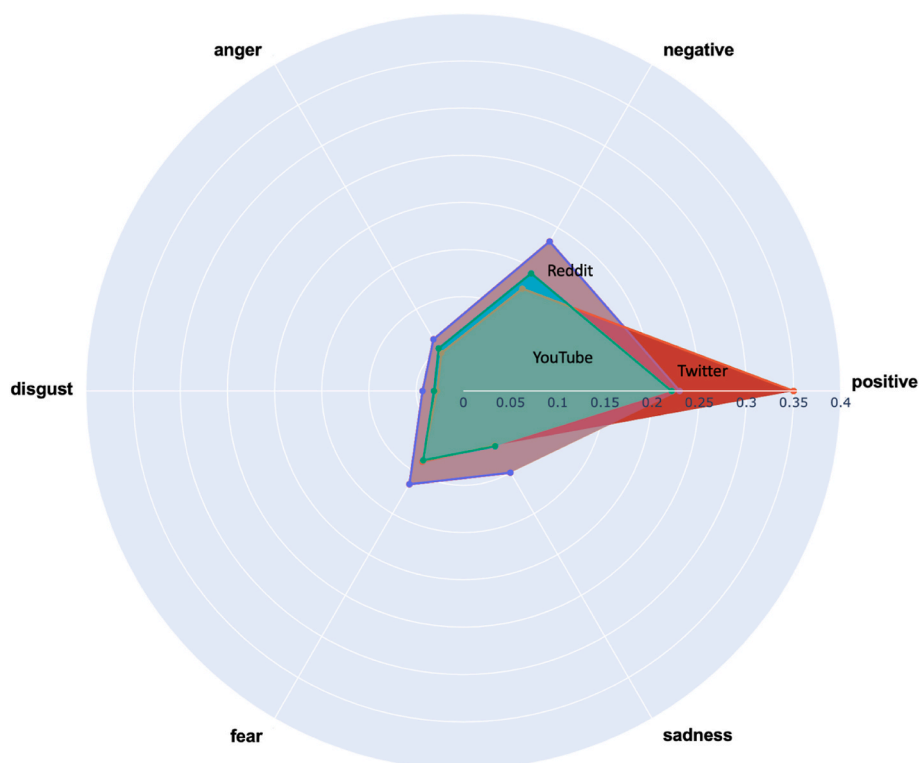


Fig. 2. NRClex emotion infographic concerning Covid vaccine side effect discussions in Reddit, Twitter and YouTube.

semantic overlapping among clusters from Reddit and YouTube than those from Twitter, indicating a distinction between content-based and profile-based networks (Yoo et al., 2020). For example, discussions on vaccine safety were salient on all three platforms, but information sharing-oriented schemas in Twitter facilitated the development of clusters regarding the sharing of misperceptions and corresponding messages intended to correct these misperceptions. That is, the type of information deemed to be safe to broadcast to heterogeneous groups is likely to be shared on profile-based networks (Theocharis et al., 2021, pp. 1–26). By contrast, the other two platforms—both of which provide anonymity affordances—allowed discussions that are likely to be reluctant on profile-based networks, such as severe cases of vaccine side effects (e.g., *some people who have been vaccinated and are experienced troubling side effects are admittedly reluctant to report them officially or unofficially for fear of dissuading others from getting vaccinated*). In addition, as indicated by module 4, anonymity affordances of content-based networks enabled users to express their true opinions and emotions (e.g., *I do want to get vaccine shots, but I'm honestly afraid of potential side effects ...*), which is less likely to occur on profile-based networks (Brown, Ng, Riedl, & Lacasa-Mas, 2018; Halpern & Gibbs, 2013).

The distinction between content-based and profile-based networks was further corroborated by emotional analysis. Compared to the other two content-based platforms (i.e., Reddit and YouTube), Twitter showed a higher proportion of positive sentiments and lower proportion of negative emotions (e.g., fear, anger, sadness, disgust). As discussed earlier, this can be derived from profile-based connections, which allow users to be exposed to those with opposing viewpoints and, thus, cause them to consider Twitter not as a safe place, such as Reddit or YouTube, to voice their true opinions about vaccine side effects (Munger, 2017; Theocharis et al., 2021, pp. 1–26). In that sense, the tendency of positivity was observed in discussions on Twitter regarding vaccine experiences; symptoms and information conveyed in tweets were mild, and some even used positive anecdotal evidence as a persuasion tactic to encourage others to get vaccinated (e.g., “#VaccinesWork”) (see Kim, 2015).

