10

# How to Manage Issues on Twitter

## *Perspectives from Twitter Users Concerned about Mis/Dis-Information*

*Joanna Sin Sei-Ching and Kyung-Sun Kim*

### INTRODUCTION

Social media has brought significant changes to people's lives. Some even suggest that social media rewires our brains (MSNBC 2022). While more evidence is needed on the latter, most would agree with the former, especially regarding changes in people's communication and information behavior. Social media platforms have become popular sources of information in everyday life, through which users stay informed and connected. A Pew Research Center study (hereafter, *Pew study*) found, for example, that almost half of US adults get news from social media at least sometimes (Walker and Matsa 2021). Unfortunately, problems on social media, such as mis/dis-information, have become increasingly pressing. Around 64 percent of US adults felt that social media has mostly negative effects on how things are going in the country (Auxier 2020). It is critical for scholars and policymakers to explore ways to tackle these issues. It is equally important to hear from the community of users themselves as they are more directly involved.

The current study focuses on Twitter, a popular microblogging platform. Among various social media platforms (e.g., Facebook, YouTube, Instagram), Twitter has the largest share of users (55 percent) who get news from the site regularly (Walker and Matsa 2021). Twitter is also a go-to platform for breaking news (Osborne and Dredze 2014) and plays a notable role in crisis communication and disaster management, such as during natural disasters and mass shootings (Acar and Muraki 2011). However, in part due to the short length, large volume, and easy and speedy distribution of postings, the information environment of Twitter can be volatile and challenging (Sankaranarayanan et al. 2009; Sin 2016). Like other social media, Twitter also faces extensive mis/dis-information issues including disinformation campaigns during elections, and vaccine misinformation (Bovet and Makse 2019;

Joanna Sin Sei-Ching is an Associate Professor at the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore. Kyung-Sun Kim is a Professor at the Information School, University of Wisconsin-Madison.



Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010




FIGURE 10.1 Visual themes from how to manage issues on Twitter: perspectives from Twitter users concerned about mis/dis-information.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

Chamberlain 2010; Linvill and Warren 2020), which pose significant and long-lasting threats to individuals and societies. It is disconcerting that falsehoods often spread on social media faster and further than truth, and that they persist and resurge even after they have been debunked, as shown in other Twitter studies (Shin et al. 2018; Vosoughi, Roy, and Aral 2018). In a Pew study, 91 percent of US adults reported encountering at least some inaccurate or misleading information on Twitter, while 33 percent said they encountered such information "a lot" (Odabaş 2022b). It is thus of interest to examine how Twitter users (also called *Twitterers*) handle these mis/dis-information issues on the platform.

As a platform where a community of users creates and shares information, social media such as Twitter can be conceptualized as knowledge commons. Knowledge commons is a complex ecosystem where a resource is shared by a group of people that is subject to social dilemmas (Hess and Ostrom 2007). Focusing on the governance structure, the Governing Knowledge Commons Framework (GKC) (Frischmann, Madison, and Strandburg 2014; Madison, Frischmann, and Strandburg 2010) views the knowledge commons as "the institutionalized community governance of the sharing and, in some cases, creation, of information, science, knowledge, data, and other types of intellectual and cultural resources" (Frischmann, Madison, and Strandburg 2014, 3). So far, little is known about the governance of social media as knowledge commons, such as what users do and want for social media platform governance (Riedl, Whipple, and Wallace 2021).

Informed by the GKC (Frischmann, Madison, and Strandburg 2014; Madison, Frischmann, and Strandburg 2010), which modified and extended Ostrom's Institutional Analysis and Development framework (IAD), this study is designed to understand Twitter's user community and governance: how Twitter users manage day-to-day issues, and what actions and actors they see as vital to governing the platform – a knowledge commons.

The first set of research questions (RQs) focus on the *action arena*, exploring what *actions* have been taken by Twitter users and what actions and actors they deemed important:

**RQ1a**: How frequently do Twitter users *take various actions* when encountering problems on Twitter?
**RQ1b**: To what extent do Twitter users *think others should take various actions* for managing problems on Twitter?
**RQ1c**: To what extent do Twitter users think different *actors should take responsibility* for managing problems on Twitter?

GKC and IAD posit that the action arena can be affected by several underlying factors, including the attributes of the community. A heterogeneous community could mean that the knowledge commons is more challenging and costly to maintain, as community members are likely to have varying values and interests (Ostrom and Hess 2007). This is pertinent to Twitter, where the community is diverse (Auxier and Anderson 2021; Chaffey 2022). Moreover, a social media user

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

may undertake multiple roles (e.g., sharers of existing resources, creator of new resources, cocreators of collective resources), which would add another layer of complexity to knowledge production and management (Frischmann, Madison, and Strandburg 2014). We cannot assume that community members will act the same way and share the same view regarding the governance of the commons. The second set of RQs thus examines whether individuals with different characteristics would act or think differently:

**RQ2**: Are there demographic differences (gender, age, education, frequency of use) in: (a) how frequently Twitter users themselves take various actions; (b) the extent to which Twitter users want various actions to be taken by others; and (c) the extent to which Twitter users think different actors should take responsibility?

Scholars and policymakers will find the study findings useful when developing plans to tackle rising challenges on Twitter, such as mis/dis-information. Findings on demographic differences could also inform policies and system designs to be more attuned to individual differences. In addition, the findings could be of interest to researchers who are collecting empirical evidence to develop hypotheses and models on users' social media information behavior. Through the study findings, Twitter users can gain insights into their fellow community members regarding their actions and preferences in managing Twitter problems. Such insights could be instrumental for users in improving Twitter and other social media communities through collective actions.

## BACKGROUND AND LITERATURE REVIEW

### *Social Media Platforms as Knowledge Commons*

Social media, "online and mobile technologies or platforms people use to interact and share content" (Chandler and Munday 2016b), is often seen as heralding a new era in how information and content are produced, communicated, and consumed (Bruns 2007). Among the different types of social media, microblogs allow users to broadcast short text messages (Chandler and Munday 2016a). Twitter, one of the most popular microblog platforms, was estimated to have about 329 million users worldwide in 2022 (Statista 2022), posting over 690 million tweets per day on average (Internet Live Stats 2022). Many US users (89 percent) keep their accounts public (McClain et al. 2021), thus allowing people without Twitter accounts to read their public tweets. As a digital space and virtual community where a group of individuals come together to share content, social media platforms are shared informational resources (i.e., knowledge commons). Corresponding analytical frameworks, such as GKC and IAD, can offer insight into understanding people's decision-making and behaviors on these platforms (Ostrom and Hess 2007).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

### The Governing Knowledge Commons Framework (GKC)

Overall, GKC puts forth the following constructs and relationships. First, it identifies three groups of interacting variables – *resource characteristics*, *attributes of the community*, and *rules-in-use* – as underlying factors. These factors would affect the *action arena*, including *action situations* and *actors*. Specifically, the action arena captures how community members make decisions within a situation. These decisions would produce various *patterns of interaction*, which feed back to the three underlying factors and the action arena. *Evaluation criteria* can be leveraged to assess the interaction patterns. Regarding details of the GKC framework, readers may consult Chapter 1 of this volume.

The current study focuses on the action arena, which Ostrom and Hess (2007) identified as "often at the heart of the analysis" and "particularly useful in analyzing specific problems or dilemmas in the process of institutional change" (45). Before zooming into the action arena, it would be helpful to contextualize the study environment. Based on GKC and IAD, the following sections will present the description of the study context and relevant literature related to the three groups of underlying factors: *resource characteristics*, *attributes of the community*, and *rules-in-use*.

