

Social Media Influencer and Source Credibility: Endorsing Content Moderation on Douyin

RENYI HE

The Chinese University of Hong Kong, China

HAO TIAN

Fudan University, China

Content moderation (CM) is widely employed by social media platforms to curb false information, but it often leads to a form of censorship that hampers discussions and lowers the accountability of CM. Short-video platform Douyin made a novel effort to enhance its accountability by launching an influencer named *Meiloufeng*, who introduces CM policies and interacts with users through entertaining short videos. The purpose of Meiloufeng is to enhance his credibility as the source of CM to increase the accountability of Douyin. The analysis found that being both a human source and an institutional source jointly shaped the construction of Meiloufeng's source credibility. This study also revealed that, rather than addressing the need to enhance the degree of transparency, the introduction of Meiloufeng raised a more important question: How can the degree of transparency be explored to better inform the public and be accepted by the platforms?

Keywords: social media influencer, source credibility, human source, institutional source, content moderation, Douyin, transparency, accountability

In recent years, social media platforms have been exerting more efforts into monitoring and regulating user-generated content to curb the spread of false and harmful information (Caplan, 2018). This organized practice of monitoring and screening content is called content moderation (CM; Roberts, 2019). It is used to provide a healthy digital space for civic participation, but the deletion of user-generated content can lead up to a form of censorship that suppresses civic participation (Helberger, Pierson, & Poell, 2018) and further trigger strong backlash from dissatisfied users. Frequent complaints include no disclosure of details about how they violated community guidelines, no information about the moderation process, and no opportunities for human interaction to save their content and accounts (West, 2018). The lack of reliable information leads users to develop their own explanations, which breeds distrust of platforms and ferments conspiracy theories that further damage platform-user relations (Suzor, Van Geelen, & Myers West, 2018).

Renyi He: renyihe@link.cuhk.edu.hk

Hao Tian: tianhao@fudan.edu.cn

Date submitted: 2023-01-03

Copyright © 2023 (Renyi He and Hao Tian). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <http://ijoc.org>.

Accordingly, major digital platforms have been working together with governments and users to find feasible methods to increase users' accountability of CM, such as joint institutions of government and platform, volunteer moderators, and more timely updates of the moderation process (Crawford & Gillespie, 2016; Gorwa, 2019a; Seering, Wang, Yoon, & Kaufman, 2019). Among these efforts, Douyin, one of the most popular short-video platforms in China whose mother company also owns TikTok, set up an account for its moderation team Douyin Safety Center (抖音安全中心) in 2017 and launched a character called *Meiloufeng* (梅楼封, which is a homophone of 没楼封, meaning no social media post to block) in 2021 to promote CM. Through short videos, Meiloufeng explains the moderation rules, introduces moderation functions, displays accounts and content that violated community rules, and elicits user feedback. He also mimics trending memes, participates in popular challenges, and makes entertaining content just like other influencers. Within 10 days of launching, Meiloufeng soon attracted a large audience, and the account's number of followers quickly increased to 5.3 million. Different from other platforms that have tried to improve or renew existing moderation systems or mechanisms, the performance of Meiloufeng is a new direction to promote CM, inform users, and cultivate platform-user relations. What does Meiloufeng do to promote CM in the short videos? How does he construct his credibility? Most importantly, what can this new direction tell us in terms of enhancing the accountability of CM? These questions are worth investigating.

For its 8 million followers (and counting), Meiloufeng has become an important source of Douyin's CM, and what it has been doing can be perceived as efforts to increase his credibility as such a source. For that reason, this study adopted the framework of source credibility theory that has been widely used in the analyses of celebrity endorsement (e.g., Djafarova & Rushworth, 2017; Jin, Muqaddam, & Ryu, 2019) to explore how Meiloufeng, as a social media influencer, endorses CM by producing informative and entertaining short videos.

Increasing User Acceptance of Content Moderation: A Core but Tough Issue

Social media users are dissatisfied about current CM practices. In the United States, only 31% of adults have some confidence in the capability of digital platforms to moderate content (Laloggia, 2019). There is also a strong dissatisfaction toward moderation in China because it is not only political information that is heavily moderated but also content like smoking and tattoos (Jing, 2019). Users do not only contest specific moderation decisions but more fundamental issues as well, such as the ultimate goal of CM and the inconsistency of the moderation system. An essential problem that causes user dissatisfaction is that platforms are still struggling to moderate contentious content without obstructing public discussions (Helberger et al., 2018). Existing CM systems encounter difficulty in determining what is "good" and distinguishing "bad" from "good." For instance, the moderation of nudity is helpful to control erotic and sexual content, but it also leads to overblocks of breastfeeding pictures that can be used to promote gender education and convert gender discrimination (Gillespie, 2018). Similarly, the moderation of violent and abusive content aims to avoid overexposure to violence that may cause mental health problems, but it also stops the use of social media to preserve evidence of mass shootings and sexual abuses (Banchik, 2021).

A major concern to enhance the accountability of CM is to increase the level of transparency because users who are well informed can better understand moderation, lower their antagonism toward platforms, reduce future postremovals, and contribute to improving existing CM practices (Fung, Graham, & Weil, 2007; Suzor, West, Quodling, & York, 2019). A healthy digital sphere relies on joint efforts from platforms and user communities (Squirrell, 2019). Enhancing the transparency of CM can prepare users to become productive community members. If there is insufficient reliable information, then users may develop their own "folk theories" and "vernacular rationalizations," which would breed distrust between users and platforms and lower the accountability of platforms (Suzor et al., 2019; West, 2018).