On the other hand, content-based networks, where connections are formed around users' shared interests and, accordingly, user profiles may matter less than the content, enable users to express their true opinions and emotions (Brown et al., 2018; Halpern & Gibbs, 2013), thereby contributing a higher proportion of negative emotions, as seen on Reddit and YouTube. In particular, Reddit was noted for its high proportion of all four negative emotions. As a collaboration platform, Reddit encourages users—including those with divergent viewpoints—to engage collectively in debating various issues. As debates lengthen on Reddit, the anonymity affordance that enables users to express their true opinions and emotions as well as one's tendency to reinforce their counter-arguments (Weeks, 2015) could make the comments with greater negative sentiments (Zollo et al., 2015). Noteworthy is how YouTube showed a significantly higher proportion of anger than Twitter. This may have resulted from users' mistrust of mainstream news agencies or health professionals, who were the presenters of most of the video content considered in this study. Users who made comments may misattribute the lack of information on extremely severe cases or long-term side effects of vaccination to the presenters' intention to hide the actual risks of vaccination (Griffith et al., 2021; Murphy et al., 2021), which could manifest in anger.

Another interesting finding is that, while Reddit and YouTube showed similar patterns for emotional contexts, a notable difference was found within the same module of sharing vaccine experiences. Unlike other modules in which Reddit showed higher proportions of emotions, sharing vaccine experiences yielded higher proportions of emotions (i.e., fear, anger, sadness) on YouTube. This could have resulted from the different communication norms developed on the platform. We found that Reddit users tended to provide empirical evidence supported by reliable and objective information, without the clear presence of subjective emotions (e.g., “records of vaccine experiences”), which could be indicative of the communication norm to share experience-based information to benefit others. However, this particular communication norm may not exist on YouTube, where emotional expression could play a greater part in the sharing of vaccination experiences (see Shao, 2009).

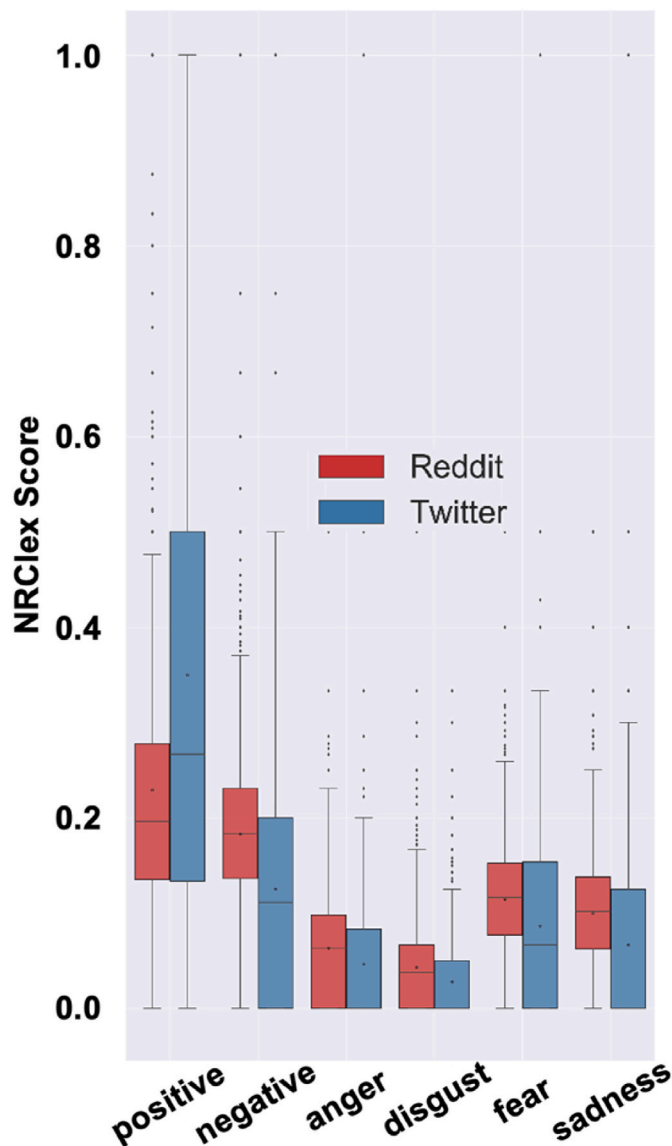


Fig. 3. Emotional affect comparison in Reddit and Twitter.

6.2. Theoretical implications

First, this study highlights the importance of a cross-platform perspective when examining public perceptions based on social media discourse. A growing body of health surveillance studies have emerged that examine the public's perceptions and sentiments based on social media discourse as individuals' health decisions become increasingly affected by their exposure to health information available on social media (Dunn, Leask, Zhou, Mandl, & Coiera, 2015; Wilson & Wiysonge, 2020). However, most previous studies have leveraged a single platform as a data source (e.g., Abd-Alrazaq et al., 2020; Hou et al., 2021; Lazard et al., 2015), and potentially overlooked the impact of the platform in shaping the nature of the information shared, which is suggested in the communication/information science domains (Hiaeshutter-Rice et al., 2021; Theocharis et al., 2021, pp. 1–26; Vraga & Bode, 2018). This study contributes to the existing literature by bridging two research silos—public health surveillance and the communication/information science domains. In so doing, we expand the understanding of public concerns during the public health crisis by providing insights into different information needs and exposures that exist across different types of platforms.