### Resource Characteristics

On resource characteristics, the Twitter site (https://twitter.com) can be viewed as the *facilities*. It is where *artifacts* (e.g., the tweets) expressing various *ideas* are stored and made available to the public (Ostrom and Hess 2007). While it is possible that some materials (e.g., photos) posted on Twitter are copyrighted and thus have restrictions on reuse, the ideas communicated in digital form through Twitter are nonrivalrous (Frischmann, Madison, and Strandburg 2014). That is, a person reading a tweet and using its ideas does not deplete the resource pool nor prevent others from accessing the same tweet and ideas. Through Twitter, users can post and access different types of information and ideas, including news (Kim, Sin, and Yoo-Lee 2014; Walker and Matsa 2021), especially breaking news (Osborne and Dredze 2014), citizen journalism (Murthy 2011), social activism information (Sandoval-Almazan and Gil-Garcia 2014), opinions and popular trends (Kim and Sin 2016; Kim, Sin, and Yoo-Lee 2021), and entertainments (McClain et al. 2021). Twitter is also a source of expert and academic information (Mohammadi et al. 2018). Especially after 2017 when Twitter added threads as a feature, experts are known to write "tweetorials," a series of tweets to explain technical concepts to a wider public audience (Breu 2019). Overall, the informational and educational potentials of Twitter as a shared knowledge commons are considerable. Unfortunately, Twitter could not escape the "dark side" of social media (Baccarella et al. 2018). Mis/dis-information, hate speech, harassment, and doxing are among the problems plaguing the Twitter community (MacAllister 2016; McClain et al. 2021; Whittaker and Kowalski 2015).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

### Attributes of the Community

Regarding knowledge commons' attributes of the community, IAD identifies *information users*, *information providers*, and *information managers/policymakers* (Ostrom and Hess 2007). A defining characteristic of social media like Twitter is that they blur the line between information users and information providers (Bruns 2007). Such blurring of roles is highlighted by the GKC (Frischmann, Madison, and Strandburg 2014). Instead of being mere receivers, social media users now have the ability to broadcast user-generated content to large masses at low costs, allowing diverse and marginalized voices that have been traditionally sidelined by mass media to come through. The democratizing and transformative potential of social media has long been lauded. However, it is increasingly apparent that along with potential, significant dangers also exist (Picard 2015). The quality of information on social media is often found wanting (McClain et al. 2021), and there is evidence of coordinated disinformation campaigns (Beskow and Carley 2020; Hindman and Barash 2018). Exacerbating the problem is that credibility assessment of social media content is often challenging. For example, the original source of the content may be obscured, rendering it difficult to apply traditionally well-used heuristics such as checking source expertise and trustworthiness (Wathen and Burkell 2002). Users are often found to rely on superficial heuristics such as the date or the length of a message when evaluating social media content (Kim, Sin, and Yoo-Lee 2021). Furthermore, poor information evaluation awareness or skills are not always the culprit (Kim and Sin 2011). It appears that, sometimes, truthfulness simply is not a high priority in users' everyday social media sharing. Users' willingness to share a message on social media is not necessarily affected by the accuracy or trustworthiness of the message. Instead, a message is shared because it is perceived as novel, eye-catching, and a good topic of conversation (Chen et al. 2015a; Leeder 2019). The unfortunate outcome is that social media users often unknowingly become spreaders of mis/dis-information.

Not all Twitter "users" are humans. Some are bots, automated accounts programmed to perform specific tasks. Bots can provide useful functions, such as adding video captions (e.g., @HeadlinerClip). News bots (Lokot and Diakopoulos 2016) (e.g., @earthquakeBot, @FintechBot, @MagicRealismBot, and @parliamentedits) help share breaking news or topical information and commentary, while others send self-care (e.g., @tinycarebot) and positive messages (e.g., @TheNiceBot). Unfortunately, the values bots can bring to the Twitter community are often overshadowed by malicious bots, which play significant roles in spreading falsehood, spamming, and phishing (Alothali et al. 2018). Bots have been found to target influential users (e.g., Twitter users with many followers) by mentioning or replying to them with links to low-credibility information, hoping the information will be reshared (Shao et al. 2018). Removing those bots could improve the quality of social media information considerably. However, researchers have emphasized that bots alone cannot account for the spread of falsehood. The actions of human users still

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

matter to the propagation of mis/dis-information (Shao et al. 2018; Vosoughi, Roy, and Aral 2018).

The management team of Twitter, Inc., the company that owns Twitter, is the primary policymaker. Government policymakers also play a role at a broader level, as both regular users and Twitter, Inc. are subject to prevailing state, national, and international laws. Together, these policy markers establish the rules that set the boundaries of users' day-to-day operations.

**Rules-In-Use**

In examining rules-in-use, the degree of openness and nature of control is a focus of GKC (Frischmann, Madison, and Strandburg 2014). Except in countries where Twitter is banned, it is relatively open to access and participation. IAD identified three levels of rulemaking: operational, policy, and constitutional (Ostrom and Hess 2007). In the context of this study, the operational level is where Twitter users make day-to-day decisions on how they act and interact with each other on the platform. This level is the focus of the current study. Regarding the policy level, it is where people, such as the management of Twitter, Inc., make rules concerning users' day-to-day operations. While Twitter is quite open to public participation, as noted above, control is ultimately vested in Twitter, Inc. Specifically, users' actions on Twitter are subjected to the Twitter User Agreement, including the *Terms of Services*, *Privacy Policy*, and *Twitter Rules and Policies* set forth by Twitter, Inc. (Twitter 2022). Beyond policy, the technology itself can constrain users through system design. For example, tweets were intended to be short, with an original 140 characters and a current 280 characters limit per tweet. Critics have observed that this design demands ideas to be simplified, rendering Twitter ill-suited for nuanced discussions (Ott 2017). The brevity also means that contextual information is often absent, making credibility assessment difficult (Sankaranarayanan et al. 2009; Sin 2016).

The challenges in credibility assessment are not helped by the fact that most US social media companies traditionally have not set strict rules on mis/dis-information (Wardle and Singerman 2021). It was only in 2020 when Twitter began taking more visible actions on mis/dis-information on specific topics, including labeling misleading information related to COVID-19 (Roth and Pickles 2020), the US 2020 election (Gadde and Beykpour 2020), and, more recently, Russia's invasion of Ukraine (Roth 2022).

Legal structures affect knowledge commons (Frischmann, Madison, and Strandburg 2014), which is salient to the current study as both Twitter users and Twitter, Inc., are bounded by relevant legislation. In the US, Section 230 of the federal code has become an issue of contention. Section 230 states:

> "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider" (United States Communications Decency Act 1996) and that they should not be liable for "any action voluntarily taken in good faith to restrict access to or

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

> availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected . . ." (United States Communications Decency Act 1996)

This gives companies such as Twitter, Inc. the power to make their own decisions on what content to keep on their site or to take down. Laws such as Section 230 operate at IAD's constitutional level of rulemaking. Section 230 defines that in the US, it is primarily Twitter, Inc., and not the government, who makes decisions regarding the handling of mis/dis-information on Twitter.

### *Individual Differences: Demographics and Use Behavior*

GKC highlights attention to the attributes of the community and their influences on the *patterns of interaction*. We thus posit that Twitter users' attributes/characteristics may be associated with different behaviors and preferences in the action arena. This aligns with the views and findings from information behavior research, which has often observed various individual and group differences (Case and Given 2016). As a beginning exploration, the current study first examined basic demographic attributes, specifically, age, gender, educational attainment, and frequency of Twitter use.

Although the findings are not unanimous, gender differences are often found in various aspects of social media usage. These include differences in strategies used for evaluating social media information (Kim and Sin 2015), motivations for sharing misinformation (Chen et al. 2015b), and perception of online risk (Lim and Kwon 2010). Furthermore, while men and women share some similarities in how they perceive online trolls, their actions differ. Men were more likely to confront or block trolls, whereas women were more likely to ignore the trolling and discourage confrontation (Fichman and Sanfilippo 2014). A meta-analysis also found that women had greater concerns about social media privacy issues than men. Unlike men, women were more likely to activate privacy settings and less likely to disclose personal information in their profiles (Tifferet 2019). Studies of gender differences in social media have primarily focused on men and women. Less is known about how different groups in the entire gender spectrum compare to each other in their social media use.

Most discussions on age differences in social media focus on the prevalence of use. For example, Twitter is found to be more popular among younger individuals (McClain et al. 2021). Beyond prevalence, some researchers have posited that older users may act differently, such as exercising less self-disclosure (Taddicken 2014) or having different norms regarding emotional expression on social media (Waterloo et al. 2017). The findings have been mixed, however.

Educational differences are less often explored. However, education level was found positively related to the support for social media content reviews (Riedl, Whipple, and Wallace 2021). Microblog users who were college underclassmen reported more difficulties in finding everyday information compared to upperclassmen, master's, and doctoral students (Sin and Kim 2014).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

Frequency of use can also have influences on users' attitudes and behaviors. For Twitter, a Pew study found several differences. For example, infrequent users followed fewer accounts than frequent users. They were also more interested in using Twitter to learn about different viewpoints rather than to express their own opinions (Odabaş 2022a). Another study found nonlinear, u-shape patterns between frequency of microblog use and everyday information seeking. Occasional microblog users reported a higher level of difficulties in finding everyday information than non- or frequent users and were less satisfied with the quality of the information found (Sin and Kim 2014). Frequent social media users tended to be less supportive of government regulation of platforms (Riedl, Whipple, and Wallace 2021).

## RESEARCH METHODS

### *Data Collection*

Study participants were Twitter users in the US who found misinformation on Twitter problematic. Data were collected using an online questionnaire. Participants answered questions on their demographics, what actions were taken or should be taken to manage problems on Twitter, and what groups should take responsibility. The survey was conducted in Fall 2018. The sample was recruited from Amazon Mechanical Turk (MTurk). MTurk is among the most frequently used channels for recruiting participants online (Aguinis, Villamor, and Ramani 2021), as it is known to facilitate the recruitment of diverse samples (Casler, Bickel, and Hackett 2013).

The current study focused on participants who considered misinformation on Twitter to be problematic. This is operationalized by participants' responses to two questions, "Regarding tweets/retweets posted by Twitter users outside your social network, how problematic is fake news?" and "Regarding tweets/retweets posted by Twitter users outside your social network, how problematic is inaccurate information?" To account for differences in how individuals perceive social boundaries, participants were to decide who was considered "outside [their] social network" instead of following a universal definition set by the researchers. Participants who have an average score of four or above on a five-point scale (indicating "Very problematic" to "Extremely problematic") on the two questions were included in the current analysis. The questionnaire was hosted on Qualtrics. A pilot test was done with ten MTurkers before data collection.

### *Data Analysis*

Descriptive statistics, including means and standard deviations, were used to address RQ1. Robust multiway ANOVAs were used for RQ2. Robust multiway ANOVA is an inferential test for comparing group means. Particularly, it analyzes the relationship between a continuous dependent variable (DV) and multiple categorical

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

independent variables (IVs). Robust multiway ANOVA has the advantage of not requiring the assumptions of regular multiway ANOVA (Field 2009). In this study, the DVs are the top five answers from RQ1a, RQ1b, and RQ1c, respectively. There are four IVs: age, gender, educational attainment, and frequency of Twitter use. Analyses were done using SPSS.

## RESULTS

### *Participant Characteristics*

The sample consisted of 400 US participants on MTurk who were registered users of Twitter and found misinformation on Twitter "very" or "extremely problematic." In terms of age, participants in their thirties constituted the majority of the sample ($n$ = 202, 50.5%), followed by those in their forties ($n$ = 80, 20%) and twenties ($n$ = 60, 15%), fifties ($n$ = 40, 10%) and sixty and above ($n$ = 18, 4.5%) (Table 10.1). There were slightly more women ($n$ = 205, 51.3%) than men ($n$ = 193, 48.3%). Two participants (0.5%) selected "Other." Both identified themselves as genderqueer. A plurality of the

TABLE 10.1 *Participant characteristics*

| | Count | Percentage (%) |
|---|---|---|
| **Age group** | | |
| 20s | 60 | 15.0 |
| 30s | 202 | 50.5 |
| 40s | 80 | 20.0 |
| 50s | 40 | 10.0 |
| 60 or above | 18 | 4.5 |
| **Gender** | | |
| Man | 193 | 48.3 |
| Woman | 205 | 51.3 |
| Other | 2 | 0.5 |
| **Educational attainment** | | |
| High school | 41 | 10.3 |
| Some college | 152 | 38.0 |
| Bachelor's degree | 161 | 40.3 |
| Master's degree or above | 45 | 11.3 |
| (Missing) | 1 | 0.3 |
| **Frequency of use** | | |
| Less than monthly | 68 | 17.0 |
| Monthly | 74 | 18.5 |
| Weekly | 116 | 29.0 |
| Daily | 116 | 29.0 |
| Hourly | 26 | 6.5 |

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

TABLE 10.2 *Actions taken by participants: frequency of taking actions*

| Rank | Actions | M | SD |
|---|---|---|---|
| 1 | Unfollow the problematic user | 3.76 | 1.19 |
| 2 | Verify/fact check the information yourself | 3.46 | 1.12 |
| 3 | Mute the account of the problematic user | 3.31 | 1.39 |
| 4 | Block the account of the problematic user | 3.30 | 1.35 |
| 5 | Check the bio and other posts of the problematic user to learn more about them | 3.19 | 1.17 |
| 6 | Wait and see what happens | 2.80 | 1.13 |
| 7 | Avoid reading tweets posted by someone outside of your social network | 2.74 | 1.21 |
| 8 | Spend less time on Twitter | 2.48 | 1.20 |
| 9 | Report the problematic user to Twitter | 2.45 | 1.35 |
| 10 | Spend more time on other social media platforms | 2.37 | 1.21 |

Scales used: 1: Almost never; 2: Occasionally; 3: Sometimes; 4: Frequently; 5: Almost always.

participants held a bachelor's degree ($n$ = 161, 40.3%). Participants with some college education were the second largest group ($n$ = 152, 38.0%), followed by those with a master's or above ($n$ = 45, 11.3%) and those with a high school diploma ($n$ = 41, 10.3%). Most of the participants indicated that they used Twitter daily or weekly.

### *Actions Taken by Participants (RQ1a)*

When encountering problems on Twitter, "Unfollow the problematic user" was the action participants reported taking most frequently (Table 10.2). The second most frequently used action was "Verify/fact check the information yourself." The top five results fall broadly into two categories. One category includes actions that can reduce future encounters with problematic tweets and users (e.g., unfollowing, muting, or blocking). The second group is related to information seeking/verification (e.g., verifying the information and checking the bio and posts of problematic users).

More assertive actions, such as reporting problematic users to Twitter, were used less frequently (ranked ninth out of ten actions). Participants also did not seem too deterred from using Twitter due to problems there. "Spend less time on Twitter" and "Spend more time on other social media platforms" were used only occasionally (ranked eighth and tenth, respectively).

### *Actions to Be Taken by Others (RQ1b)*

Regarding the actions that participants want others to take (Table 10.3), an institutional effort, "Ban problematic users," received the highest average score ($M$ = 4.04, $SD$ = 1.24). This is followed by actions in system design, including

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

TABLE 10.3 *Actions to be taken by others: extent to which participants want others to take them*

| Rank | Actions | M | SD |
|---|---|---|---|
| 1 | Institutional efforts: Ban problematic users from Twitter | 4.04 | 1.24 |
| 2 | System design: Provide better functions for users to filter things easily | 4.01 | 0.99 |
| 3 | System design: Provide better functions for users to report problematic users/tweets easily | 3.89 | 1.16 |
| 4 | System design: Provide functions for moderators to issue warnings to problematic users | 3.81 | 1.23 |
| 5 | Institutional efforts: Employ people to monitor and problematic user/tweet | 3.69 | 1.27 |
| 6 | System design: Provide functions for users to downvote or dislike problematic users/tweets | 3.55 | 1.33 |
| 7 | System design: Provide functions to automatically and silently block problematic users/tweets without requiring users' inputs | 3.44 | 1.43 |
| 8 | Institutional efforts: Remind users regularly of Twitter's rules, policies and terms of services | 3.18 | 1.28 |
| 9 | System design: Provide functions to reward users who have rightfully flagged a problematic user/tweet | 3.01 | 1.42 |
| 10 | Institutional efforts: Remind users regularly of microblogging etiquettes | 2.95 | 1.32 |
| 11 | Institutional efforts: Enact and enforce laws to deter and punish problematic behavior on Twitter | 2.91 | 1.47 |
| 12 | Institutional efforts: Provide training and resources to users on how to use Twitter properly and safely | 2.80 | 1.38 |
| 13 | Users' efforts: Twitter users to form volunteer groups to monitor and flag problems | 2.71 | 1.24 |
| 14 | Users' efforts: Individual Twitter user to rebut and deter other users' problematic behaviors | 2.69 | 1.22 |
| 15 | Institutional efforts: Require all tweets to be posted with real names | 2.57 | 1.57 |

Scales used: 1: Not at all; 2: To a small extent; 3: To a moderate extent; 4: To a great extent; 5: To a very great extent.

providing better functions for individual users "to filter things easily" ($M = 4.01$, $SD = 0.99$) and "to report problematic users/tweets easily" ($M = 3.89$, $SD = 1.16$). The fifth highest score went to an institutional effort, "Employ people to monitor and problematic user/tweet" ($M = 3.69$, $SD = 1.27$).