Second, this study suggests a theoretical underpinning for

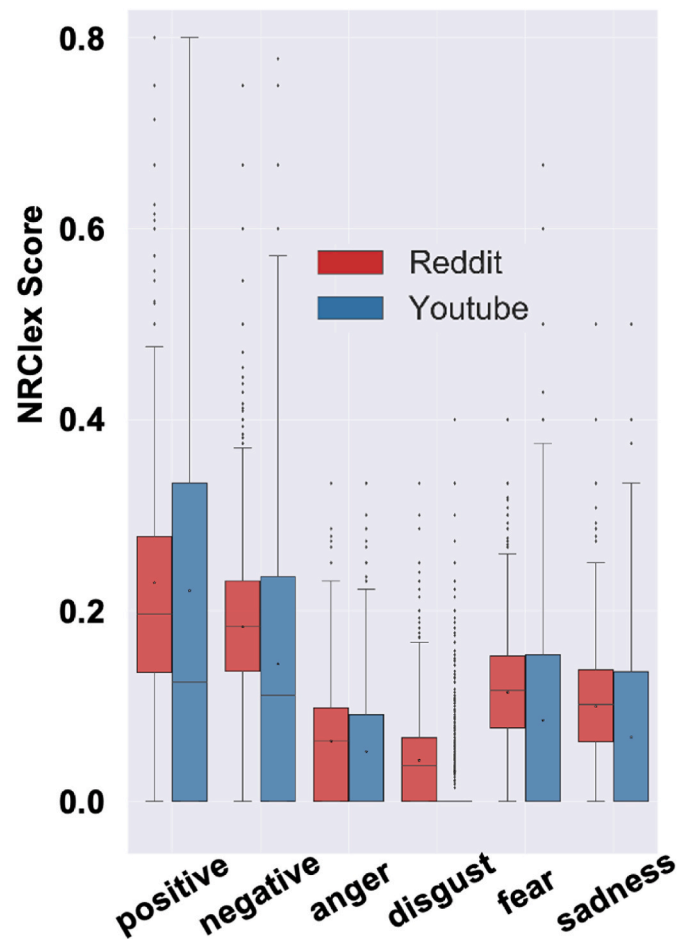


Fig. 4. Emotional affect comparison in Reddit and YouTube.

understanding the nature of information that varies by different types of platforms. Recently, a growing but limited body of research has emerged that employs a cross-platform perspective with the recognition of the public's multi-platform usage for obtaining health information (Lim, Molenaar, Brennan, Reid, & McCaffrey, 2022; O'Leary, Coulson, Perez-Vallejos, & McAuley, 2020; Wawrzuta, Klejdysz, Jaworski, Gotlib, & Panczyk, 2022). However, little research attempts have been made to understand the characteristics of different platform types and theorize how they subsequently lead to systematic variations in information. Drawing on the schema theory (Fiske & Taylor, 1991; Rumelhart, 1980) as an overarching theoretical framework, we suggest two dimensions of the platform based on its typology (Zhu & Chen, 2015) as a new approach to identify platform characteristics and provide a theoretical underpinning of how the platform's relationship and content characteristics determine thematic and emotional differences across platforms.

Third, drawing on the schema theory, this study offers a novel notion of 'social media platform schema'. Building on the view that the context-specific cognitive structures affect information processing (i.e., schema theory) (Rumelhart, 1980), a body of literature has developed to explore how people process information regarding various stimulus domains. For example, consumers develop cognitive schemata for various market-related elements, such as product category schema, brand schema, and ad schema (Halkias, 2015). We expand the scope of the schema-based approach by applying it to the social media platform context, concluding with insights into how users' platform schema shapes thematic and emotional differences across platforms. To the best of authors' knowledge, this study is the first to suggest the notion of social media platform schema and apply it to examine health information in social media.