Overall, actions related to the system-design category tend to score higher, with average scores ranging from three (*to a moderate extent*) to four (*to a great extent*). In contrast, participants did not prefer actions in the users' efforts category. "Twitter users to form volunteer groups to monitor and flag problems" and "Individual Twitter user to rebut and deter other users' problematic behaviors" scored the second and the third lowest out of fifteen strategies. Actions in the institutional efforts category saw wider variations. For example, while the highest-scoring effort

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

TABLE 10.4 *Actors to take responsibility: extent to which participants want them to be responsible*

| Rank | Actors | M | SD |
|---|---|---|---|
| 1 | The Twitter company | 4.55 | 0.93 |
| 2 | Individual Twitter user | 3.32 | 1.27 |
| 3 | Prominent Twitter users/influencers | 3.02 | 1.40 |
| 4 | Search engines, social media, and internet companies | 2.93 | 1.39 |
| 5 | Social media experts, subject specialists, and scholars | 2.86 | 1.35 |
| 6 | All members of the public | 2.73 | 1.31 |
| 7 | The media and journalists | 2.46 | 1.27 |
| 8 | Social activists | 2.43 | 1.31 |
| 9 | Law enforcement agencies | 2.32 | 1.18 |
| 10 | Schools, educators, and librarians | 2.21 | 1.30 |
| 11 | The government, politicians, and elected officials | 1.91 | 2.15 |
| 12 | Nonprofit organizations | 1.62 | 1.75 |

Scales used: 1: Not at all; 2: To a small extent; 3: To a moderate extent; 4: To a great extent; 5: To a very great extent.

(i.e., "Ban problematic users") was an institutional effort, the lowest-scoring was also an institutional effort (i.e., "Require all tweets to be posted with real names").

### *Actors to Take Responsibility (RQ1c)*

Regarding the extent to which different groups should take responsibility for managing problems on Twitter, the answer for "The Twitter company" (Twitter, Inc.) stood out ($M = 4.55$, $SD = 0.93$), indicating an average score ranging from *a great extent* to *a very great extent*. Notably, Twitter, Inc. scored considerably higher than any other groups on the list (Table 10.4).

Results on the top four highest-scoring groups show that participants centered the responsibility on companies and users. Beyond Twitter, Inc., mentioned above, "Search engines, social media, and internet companies" ranked fourth. The second- and third-highest groups were related to users: "Individual Twitter user" and "Prominent Twitter users/influencers." Experts ("Social media experts, subject specialists, and scholars") rounded out the fifth. In contrast, governmental groups ("Law enforcement agencies" and "The government, politicians, and elected officials") were seen as having responsibilities to only *a small extent*. Their average scores ranked ninth and eleventh out of twelve groups.

### *Demographic Differences in Actions Taken by Participants (RQ2a)*

"Unfollow the problematic user" was the action most frequently taken by participants (in RQ1a). The robust multiway ANOVA analysis revealed a statistically

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

| | Unfollow the problematic user | | Verify/fact check the information yourself | | Mute the account of the problematic user | | Block the account of the problematic user | | Check the bio and other posts of the problematic user to learn more about them | |
|---|---|---|---|---|---|---|---|---|---|---|
| | B | partial η2 | B | partial η2 | B | partial η2 | B | partial η2 | B | partial η2 |
| Intercept | 3.99 *** | 0.17 | 3.34 *** | 0.14 | 4.01 *** | 0.15 | 3.95 *** | 0.16 | 3.64 *** | 0.17 |
| **Age group** | | | | | | | | | | |
| 20s | 0.08 | 0.00 | -0.06 | 0.00 | -0.16 | 0.00 | -0.44 | 0.00 | -0.28 | 0.00 |
| 30s | 0.03 | 0.00 | -0.11 | 0.00 | 0.06 | 0.00 | -0.18 | 0.00 | -0.07 | 0.00 |
| 40s | -0.08 | 0.00 | -0.06 | 0.00 | 0.06 | 0.00 | -0.40 | 0.00 | -0.32 | 0.00 |
| 50s | 0.37 | 0.00 | 0.06 | 0.00 | 0.18 | 0.00 | 0.10 | 0.00 | 0.13 | 0.00 |
| 60 or above (reference group) | | | | | | | | | | |
| **Gender** | | | | | | | | | | |
| Man | -0.32 * | 0.02 | 0.02 | 0.00 | -0.34 * | 0.01 | -0.31 * | 0.01 | 0.02 | 0.00 |
| Woman (reference group) | | | | | | | | | | |
| **Educational attainment** | | | | | | | | | | |
| High school | 0.27 | 0.00 | 0.17 | 0.00 | -0.06 | 0.00 | -0.06 | 0.00 | -0.20 | 0.00 |
| Some college | 0.13 | 0.00 | 0.05 | 0.00 | -0.04 | 0.00 | 0.11 | 0.00 | 0.14 | 0.00 |
| Bachelor's degree | 0.15 | 0.00 | 0.02 | 0.00 | -0.05 | 0.00 | 0.01 | 0.00 | 0.15 | 0.00 |
| Master's degree or above (reference group) | | | | | | | | | | |
| **Frequency of use** | | | | | | | | | | |
| Less than monthly | -0.43 | 0.01 | -0.02 | 0.00 | -0.93 * | 0.02 | -0.56 | 0.01 | -0.79 ** | 0.02 |
| Monthly | -0.38 | 0.00 | 0.24 | 0.00 | -0.93 * | 0.02 | -0.75 * | 0.01 | -0.53 | 0.01 |
| Weekly | -0.30 | 0.00 | 0.09 | 0.00 | -0.53 | 0.01 | -0.31 | 0.00 | -0.62 * | 0.02 |
| Daily | -0.05 | 0.00 | 0.26 | 0.00 | -0.15 | 0.00 | 0.04 | 0.00 | -0.08 | 0.00 |
| Hourly (reference group) | | | | | | | | | | |

* $p < .05$ ** $p < .01$ *** $p < .001$

FIGURE 10.2 Robust multiway ANOVA results: demographic differences in actions taken by participants.

significant demographic difference (Figure 10.2). Specifically, men tended to use this action less frequently than women. In contrast, the second most used action, "Verify/fact check the information yourself," saw no significant differences.

Gender differences were also found in the third and fourth most frequently taken actions: Men used "Mute the account of the problematic user" and "Block the account of the problematic user" less. Frequency of Twitter use was also statistically significant. Compared to the reference group (hourly users), participants who tweeted infrequently (e.g., less than monthly or monthly) tended to mute or block others less.

For "Check the bio and other posts of the problematic user," the frequency of use was again statistically significant. Infrequent users (e.g., less than monthly) and weekly users applied the above action less than the reference group.

### *Demographic Differences in Actions to Be Taken by Others (RQ2b)*

In contrast to the above results, where gender and frequency of use differences were found, no such difference was found regarding the top actions participants wanted others to take.