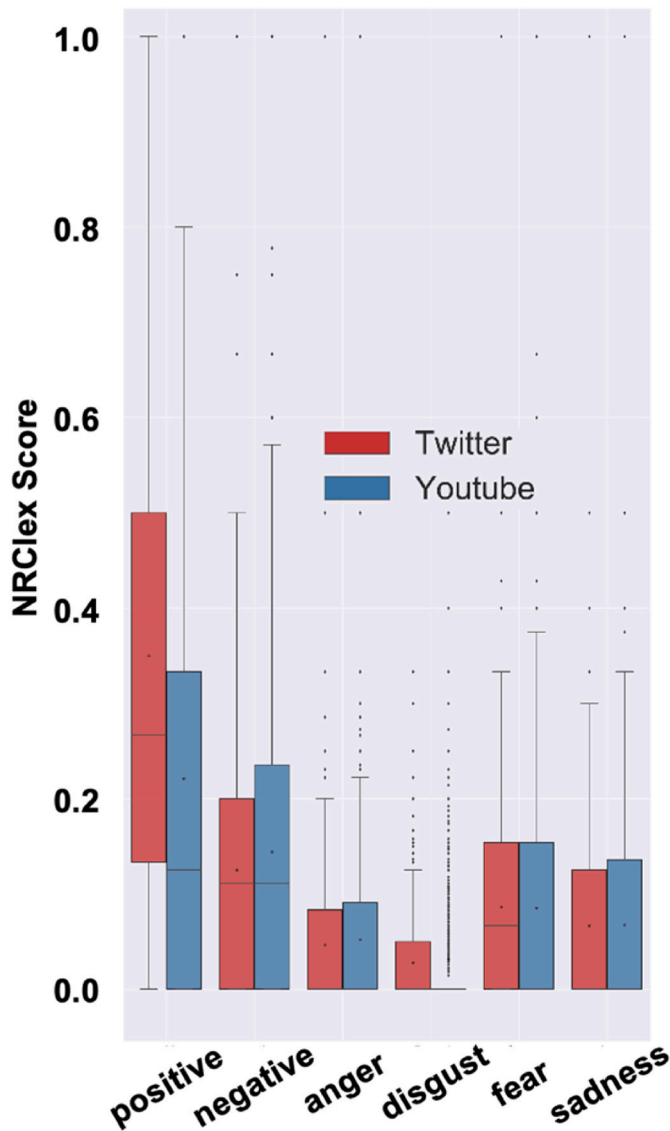


Fig. 5. Emotional affect comparison in Twitter and YouTube.

6.3. Public health implications

This research also provides important public health implications. First, this study challenges public health practitioners to recognize that platform-specific-biased conclusions can be drawn if only a single social media platform is used as a data source. For example, practitioners who rely on Twitter only may mistakenly conclude that the public’s vaccine experiences are positive based on positive sentiments detected along with positive anecdotal evidence (i.e., “#VaccineWork”). However, practitioners can expand their understanding of public perceptions by

including other platforms from the typology of social media platforms. As demonstrated in our study, other platforms can highlight different aspects of public concerns, such as various debatable issues around COVID-19 vaccine safety and extremely rare but severe cases of side effects that can act as a major source of public concern. Thus, this research offers insights into the importance of cross-platform examination based on the different aspects of public concerns captured in different types of platforms.

Second, public health practitioners should also plan strategies across platforms. For example, Reddit can be used to detect the debatable issues that spur public concerns. Among those issues, those that are grounded in misinformation or conspiracy theories can be disentangled by using Twitter, which is an efficient medium to inform the public about the prevalent misconceptions and corresponding corrected information. By doing so, practitioners can effectively allocate adequate resources toward communication strategies that pinpoint and address information needs posed by the public during a public health crisis.

Third, we recommend that practitioners develop communication strategies counteracting certain negative emotions prevalent in public discourse. The results of this study demonstrated that fear was the most prominent emotion about COVID-19 vaccine side effects, unlike other health crises where anger is often leveraged to sow mistrust in medical professionals and increase the public’s susceptibility to conspiracy theories (Mitra, Counts, & Pennebaker, 2016). Thus, communication strategies for the general public aimed at combating misinformation about COVID-19 vaccines should tailor towards minimizing their fear (Chou & Budenz, 2020). Unlike anger, which generates approach tendency, fear can result from uncertain and uncontrollable situations and is related to an avoidance tendency which may manifest as vaccine hesitancy (Carver & Harmon-Jones, 2009; Weeks, 2015). Thus, effective communication strategies, including concrete and actionable messages to reduce fear, can benefit the general public by increasing their self-efficacy (Maloney, Lapinski, & Witte, 2011; Witte & Allen, 2000).

7. Limitations and future research

This study highlights the importance of a cross-platform perspective, explains the differences across platforms, and offers a novel ‘social media platform schema’ as a theoretical framework for future research. There are several ways to strengthen the findings of our study.

First, the current study is based on the typology of social media platforms developed by Zhu and Chen (2015). However, we excluded a profile-based and customized-content based platform, because its best-known example, Facebook, restricts data crawling in their terms of service. If another widely used profile-based and customized-content based platform emerges, future research should examine the new platform through the theoretical framework of this study. Similarly, given that our findings could be due to the unique characteristics of that platform, cross-examination using a secondary platform within the same typology can be further performed.

Second, while this study examined information characteristics at the platform level, it is worth investigating whether our theoretical framework holds at the individual level. For example, future research might

Table 5 Proportion of emotions and pairwise comparisons between Reddit and YouTube within the same module.