### *Demographic Differences in Actors Considered Responsible (RQ2c)*

"The Twitter company," the highest-ranked group that participants consider most responsible, saw a statistically significant difference in frequency of Twitter use

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

| | The Twitter company | | Individual Twitter user | | Prominent Twitter users/influencers | | Search engines, social media, and Internet companies | | Social media experts, subject specialists, and scholars | |
|---|---|---|---|---|---|---|---|---|---|---|
| | B | partial η2 | B | partial η2 | B | partial η2 | B | partial η2 | B | partial η2 |
| Intercept | 4.18 *** | 0.26 | 4.30 *** | 0.23 | 4.02 *** | 0.14 | 3.94 *** | 0.14 | 3.77 *** | 0.15 |
| **Age group** | | | | | | | | | | |
| 20s | 0.00 | 0.00 | -0.55 | 0.01 | -0.13 | 0.00 | -0.59 | 0.01 | -0.13 | 0.00 |
| 30s | 0.07 | 0.00 | -0.36 | 0.01 | -0.21 | 0.00 | -0.49 | 0.01 | -0.20 | 0.00 |
| 40s | 0.14 | 0.00 | -0.44 | 0.01 | -0.62 | 0.01 | -0.68 * | 0.01 | -0.51 | 0.01 |
| 50s | 0.29 | 0.00 | -0.44 | 0.01 | -0.31 | 0.00 | -0.34 | 0.00 | -0.30 | 0.00 |
| 60 or above (reference group) | | | | | | | | | | |
| **Gender** | | | | | | | | | | |
| Man | -0.12 | 0.00 | -0.04 | 0.00 | -0.33 * | 0.01 | -0.36 * | 0.02 | -0.29 * | 0.01 |
| Woman (reference group) | | | | | | | | | | |
| **Educational attainment** | | | | | | | | | | |
| High school | 0.14 | 0.00 | 0.15 | 0.00 | 0.01 | 0.00 | -0.35 | 0.00 | -0.41 | 0.00 |
| Some college | 0.20 | 0.00 | -0.11 | 0.00 | -0.15 | 0.00 | -0.13 | 0.00 | -0.27 | 0.00 |
| Bachelor's degree | -0.13 | 0.00 | 0.08 | 0.00 | -0.04 | 0.00 | -0.05 | 0.00 | -0.34 | 0.01 |
| Master's degree or above (reference group) | | | | | | | | | | |
| **Frequency of use** | | | | | | | | | | |
| Less than monthly | 0.46 * | 0.01 | -0.61 * | 0.01 | -0.65 | 0.01 | -0.15 | 0.00 | -0.07 | 0.00 |
| Monthly | 0.25 | 0.00 | -0.75 ** | 0.02 | -0.75 * | 0.01 | -0.56 | 0.01 | -0.55 | 0.01 |
| Weekly | 0.26 | 0.00 | -0.58 * | 0.01 | -0.44 | 0.01 | -0.30 | 0.00 | -0.35 | 0.00 |
| Daily | 0.32 | 0.01 | -0.52 * | 0.01 | -0.37 | 0.00 | -0.03 | 0.00 | -0.09 | 0.00 |
| Hourly (reference group) | | | | | | | | | | |

* $p < .05$ ** $p < .01$ *** $p < .001$

FIGURE 10.3 Robust multiway ANOVA results: demographic differences in actors to take responsibility.

(Figure 10.3). That is, infrequent users (e.g., less than monthly) tended to rate the company as having higher responsibilities than did the reference group (hourly users). Frequency of use was again significant for the second-highest response – "Individual Twitter user." This time, however, infrequent users rated "Individual Twitter user" lower than the reference group.

"Prominent Twitter users/influencers" also saw a significant difference in frequency of Twitter use. Similar to the "Individual Twitter user," the infrequent users (e.g., monthly) rated "Prominent Twitter users/influencers" as having less responsibility than did the reference group. In addition, gender differences were found. Women tended to rate "Prominent Twitter users/influencers" as having higher responsibilities than men did.

For "Search engines, social media, and internet companies," a similar gender difference was observed: Women rated them higher than men did. Furthermore, age differences emerged. Compared to younger participants, particularly those in their forties, participants who were sixty or above (the reference group) saw these companies as having higher responsibilities.

For the fifth-highest group, "Social media experts, subject specialists, and scholars," a gender difference was found. Women rated the responsibilities of this group higher than men did.

To summarize, women, more than men, tended to view the following three groups as having higher responsibilities in handling Twitter problems: "Prominent Twitter users/influencers"; "Search engines, social media, and internet companies"; "Social media experts, subject specialists, and scholars." Infrequent Twitter users

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

rated "The Twitter company" highly while rating "Individual Twitter user" and "Prominent Twitter users/influencers" lower than did the reference group (hourly users). Age difference was found only in "Search engines, social media, and Internet companies."

## DISCUSSION

### *Action Arena*

Participants' top actions taken when encountering problems on Twitter (RQ1a) can be categorized in two groups: information filtering (including unfollowing, muting, and blocking) and information verification (including fact-checking and checking the bio and posts of problematic users). From the perspective of information literacy education, it is promising that fact-checking, one of the highly recommended strategies, is among the top actions. Interestingly, while all participants considered fake news and inaccurate information on Twitter "very problematic" or "extremely problematic," they did not frequently take actions. Even the top action – unfollowing – averaged only 3.76, and fact-checking averaged 3.46, which translates to between *sometimes* and *frequently* used. A possible reason for this may be that the information filtering strategies used by participants may have helped reduce the chance of encountering problematic users and tweets. Twitter's algorithms (Newberry and Sehl 2022) may also have reinforced the users' filtering actions and reduced unpreferred/unwanted encounters. Even for earlier information systems, "the need for quality filtering of information" was identified as an essential need (Faibisoff and Ely 1976, 6). Quality filtering is no less vital today. For social media platforms such as Twitter, how to ensure necessary filtering does not exacerbate current problems of "echo chamber" and polarization (Cinelli et al. 2020) would require the attention of policymakers and platform designers.

Banning problematic users emerged as the top action that participants wanted others to take. Whether social media companies should ban (or be allowed to ban) problematic users has become increasingly politicized. For example, Republicans and Democrats differed significantly on whether it was right or wrong for some social media companies to ban Donald Trump from their platforms. Among Republicans and Republican leaners, 78 percent said the ban was wrong. In contrast, 89 percent of Democrats and Democrat leaners said the ban was right. Republicans were especially more likely than Democrats to believe that social media platforms intentionally censor political viewpoints (McClain and Anderson 2021). Two states, Texas and Florida, passed laws barring large social media companies from banning people based on their viewpoints. Currently, these laws are on hold as they are being contested in courts, but the debate is expected to continue (Robertson 2022). Unlike the ongoing legal contests, participants' views on the topic have been quite similar. They notably favored such bans *to a large extent* (mean

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

score of 4.04 out of 5). Future studies may explore if this preference holds and if respondents' political affiliation plays a role.

This result on which participants wanted Twitter, Inc. to ban problematic users can be viewed in tandem with the findings from RQ1a about participants' own actions. Participants often unfollowed, muted, or blocked problematic users, but they did not frequently report problematic users to Twitter. Reporting to Twitter ranked only ninth, with a mean score between *occasionally* and *sometimes*. The findings suggest that participants did not take the power of banning users lightly. They appear to be using the function judiciously while recognizing the need to do something to prevent problematic users from disrupting the community further.

Participants saw Twitter, Inc. as most responsible for handling Twitter issues. The mean score (4.55, *to a great extent*) of their level of responsibility was considerably higher than those for other actors such as individual Twitter users or influencers. This finding confirms the calls from the public, scholars, and policymakers for social media companies to take more responsibility (Cusumano, Gawer, and Yoffie 2021; Suzor et al. 2019; Wakabayashi 2019). On the other end of the spectrum, participants did not emphasize the government's role. "The government, politicians, and elected officials" ranked eleventh out of twelve actors to take responsibility. Law enforcement ranked ninth. The option of enacting and enforcing laws also drew little interest (eleventh out of fifteen actions to be taken by others). More studies can be conducted to investigate the reasons behind these preferences. Areas to explore are whether users' preferences are affected by their understanding of internet laws, users' familiarity with the roles different actors play in social media governance, and what users think different actors can do.

The US public has complicated views on government regulation of social media. While the view that the government should restrict false information online has gained more support recently (from 39 percent in 2018 to 48 percent in 2021), a larger share of the public considered that tech companies should take care of the issue (56 percent in 2018 and 59 percent in 2021) (Mitchel and Walker 2021). A more recent Pew study shows a steep decline in the preference for government regulation of tech companies, however. The share of support dropped from 56 percent in 2021 to 44 percent in 2022. This drop was observed not only among Republicans but also among Democrats (Vogels 2022). The issue is likely made more complex by what different levels of governments would regulate. Now, government regulation may not be limited to restricting false or harmful materials. It could also include government regulation that restricts a company's power to handle false information (e.g., the Texas and Florida state laws mentioned above). The *constitutional level of rules making*, which defines "who must, may, or must not participate in making collective choices" (Ostrom and Hess 2007, 50), would become the space to watch. If the center of action indeed shifts to the fights between governments and big tech companies in the legal arena – and away from the day-to-day operations and

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

relationships between a company and its users – whether users may feel sidelined and subsequently disengaged could become a matter of concern. With GKC's attention to the character of control and the "nestness" (Frischmann, Madison, and Strandburg 2014, 32) of knowledge commons in broader sociocultural and institutional settings, future studies may apply GKC to further probe whether exogenous institutions such as the government may gain more control of platform governance over internal governing mechanisms within the Twitter community.