	Module1			Module3			Module4		
	Reddit	YouTube	Reddit vs YouTube	Reddit	YouTube	Reddit vs YouTube	Reddit	YouTube	Reddit vs YouTube
Positive	26.62	24.39	2.72**	16.04	9.74	9.46***	20.66	19.97	0.46
Negative	17.09	12.71	8.89***	20.63	22.87	-3.46***	18.93	17.19	1.55
Fear	10.91	7.42	10.06***	12.34	14.56	-4.39***	11.73	8.91	4.08***
Anger	5.33	4.22	4.27***	8.38	10.51	-6.06***	6.40	6.23	0.24
Sadness	8.72	5.45	12.41***	12.27	13.92	-4.31***	10.94	7.50	5.18***
Disgust	4.02	2.77	6.90***	4.80	4.37	1.44	4.61	4.65	-0.07

Note. *p < 0.05, **p < 0.01, ***p < 0.001.

examine individuals participating in multiple platforms, and characterize their information sharing behavior with respect to different typology of social media platforms. Our findings would be bolstered if individuals followed the communication norms of the platform.

Third, the data may be subjective to self-selection bias, since not everyone uses social media. Similarly, each social media platform has different user demographics, which could also contribute to their information needs and sharing behavior. Thus, replicating a similar study with different social media platforms that consist of different user demographics can bolster our understanding of social media platform schema during a health crisis.

Appendix A

The Complete Label Conversion Table for Fig. 1.

	Full Labels	Shorten Labels (Fig. 1)
Twitter	#VaccinesWork	#VaccinesWork_Twitter
	Symptoms	Symptoms_Twitter
	Vaccine safety	Safety_Twitter
	Information about COVID-19 vaccines	VaccineInfo_Twitter
	Vaccine myths	Myths_Twitter
	Information about side effects 1	SideEffectInfo1_Twitter
	Information about side effects 2	SideEffectInfo2_Twitter
Reddit	Record of vaccine experiences	Record_Reddit
	Comparing risks of COVID-19 to vaccine side effects	Comparison_Reddit
	Debate on vaccine safety	DebateSafety_Reddit
	Need for information	InfoNeeds_Reddit
	Vaccination debate	Debate_Reddit
	Negative emotions around vaccination	NegEmotion_Reddit
	Symptoms	Symptoms_Reddit
YouTube	Mistrust of vaccine safety 1	Mistrust1_YouTube
	Mistrust of vaccine safety 2	Mistrust2_YouTube
	Conspiracy theory	Conspiracy_YouTube
	Symptoms	Symptoms_YouTube
	Mistrust of vaccine safety 3	Mistrust3_YouTube
	Vaccine experiences	Experiences_YouTube
	Vaccine hesitancy	Hesitancy_YouTube

References

Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Z. (2020). Top concerns of tweeters during the COVID-19 pandemic: Infoveillance study. *Journal of Medical Internet Research*, 22(4), Article e19016.

Al-Ramahi, M., Elnohokaty, A., El-Gayar, O., Nasralah, T., & Wahbeh, A. (2021). Public discourse against masks in the COVID-19 era: Infodemiology study of Twitter data. *JMIR Public Health and Surveillance*, 7(4), Article e26780.

Anderson, K. E. (2015). Ask me anything: What is Reddit? *Library Hi Tech News*, 5, 8–11.

Babicki, M., Malchrzak, W., & Mastalerz-Migas, A. (2022). Assessment of attitudes, main concerns and sources of knowledge regarding COVID-19 vaccination in Poland in the unvaccinated individuals—a nationwide survey. *Vaccines*, 10(3), 381.

Baumgartner, J., Zannettou, S., Squire, M., & Blackburn, J. (2020). The pushshift telegram dataset. May. In *Proceedings of the international AAAI Conference on Web and social media* (pp. 840–847).

Bird, S. (2006). Nltk: The natural language toolkit. July. In *Proceedings of the COLING/ACL on interactive presentation sessions* (pp. 69–72). Association for Computational Linguistics.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), Article P10008.

Bode, L., & Vraga, E. K. (2018). Studying politics across media. *Political Communication*, 35(1), 1–7.

Bora, K., Das, D., Barman, B., & Borah, P. (2018). Are internet videos useful sources of information during global public health emergencies? A case study of YouTube videos during the 2015–16 Zika virus pandemic. *Pathogens and Global Health*, 112(6), 320–328.

Brown, D. K., Ng, Y. M. M., Riedl, M. J., & Lacasa-Mas, I. (2018). Reddit’s veil of anonymity: Predictors of engagement and participation in media environments with hostile reputations. *Social Media+ Society*, 4(4), 1–9.

Credit author statement

Soyeon Kwon: Conceptualization, Validation, Formal analysis, Writing - Original Draft, Project administration; **Albert Park:** Conceptualization, Software, Formal analysis, Data Curation, Writing - Review & Editing.

Data availability

The authors do not have permission to share data.

Calac, A. J., Haupt, M. R., Li, Z., & Mackey, T. (2022). Spread of COVID-19 vaccine misinformation in the ninth inning: Retrospective observational infodemic study. *JMIR Infodemiology*, 2(1), Article e33587.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, 135(2), 183–204.

Chen, A. T. (2012). Exploring online support spaces: Using cluster analysis to examine breast cancer, diabetes and fibromyalgia support groups. *Patient Education and Counseling*, 87(2), 250–257.

Chew, C., & Eysenbach, G. (2010). Pandemics in the age of twitter: Content analysis of tweets during the 2009 H1N1 outbreak. *PLoS One*, 5(11), Article e14118.

Chou, W. Y. S., & Budenz, A. (2020). Considering emotion in COVID-19 vaccine communication: Addressing vaccine hesitancy and fostering vaccine confidence. *Health Communication*, 35(14), 1718–1722.

Dror, A. A., Eisenbach, N., Taiber, S., Morozov, N. G., Mizrachi, M., Zigran, A., ... Sela, E. (2020). Vaccine hesitancy: The next challenge in the fight against COVID-19. *European Journal of Epidemiology*, 35(8), 775–779.

Dunn, A. G., Leask, J., Zhou, X., Mandl, K. D., & Coiera, E. (2015). Associations between exposure to and expression of negative opinions about human papillomavirus vaccines on social media: An observational study. *Journal of Medical Internet Research*, 17(6), e4343.

Dunn, A. G., Surian, D., Leask, J., Dey, A., Mandl, K. D., & Coiera, E. (2017). Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. *Vaccine*, 35(23), 3033–3040.

Evans, S. K., Pearce, K. E., Vitak, J., & Treem, J. W. (2017). Explicating affordances: A conceptual framework for understanding affordances in communication research. *Journal of Computer-Mediated Communication*, 22(1), 35–52.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill.

Fox, S., & Duggan, M. (2013). *Health online 2013. Pew internet & American life Project*. Retrieved from Accessed 2 May, 2022 <https://www.pewresearch.org/internet/2013/01/15/health-online-2013/>.

Fruchterman, T. M., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11), 1129–1164.