Participants' preference for content moderation is also worth noting. Twitter uses commercial content moderation, including algorithmic moderation. Such moderations are often found to be opaque (Gorwa, Binns, and Katzenbach 2020; Roberts 2018), especially when compared to content moderation done by user volunteers (such as on Reddit and Twitch) (Cook, Patel, and Wohn 2021). In the current study, participants indeed prefer human moderation (fifth) over algorithmic moderation (seventh out of fifteen actions to be taken by others). However, they did not like involving volunteers to monitor or rebuke problems (ranked thirteenth and fourteenth). Instead, they prefer Twitter to employ people to monitor (ranked fifth) and issue warnings to problematic users (ranked fourth). This preference for the social media company (instead of users) to do more on content moderation is also observed in Cook and her colleagues' cross-platform study about content moderation for dealing with toxicity on social media (Cook, Patel, and Wohn 2021).

Unlike their study concluding that users seem to "want to be taken care of when it comes to content moderation as opposed to engaging themselves" (Cook, Patel, and Wohn 2021, 1), participants of the current study seem to want others to provide more ways to engage. The top actions they wanted others to take involved system functions that give them the power to filter things (ranked second), to report users or tweets (third), and to downvote them (sixth out of fifteen actions to be taken by others). It appears that participants do not want to burden users with the *responsibility* of wrangling Twitter problems, as noted above. Instead, they want system features and tools at their disposal, which would allow them to participate in tackling problematic users and tweets whenever they *choose* to.

Differences between previous and current study findings could be due to the differences in study focus. Cook and her colleagues' study, for example, was on social media toxicity, and their data covered multiple platforms, not just Twitter. Another possibility is due to potential differences in motivation levels. The current study focused on participants who considered dis/misinformation on Twitter very or extremely problematic. The perceived seriousness of the problem may have contributed to a higher interest in having the tools available to combat problems themselves. While any user engagement features could potentially be targeted for manipulation by malicious actors (Lee, Tamilarasan, and Caverlee 2021), companies should appreciate users' desire for more agency in their day-to-day interactions on social media.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

### Demographic Differences

Among the four user attributes, frequency of Twitter use and gender yielded more significant differences. Frequent users tended to take actions more frequently and think various actors should take responsibility to a larger extent than infrequent users. The exception is related to Twitter, Inc. Compared to infrequent Twitter users, frequent users rated the company's responsibility lower (while rating individual users' and influencers' responsibilities higher). A possible reason may be that frequent users are more familiar with the dynamics of the platform. They may be more cognizant of the impacts of individual users and influencers on compounding (or lessening) social media problems, such as rumor propagation (Mikhaeil and El Mougy 2020).

The study found that women tended to take actions more frequently and rated several actors with higher responsibilities. One contributing factor may be that women face more gender-based harassment than men on social media (Simons 2015; Winkelman et al. 2015); they have been the targets of disinformation and character attack (Bradshaw and Henle 2021). This may have led women to take more frequent actions to filter out problematic users and believe more strongly that various actors should take more active measures. It is also observed that women of color and other minorities often face situations where they have to provide unpaid labor to actively call out misogyny and racism on social media and end up facing harassment and threats as a result (Nakamura 2015). Along this line, more in-depth research may be needed to understand the lived experience and action arenas of user groups who have been found more targeted on social media, such as racial/ethnic minorities, LGBTQIA+, and people with disabilities (Silva et al. 2016).

### Limitations and Further Studies

The study participants were a self-selected sample. Studies suggested that US MTurkers tend to be younger, have a higher female-to-male ratio, are more educated, and have lower median household income than the overall US population (Aguinis, Villamor, and Ramani 2021; Difallah, Filatova, and Ipeirotis 2018). They have also been found to be more ethnically diverse than other US internet panels (Smith et al. 2016). Overall, the sample may not represent all US adult Twitter users who considered mis/dis-information problematic. For instance, a limitation of this study sample is that there are too few gender-nonconforming participants. Further studies on different samples will shed light on whether they share similar patterns of behavior and preferences. As an initial step, this study focused on only a few user attributes. Future research may include more user attributes such as race/ethnicity, political ideology, and problem-solving styles. Studies may also expand the analysis to include other parts of GKC (e.g., patterns of interaction and evaluation criteria). A more diverse range of methods can also be employed. More inductive and qualitative methods such as interviews can be used to reveal the reasons behind

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

participants' actions and preferences and to investigate how the action arena in turn influences users' thoughts and experiences. Social media data may also be crawled and collected for triangulation. As the socioecological, political, and legal milieus may evolve, longitudinal research would be beneficial to capture changes in the underlying factors and action arena over time. Regarding theoretical development, this study suggests that using GKC to study the Twitter community is promising. More research about this platform with GKC would contribute to (1) understanding the challenges of such research, and (2) identifying areas of development for using GKC to study Twitter and other social media as knowledge commons.

## CONCLUSION

The current study is an initial exploration of the actions and preferences of Twitter users who found mis/dis-information on Twitter problematic. The findings have implications for Twitter platform governance, system design, and everyday management of mis/dis-information on social media. For instance, the study found that participants took more information filtering and fact-checking steps than assertive actions such as reporting problematic users or tweets. Participants wanted problematic users/contents to be banned, as well as having Twitter, Inc. employ people to moderate problematic users/contents. They also sought improvements in system design to provide users with better tools to engage, such as easier reporting and upvoting/downvoting capabilities. Participants considered Twitter, Inc. most responsible for managing problems on Twitter. There was less appetite for actions by governments, however. Some demographic differences were also found in the study. While age showed few differences, frequency of Twitter use and gender differences were more pronounced.

By collecting data from Twitter users who are concerned about mis/dis-information, this study sheds light on the actions they take, and their views on how to and who should take responsibility to confront these issues. Social media policymakers, system designers, researchers, and Twitter users may find insights from the study results. As challenges such as mis/dis-information continue to rage, and the governance of social media appears more and more politicized, it is of paramount importance to understand the everyday decision-making of diverse users and amplify their voices, so as to inform social media policies and system designs to take into consideration users' well-being and agency.

## REFERENCES

Acar, Adam, and Yuya Muraki. 2011. "Twitter for Crisis Communication: Lessons Learned from Japan's Tsunami Disaster." *International Journal of Web Based Communities* 7 (3): 392–402.

Aguinis, Herman, Isabel Villamor, and Ravi S. Ramani. 2021. "MTurk Research: Review and Recommendations." *Journal of Management* 47 (4): 823–837.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

Alothali, Eiman, Nazar Zaki, Elfadil A Mohamed, and Hany Alashwal. 2018. "Detecting Social Bots on Twitter: A Literature Review." In *Proceedings of the 2018 International Conference on Innovations in Information Technology (IIT)*, 175–180. Piscataway, NJ: IEEE.

Auxier, Brooke. 2020. "64% of Americans Say Social Media Have a Mostly Negative Effect on the Way Things Are Going in the US Today." Pew Research Center. www.pewresearch.org/fact-tank/2020/10/15/64-of-americans-say-social-media-have-a-mostly-negative-effect-on-the-way-things-are-going-in-the-u-s-today/.

Auxier, Brooke, and Monica Anderson. 2021. "Social Media Use in 2021." Pew Research Center. www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/.

Baccarella, Christian V., Timm F. Wagner, Jan H. Kietzmann, and Ian P. McCarthy. 2018. "Social Media? It's Serious! Understanding the Dark Side of Social Media." *European Management Journal* 36 (4): 431–438.

Beskow, David M., and Kathleen M. Carley. 2020. "Characterization and Comparison of Russian and Chinese Disinformation Campaigns." In *Disinformation, Misinformation, and Fake News in Social Media: Emerging Research Challenges and Opportunities*, ed. Kai Shu, Suhang Wang, Dongwon Lee, and Huan Liu, 63–81. Cham: Springer.