- Gertz, A., Sewalk, K., & Brownstein, J. (2022). Assessing the shift in reasoning for COVID-19 vaccine hesitancy in the United States using a six-month cross-sectional analysis from December 2020 to June 2021. *International Journal of Infectious Diseases*, 116, S59.
- Griffith, J., Marani, H., & Monkman, H. (2021). COVID-19 vaccine hesitancy in Canada: Content analysis of tweets using the theoretical domains framework. *Journal of Medical Internet Research*, 23(4), Article e26874.
- Gui, X., Kou, Y., Pine, K. H., & Chen, Y. (2017). Managing uncertainty: Using social media for risk assessment during a public health crisis. May. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 4520–4533). New York, NY: ACM.
- Halkias, G. (2015). Mental representation of brands: A schema-based approach to consumers' organization of market knowledge. *The Journal of Product and Brand Management*, 24(5), 438–448.
- Halkias, G., & Kokkinaki, F. (2014). The degree of ad-brand incongruity and the distinction between schema-driven and stimulus-driven attitudes. *Journal of Advertising*, 43(4), 397–409.
- Halpern, D., & Gibbs, J. (2013). Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior*, 29(3), 1159–1168.
- Hiaeshutter-Rice, D., Chinn, S., & Chen, K. (2021). Platform effects on alternative influencer content: Understanding how audiences and channels shape misinformation online. *Frontiers in Political Science*, 3, 1–13.
- Hou, Z., Tong, Y., Du, F., Lu, L., Zhao, S., Yu, K., ... Lin, L. (2021). Assessing COVID-19 vaccine hesitancy, confidence, and public engagement: A global social listening study. *Journal of Medical Internet Research*, 23(6), Article e27632.
- Huangfu, L., Mo, Y., Zhang, P., Zeng, D. D., & He, S. (2022). COVID-19 vaccine Tweets after vaccine rollout: Sentiment-based topic modeling. *Journal of Medical Internet Research*, 24(2), Article e31726.
- Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., Gogate, M., Dashtipour, K., ... Sheikh, A. (2021). Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward covid-19 vaccines in the United Kingdom and the United States: Observational study. *Journal of Medical Internet Research*, 23(4), Article e26627.
- Hu, T., Wang, S., Luo, W., Zhang, M., Huang, X., Yan, Y., ... Li, Z. (2021). Revealing public opinion towards COVID-19 vaccines with Twitter data in the United States: Spatiotemporal perspective. *Journal of Medical Internet Research*, 23(9), Article e30854.
- Jamison, A. M., Broniatowski, D. A., Dredze, M., Sangraula, A., Smith, M. C., & Quinn, S. C. (2020). Not just conspiracy theories: Vaccine opponents and proponents add to the COVID-19 'infodemic' on Twitter. *Harvard Kennedy School Misinformation Review*, 1, 1–22.
- Jiang, X., Su, M. H., Hwang, J., Lian, R., Brauer, M., Kim, S., et al. (2021). Polarization over vaccination: Ideological differences in Twitter expression about COVID-19 vaccine favorability and specific hesitancy concerns. *Social Media+ Society*, 7(3), 1–14.
- Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior*, 66, 236–247.
- Kim, H. S. (2015). Attracting views and going viral: How message features and news-sharing channels affect health news diffusion. *Journal of Communication*, 65(3), 512–534.
- Kwon, S., & Park, A. (2022). Understanding user responses to the COVID-19 pandemic on Twitter from a terror management theory perspective: Cultural differences among the US, UK and India. *Computers in Human Behavior*, 128, Article 107087.
- Lange, P. G. (2007). Commenting on comments: Investigating responses to antagonism on YouTube. March. In *Society for applied anthropology conference* (pp. 163–190) (Tampa, Florida).
- Lazard, A. J., Scheinfeld, E., Bernhardt, J. M., Wilcox, G. B., & Suran, M. (2015). Detecting themes of public concern: A text mining analysis of the centers for disease control and prevention's Ebola live twitter chat. *American Journal of Infection Control*, 43(10), 1109–1111.
- Lim, M. S., Molenaar, A., Brennan, L., Reid, M., & McCaffrey, T. (2022). Young adults' use of different social media platforms for health information: Insights from web-based conversations. *Journal of Medical Internet Research*, 24(1), Article e23656.
- Lyu, J. C., Le Han, E., & Luli, G. K. (2021). COVID-19 vaccine-related discussion on Twitter: Topic modeling and sentiment analysis. *Journal of Medical Internet Research*, 23(6), Article e24435.
- Maloney, E. K., Lapinski, M. K., & Witte, K. (2011). Fear appeals and persuasion: A review and update of the extended parallel process model. *Social and Personality Psychology Compass*, 5(4), 206–219.
- Mamykina, L., Nakik, D., & Elhadad, N. (2015). Collective sensemaking in online health forums. April. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 3217–3226).
- Mangiò, F., Pedeliento, G., & Andreini, D. (2021). Branding rhetoric in times of a global pandemic: A text-mining analysis. *Journal of Advertising*, 50(3), 240–252.
- Marshall, S. P. (1995). *Schemas in problem solving*. New York: Cambridge University Press.
- McNeill, A., Harris, P. R., & Briggs, P. (2016). Twitter influence on UK vaccination and antiviral uptake during the 2009 H1N1 pandemic. *Frontiers in Public Health*, 4, 26.
- Melton, C. A., Olusanya, O. A., Ammar, N., & Shaban-Nejad, A. (2021). Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. *Journal of Infection and Public Health*, 14(10), 1505–1512.
- Metzger, M. J., & Flanagin, A. J. (2011). Using Web 2.0 technologies to enhance evidence-based medical information. *Journal of Health Communication*, 16(1), 45–58.
- Mitra, T., Counts, S., & Pennebaker, J. W. (2016). Understanding anti-vaccination attitudes in social media. March. In *Proceedings of the 10th international conference on web and social media* (pp. 269–278) (Cologne: Germany).
- Mohammad, S., & Turney, P. (2010). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. June. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text* (pp. 26–34).
- Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), 436–465.
- Morales, G. D. F., Monti, C., & Starnini, M. (2021). No echo in the chambers of political interactions on Reddit. *Scientific Reports*, 11(1), 1–12.
- Munger, K. (2017). Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Political Behavior*, 39(3), 629–649.
- Murphy, J., Vallières, F., Bentall, R. P., Shevlin, M., McBride, O., Hartman, T. K., ... Hyland, P. (2021). Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. *Nature Communications*, 12(1), 1–15.
- Neiger, B. L., Thackeray, R., Burton, S. H., Thackeray, C. R., & Reese, J. H. (2013). Use of Twitter among local health departments: An analysis of information sharing, engagement, and action. *Journal of Medical Internet Research*, 15(8), e177.
- Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2), Article 026113.
- O'Leary, K., Coulson, N., Perez-Vallejos, E., & McAuley, D. (2020). Towards understanding how individuals with inflammatory bowel disease use contemporary social media platforms for health-related discourse. *Computers in Human Behavior*, 112, Article 106463.
- Park, A., Bowling, J., Shaw, G., Li, C., & Chen, S. (2019). Adopting social media for improving health. *North Carolina Medical Journal*, 80(4), 240–243.
- Park, A., Conway, M., & Chen, A. T. (2018). Examining thematic similarity, difference, and membership in three online mental health communities from Reddit: A text mining and visualization approach. *Computers in Human Behavior*, 78, 98–112.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Poland, C. M., & Brunson, E. K. (2015). The need for a multi-disciplinary perspective on vaccine hesitancy and acceptance. *Vaccine*, 33(2), 277–279.
- Prakasam, N., & Huxtable-Thomas, L. (2021). Reddit: Affordances as an enabler for shifting loyalties. *Information Systems Frontiers*, 23(3), 723–751.
- Puligadda, S., Ross, W. T., Jr., & Grewal, R. (2012). Individual differences in brand schematicity. *Journal of Marketing Research*, 49(1), 115–130.
- Reich, B. J., & Pittman, M. (2020). An appeal to intimacy: Consumer response to platform-appeal fit on social media. *Journal of Consumer Psychology*, 30(4), 660–670.
- Roeslein, J. (2009). *Tweeepy*. Retrieved from <http://www.tweeepy.org>.
- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Schulze, C., Schöler, L., & Skiera, B. (2014). Not all fun and games: Viral marketing for utilitarian products. *Journal of Marketing*, 78(1), 1–19.
- Shao, G. (2009). Understanding the appeal of user-generated media: A uses and gratification perspective. *Internet Research*, 19(1), 7–25.
- Shi, Z., Rui, H., & Whinston, A. B. (2014). Content sharing in a social broadcasting environment: Evidence from twitter. *MIS Quarterly*, 38(1), 123–142.
- Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PLoS One*, 6(5), Article e19467.
- Sinnenberg, L., Buttenheim, A. M., Padrez, K., Mancheno, C., Ungar, L., & Merchant, R. M. (2017). Twitter as a tool for health research: A systematic review. *American Journal of Public Health*, 107(1), e1–e8.
- Song, Y., Kwon, K. H., Lu, Y., Fan, Y., & Li, B. (2021). The "parallel pandemic" in the context of China: The spread of rumors and rumor-corrections during COVID-19 in Chinese social media. *American Behavioral Scientist*, 65(14), 2014–2036.
- Song, H., Omori, K., Kim, J., Tenzek, K. E., Hawkins, J. M., Lin, W. Y., ... Jung, J. Y. (2016). Trusting social media as a source of health information: Online surveys comparing the United States, Korea, and Hong Kong. *Journal of Medical Internet Research*, 18(3), e4193.
- Son, J., Lee, H. K., Jin, S., & Lee, J. (2019). Content features of tweets for effective communication during disasters: A media synchronicity theory perspective. *International Journal of Information Management*, 45, 56–68.
- Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4), 217–248.
- Taecharungroj, V. (2017). Starbucks' marketing communications strategy on Twitter. *Journal of Marketing Communications*, 23(6), 552–571.
- Theocharis, Y., Cardenal, A., Jin, S., Aalberg, T., Hopmann, D. N., Strömbäck, J., ... Štětka, V. (2021). Does the platform matter? Social media and COVID-19 conspiracy theory beliefs in 17 countries. *New Media & Society*.
- Vraga, E. K., & Bode, L. (2018). I do not believe you: How providing a source corrects health misperceptions across social media platforms. *Information, Communication & Society*, 21(10), 1337–1353.
- Wawrzuta, D., Klejdysz, J., Jaworski, M., Gotlib, J., & Panczyk, M. (2022). Attitudes toward COVID-19 vaccination on social media: A cross-platform analysis. *Vaccines*, 10(8), 1190.
- Weeks, B. E. (2015). Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication*, 65(4), 699–719.
- White, M., & Dorman, S. M. (2001). Receiving social support online: Implications for health education. *Health Education Research*, 16(6), 693–707.
- Wilson, S. L., & Wiysonge, C. (2020). Social media and vaccine hesitancy. *BMJ Global Health*, 5(10), Article e004206.

- Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: Implications for effective public health campaigns. *Health Education & Behavior, 27*(5), 591–615.
- Wong, L. P., Alias, H., Danaee, M., Ahmed, J., Lachyan, A., Cai, C. Z., ... Zhao, Q. (2021). COVID-19 vaccination intention and vaccine characteristics influencing vaccination acceptance: A global survey of 17 countries. *Infectious Diseases of Poverty, 10*(1), 1–14.
- Wu, W., Lyu, H., & Luo, J. (2021). Characterizing discourse about COVID-19 vaccines: A Reddit version of the pandemic story. *Health Data Science, 2021*, 1–11.
- Yoo, W., Paek, H. J., & Hove, T. (2020). Differential effects of content-oriented versus user-oriented social media on risk perceptions and behavioral intentions. *Health Communication, 35*(1), 99–109.
- Zhang, J., Wang, Y., Shi, M., & Wang, X. (2021). Factors driving the popularity and virality of COVID-19 vaccine discourse on Twitter: Text mining and data visualization study. *JMIR Public Health and Surveillance, 7*(12), Article e32814.
- Zheng, C., Xue, J., Sun, Y., & Zhu, T. (2021). Public opinions and concerns regarding the Canadian prime minister's daily COVID-19 briefing: Longitudinal study of YouTube comments using machine learning techniques. *Journal of Medical Internet Research, 23*(2), Article e23957.
- Zhu, Y. Q., & Chen, H. G. (2015). Social media and human need satisfaction: Implications for social media marketing. *Business Horizons, 58*(3), 335–345.
- Zollo, F., Novak, P. K., Del Vicario, M., Bessi, A., Mozetič, I., Scala, A., ... Quattrociochi, W. (2015). Emotional dynamics in the age of misinformation. *PLoS One, 10*(9), Article e0138740.