Bovet, Alexandre, and Hernán A. Makse. 2019. "Influence of Fake News in Twitter during the 2016 US Presidential Election." *Nature Communications* 10 (7): 1–14.

Bradshaw, Samantha, and Amélie Henle. 2021. "The Gender Dimensions of Foreign Influence Operations." *International Journal of Communication* 15: 4598–4618.

Breu, Anthony C. 2019. "Why Is a Cow? Curiosity, Tweetorials, and the Return to Why." *New England Journal of Medicine* 381 (12): 1097–1098.

Bruns, Axel. 2007. "Produsage: Towards a Broader Framework for User-Led Content Creation." In *Proceedings of the 6th ACM SIGCHI Conference on Creativity and Cognition*, 99–106. New York: ACM.

Case, Donald Owen, and Lisa M. Given. 2016. *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior*. 4th ed. Bingley: Emerald.

Casler, Krista, Lydia Bickel, and Elizabeth Hackett. 2013. "Separate but Equal? A Comparison of Participants and Data Gathered Via Amazon's MTurk, Social Media, and Face-to-Face Behavioral Testing." *Computers in Human Behavior* 29 (6): 2156–2160.

Chaffey, Dave. 2022. "Global Social Media Statistics Research Summary 2022." Smart Insights. www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/.

Chamberlain, P. R. 2010. "Twitter as a Vector for Disinformation." *Journal of Information Warfare* 9 (1): 11–17.

Chandler, Daniel, and Rod Munday. 2016a. "Microblogging." In *A Dictionary of Social Media*. Oxford: Oxford University Press.

2016b. "Social Media." In *A Dictionary of Social Media*. Oxford: Oxford University Press.

Chen, Xinran, Sei-Ching Joanna Sin, Yin-Leng Theng, and Chei Sian Lee. 2015a. "Why Do Social Media Users Share Misinformation?" In *Proceedings of the 15th ACM/IEEE–CS Joint Conference on Digital Libraries*, 111–114. New York: ACM.

2015b. "Why Students Share Misinformation on Social Media: Motivation, Gender, and Study-Level Differences." *Journal of Academic Librarianship* 41 (5): 583–592.

Cinelli, Matteo, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. 2020. "Echo Chambers on Social Media: A Comparative Analysis." *arXiv preprint*. https://doi.org/10.48550/arXiv.2004.09603.

Cook, Christine L., Aashka Patel, and Donghee Yvette Wohn. 2021. "Commercial Versus Volunteer: Comparing User Perceptions of Toxicity and Transparency in Content Moderation across Social Media Platforms." *Frontiers in Human Dynamics* 3: 1–8.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

Cusumano, Michael, Annabelle Gawer, and David Yoffie. 2021. "Social Media Companies Should Self-Regulate. Now." *Harvard Business Review*. https://hbr.org/2021/01/social-media-companies-should-self-regulate-now.

Difallah, Djellel, Elena Filatova, and Panos Ipeirotis. 2018. "Demographics and Dynamics of Mechanical Turk Workers." In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 135–143. New York: ACM.

Faibisoff, Sylvia G., and Donald P. Ely. 1976. "Information and Information Needs." *Information Reports and Bibliographies* 5: 2–16.

Fichman, Pnina, and Madelyn Rose Sanfilippo. 2014. "The Bad Boys and Girls of Cyberspace: How Gender and Context Impact Perception of and Reaction to Trolling." *Social Science Computer Review* 33 (2): 163–180.

Field, Andy P. 2009. *Discovering Statistics Using SPSS*. 3rd ed. Los Angeles: SAGE.

Frischmann, Brett M, Michael J. Madison, and Katherine Jo Strandburg. 2014. "Governing Knowledge Commons." In *Governing Knowledge Commons*, ed. Brett M. Frischmann, Michael J. Madison, and Katherine Jo Strandburg, 1–43. Oxford: Oxford University Press.

Gadde, Vijaya, and Kayvon Beykpour. 2020. "Additional Steps We're Taking Ahead of the 2020 US Election." Twitter. https://blog.twitter.com/en_us/topics/company/2020/2020-election-changes.

Gorwa, Robert, Reuben Binns, and Christian Katzenbach. 2020. "Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance." *Big Data and Society* 7 (1): 1–15.

Hess, Charlotte, and Ostrom Elinor. 2007. "Introduction: An Overview of the Knowledge Commons." In *Understanding Knowledge as a Commons: From Theory to Practice*, 3–26. Cambridge, MA: MIT.

Hindman, Matthew, and Vlad Barash. 2018. "Disinformation, and Influence Campaigns on Twitter." *Knight Foundation: George Washington University*. www.issuelab.org/resources/33163/33163.pdf.

Internet Live Stats. 2022. "Twitter Usage Statistics." www.internetlivestats.com/twitter-statistics/.

Kim, Kyung-Sun, and Sei-Ching Joanna Sin. 2011. "Selecting Quality Sources: Bridging the Gap between the Perception and Use of Information Sources." *Journal of Information Science* 37 (2): 178–188.

2015. "Use of Social Media in Different Contexts of Information Seeking: Effects of Sex and Problem-Solving Style." *Information Research* (online) 20 (1). http://informationr.net/ir/20-1/isic2/isic24.html.

2016. "Use and Evaluation of Information from Social Media in the Academic Context: Analysis of Gap between Students and Librarians." *Journal of Academic Librarianship* 42 (1): 74–82.

Kim, Kyung-Sun, Sei-Ching Joanna Sin, and EunYoung Yoo-Lee. 2014. "Undergraduates' Use of Social Media as Information Sources." *College and Research Libraries* 75 (4): 442–457.

2021. "Use and Evaluation of Information from Social Media: A Longitudinal Cohort Study." *Library and Information Science Research* 43 (3): 1–11.

Lee, Kyumin, Prithivi Tamilarasan, and James Caverlee. 2021. "Crowdturfers, Campaigns, and Social Media: Tracking and Revealing Crowdsourced Manipulation of Social Media." In *Proceedings of the International AAAI Conference on Web and Social Media*, 331–340. Palo Alto: AAAI.

Leeder, Chris. 2019. "How College Students Evaluate and Share 'Fake News' Stories." *Library and Information Science Research* 41 (3): 1–11.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

Lim, Sook, and Nahyun Kwon. 2010. "Gender Differences in Information Behavior Concerning Wikipedia, an Unorthodox Information Source?" *Library and Information Science Research* 32 (3): 212–220.
Linvill, Darren L., and Patrick L. Warren. 2020. "Troll Factories: Manufacturing Specialized Disinformation on Twitter." *Political Communication* 37 (4): 447–467.
Lokot, Tetyana, and Nicholas Diakopoulos. 2016. "News Bots." *Digital Journalism* 4 (6): 682–699.
MacAllister, Julia M. 2016. "The Doxing Dilemma: Seeking a Remedy for the Malicious Publication of Personal Information." *Fordham Law Review* 85 (5): 2451–2483.
McClain, Colleen, and Monica Anderson. 2021. "Republicans, Democrats at Odds over Social Media Companies Banning Trump." Pew Research Center. https://pewrsr.ch/3pnZ9d7.
McClain, Colleen, Regina Widjaya, Gonzalo Rivero, and Aaron Smith. 2021. "The Behaviors and Attitudes of U.S. Adults on Twitter." Pew Research Center. www.pewresearch.org/internet/2021/11/15/the-behaviors-and-attitudes-of-u-s-adults-on-twitter/.
Madison, Michael J., Brett M. Frischmann, and Katherine J. Strandburg. 2010. "Constructing Commons in the Cultural Environment." *Cornell Law Review* 95 (4): 657–710.
Mikhaeil, Merna, and Amr El Mougy. 2020. "The Effects of Personality Traits on the Lifetime of a Rumor." In *Highlights in Practical Applications of Agents, Multi-Agent Systems, and Trust-worthiness. The PAAMS Collection*, 108–117. Cham: Springer.
Mitchel, Amy, and Mason Walker. 2021. "More Americans Now Say Government Should Take Steps to Restrict False Information Online Than in 2018." https://pewrsr.ch/37TynlF.
Mohammadi, Ehsan, Mike Thelwall, Mary Kwasny, and Kristi L. Holmes. 2018. "Academic Information on Twitter: A User Survey." *PLoS ONE* 13 (5): e0197265. https://doi.org/10.1371/journal.pone.0197265.
MSNBC. 2022. "Social Media Is Rewiring Our Brains and Changing Our Brain Chemistry." www.msnbc.com/american-voices/watch/social-media-is-rewiring-our-brains-and-changing-our-brain-chemistry-148741189504.
Murthy, Dhiraj. 2011. "Twitter: Microphone for the Masses?" *Media, Culture and Society* 33 (5): 779–789.
Nakamura, Lisa. 2015. "The Unwanted Labour of Social Media: Women of Colour Call out Culture as Venture Community Management." *New Formations* 86: 106–112. www.muse.jhu.edu/article/604492.
Newberry, Christina, and Katie Sehl. 2022. "How the Twitter Algorithm Works in 2022 and How to Make It Work for You." Hootsuite. https://blog.hootsuite.com/twitter-algorithm/.
Odabaş, Meltem. 2022a. "5 Facts About Twitter 'Lurkers'." Pew Research Center. https://pewrsr.ch/3u5R54f.
2022b. "10 Facts About Americans and Twitter." Pew Research Center. https://pewrsr.ch/3vM1gga.
Osborne, Miles, and Mark Dredze. 2014. "Facebook, Twitter and Google Plus for Breaking News: Is There a Winner?" In *Proceedings of the International AAAI Conference on Web and Social Media*, 611–614. Palo Alto: AAAI.
Ostrom, Elinor, and Charlotte Hess. 2007. "A Framework for Analyzing the Knowledge Commons." In *Understanding Knowledge as a Commons: From Theory to Practice*, ed. Charlotte Hess and Elinor Ostrom, 41–81. Cambridge, MA: MIT.
Ott, Brian L. 2017. "The Age of Twitter: Donald J. Trump and the Politics of Debasement." *Critical Studies in Media Communication* 34 (1): 59–68.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

Picard, Robert G. 2015. "The Humanisation of Media? Social Media and the Reformation of Communication." *Communication Research and Practice* 1 (1): 32–41.

Riedl, Martin J., Kelsey N. Whipple, and Ryan Wallace. 2021. "Antecedents of Support for Social Media Content Moderation and Platform Regulation: The Role of Presumed Effects on Self and Others." *Information, Communication and Society* 25 (11): 1632–1649.

Roberts, Sarah T. 2018. "Digital Detritus: 'Error' and the Logic of Opacity in Social Media Content Moderation." *First Monday* (online) 23 (3). https://journals.uic.edu/ojs/index.php/fm/article/view/8283.

Robertson, Adi. 2022. "Supreme Court Blocks Texas Social Media Moderation Ban." *The Verge*. www.theverge.com/2022/5/31/23124281/supreme-court-texas-hb20-social-media-ban-response.

Roth, Yoel. 2022. "Introducing Our Crisis Misinformation Policy." Twitter. https://blog.twitter.com/en_us/topics/company/2022/introducing-our-crisis-misinformation-policy.

Roth, Yoel, and Nick Pickles. 2020. "Updating Our Approach to Misleading Information." Twitter. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.

Sandoval-Almazan, Rodrigo, and J. Ramon Gil-Garcia. 2014. "Towards Cyberactivism 2.0? Understanding the Use of Social Media and Other Information Technologies for Political Activism and Social Movements." *Government Information Quarterly* 31 (3): 365–378.

Sankaranarayanan, Jagan, Hanan Samet, Benjamin E. Teitler, Michael D. Lieberman, and Jon Sperling. 2009. "Twitterstand: News in Tweets." In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 42–51. New York: ACM.

Shao, Chengcheng, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. 2018. "The Spread of Low-Credibility Content by Social Bots." *Nature Communications* 9: 4787. https://doi.org/10.1038/s41467-018-06930-7.

Shin, Jieun, Lian Jian, Kevin Driscoll, and François Bar. 2018. "The Diffusion of Misinformation on Social Media: Temporal Pattern, Message, and Source." *Computers in Human Behavior* 83: 278–287.

Silva, Leandro, Mainack Mondal, Denzil Correa, Fabrício Benevenuto, and Ingmar Weber. 2016. "Analyzing the Targets of Hate in Online Social Media." In *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*, 687–690. Palo Alto: AAAI.

Simons, Rachel Noelle. 2015. "Addressing Gender-Based Harassment in Social Media: A Call to Action." *iConference 2015 Proceedings*. http://hdl.handle.net/2142/73743.

Sin, Sei-Ching Joanna. 2016. "Social Media and Problematic Everyday Life Information-Seeking Outcomes: Differences across Use Frequency, Gender, and Problem-Solving Styles." *Journal of the Association for Information Science and Technology* 67 (8): 1793–1807.

Sin, Sei-Ching Joanna, and Kyung-Sun Kim. 2014. "Impacts of Social Media Usage on the Outcomes of Students' Everyday Life Information Seeking." In *Proceedings of the American Society for Information Science and Technology*, 1–4. Silver Spring, MD: Association for Information Science and Technology.

Smith, Scott M., Catherine A. Roster, Linda L. Golden, and Gerald S. Albaum. 2016. "A Multi-Group Analysis of Online Survey Respondent Data Quality: Comparing a Regular USA Consumer Panel to MTurk Samples." *Journal of Business Research* 69 (8): 3139–3148.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010

Statista. 2022. "Number of Twitter Users Worldwide from 2019 to 2024." www.statista.com/statistics/303681/twitter-users-worldwide/.
Suzor, Nicolas, Molly Dragiewicz, Bridget Harris, Rosalie Gillett, Jean Burgess, and Tess Van Geelen. 2019. "Human Rights by Design: The Responsibilities of Social Media Platforms to Address Gender-Based Violence Online." *Policy and Internet* 11 (1): 84–103.
Taddicken, Monika. 2014. "The 'Privacy Paradox' in the Social Web: The Impact of Privacy Concerns, Individual Characteristics, and the Perceived Social Relevance on Different Forms of Self-Disclosure." *Journal of Computer-Mediated Communication* 19 (2): 248–273.
Tifferet, Sigal. 2019. "Gender Differences in Privacy Tendencies on Social Network Sites: A Meta-Analysis." *Computers in Human Behavior* 93: 1–12.
Twitter. 2022. "Terms of Service." https://twitter.com/en/tos.
United States Communications Decency Act. 1996. "U.S. Code 47 (1996), § 230." https://uscode.house.gov/view.xhtml?req = (title:47%20section:230% 20edition:prelim).
Vogels, Emily A. 2022. "Support for More Regulation of Tech Companies Has Declined in U.S., Especially among Republicans." Pew Research Center. https://pewrsr.ch/3l8NYEs.
Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018. "The Spread of True and False News Online." *Science* 359 (6380): 1146–1151.
Wakabayashi, Daisuke. 2019. "Legal Shield for Websites Rattles under Onslaught of Hate Speech." *The New York Times*. www.nytimes.com/2019/08/06/technology/section-230-hate-speech.html.
Walker, Mason, and Katerina Eva Matsa. 2021. "News Consumption across Social Media in 2021." Pew Research Center. www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/.
Wardle, Claire, and Eric Singerman. 2021. "Too Little, Too Late: Social Media Companies' Failure to Tackle Vaccine Misinformation Poses a Real Threat." *British Medical Journal* 372 26: 1–3.
Waterloo, Sophie F., Susanne E. Baumgartner, Jochen Peter, and Patti M. Valkenburg. 2017. "Norms of Online Expressions of Emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp." *New Media and Society* 20 (5): 1813–1831.
Wathen, C. Nadine, and Jacquelyn Burkell. 2002. "Believe It or Not: Factors Influencing Credibility on the Web." *Journal of the American Society for Information Science and Technology* 53 (2): 134–144.
Whittaker, Elizabeth, and Robin M. Kowalski. 2015. "Cyberbullying Via Social Media." *Journal of School Violence* 14 (1): 11–29.
Winkelman, Sloane Burke, Jody Oomen-Early, Ashley D. Walker, Lawrence Chu, and Alice Yick-Flanagan. 2015. "Exploring Cyber Harassment among Women Who Use Social Media." *Universal Journal of Public Health* 3 (5): 194–201.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 06:11:59, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009255165.010