However, CM remains an unknown process for social media users. Existing CM relies heavily on (1) algorithms that aim to match content to existing data or predict content as belonging to existing categories and (2) outsourced laborers who manually assess content flagged by users or unrecognized by algorithms to check whether it conforms to the platform community guidelines (Gorwa, Binns, & Katzenbach, 2020; Roberts, 2019). Owing to the claimed necessity to protect their business and profits from being exploited by competitors, platforms have released very little information about their core commodities and CM technology and prohibit outsourced labor from revealing any procedural information (Gillespie, 2018; Roberts, 2019).

Another approach to enhance the accountability of CM is to renew existing moderation mechanisms to incorporate social media users and other institutions. CM is more than a regulatory issue of digital platforms but also an essential subject of public interest. Therefore, to improve existing moderation, we also need to incorporate the public into the design of the CM system (Gorwa, 2019b). Some scholars proposed building "cooperative responsibility" between platforms and users to ensure that users are also aware of their responsibility to protect the public space (Helberger et al., 2018). Some called for enhanced "platform dialectic" between the moderation team and users to construct a shared understanding of CM (Squirrell, 2019). Still others argued for building a "governance triangle," incorporating states and nongovernment organizations (NGOs) in the policy-making process (Gorwa, 2019a). The essential concern of these designs is to form a mechanism that can continuously improve CM by incorporating users into the rulemaking process. These efforts promote principles like rules of law to construct governance norms, aim to build new standards of governance legitimacy, and try to constrain the exercise of power, which is called "digital constitutionalism" (Suzor, 2018; Suzor et al., 2018).

With similar concerns, digital platforms also tried to increase platform-user collaborations to form a long-standing mechanism that can enhance transparency. However, very few practices can fulfill the goal. For instance, the Twitter Trust and Safety Council has brought together user community groups, NGOs, and a handful of academics, but there was almost no public information about how they contribute to improving the CM products and policies of Twitter (Tusikov, 2017). The Facebook Oversight Body has similar problems. It aims to let users appeal takedown decisions to an independent body that includes user representatives, but legal theorists believe that it would transform the company into something resembling a governmental or judicial entity (Kadri & Klonick, 2019).

The launch of Meiloufeng by Douyin Safety Center provides us a novel effort to improve the accountability of CM by enhancing transparency and increasing platform-user collaborations. Meiloufeng

performs moderation procedures, displays the working environment, and shares the difficulties of designing a CM policy. He also interacts with users and encourages users to contribute to CM policy making through the various functions of Douyin. Indeed, Meiloufeng does not reveal everything about CM to the audience, and the audience—especially those who have suffered from CM—are not easily persuaded by Meiloufeng’s endorsement. Nonetheless, as our analysis would show, this new effort has great potential to enhance the accountability of CM.

Content Moderation and Censorship in China

The difference between moderation and censorship should be made clear since the serious challenge for CM is to filter information without becoming a form of censorship (Helberger et al., 2018). The distinction is even more important in the Chinese context, given that censorship is extensively used to screen online discussions. By combining Sarah Roberts’ (2019) definition of online commercial CM and James Grimmelmann’s (2015) definition of moderation, we define content moderation as the organized governing practice of social media platforms to facilitate communication and prevent abuse. The goal of CM is to curb the influence of toxic speech, terrorist content, and copyright violation (Gorwa et al., 2020), while censorship is often taken as an information control strategy that aims to sustain political control (Roberts, 2018). The moderation of social media content in China often includes both CM and censorship. This is because social media companies are driven by the regulatory purpose to offer a better communication environment for their users and by the political pressure from the party-state to suppress dissent, given that they are legally licensed by the party-state to operate in China. This study primarily focuses on the former, namely, the CM of Douyin.

Like their western counterparts, Chinese social media platforms also rely on algorithms and outsourced labor to remove, filter, and suspend content, and their moderation process lacks transparency (Zeng, Chan, & Fu, 2017). Moreover, the moderators work overtime at the cost of their health (Luo, 2022). Existing studies have shown two interesting strategies adopted by Chinese platforms to moderate content. The first strategy is “open refuting,” that is, openly marking misinformation and sharing them to the public (Zeng et al., 2017). Platforms may tag content with labels like rumors or unverified information, and their official accounts would announce and promote their decisions to the public. The second strategy is to introduce and endorse certain “positive” trends to secure visibility from other content. For instance, Douyin has launched and propagated a nationwide “positive energy” campaign to counterbalance erotic and sexual content with entertaining patriotic short videos (Chen, Valdovinos Kaye, & Zeng, 2021). The campaign aims to direct content-creation practices, nudge users toward “positive” trends, and seize online visibility (Zeng & Kaye, 2022). Launching Meiloufeng can be regarded as a combination of these strategies. On the one hand, Meiloufeng is the official voice that labels controversial content and displays them to the audience. On the other hand, the character itself, being entertaining and positive, is part of the “positive energy” campaigns that are advocated by the platform.

Source Credibility and Endorsement of Social Media Influencer

Hiring celebrities to promote products and services is a common practice in business because the credibility of a celebrity endorser can positively influence the credibility of the endorsed brand, which

can, in turn, increase product sales and gain an edge with competitors (Elberse & Verleun, 2012; Spry, Pappu, & Cornwell, 2011). A huge rise in a new type of digital celebrity—the social media influencer—has emerged in celebrity endorsement in recent years. Social media influencers are individuals who become famous through their social media presence (Hearn & Schoenhoff, 2016; Khamis, Ang, & Welling, 2017). The uniqueness of social media influencers is that they build their personal brands through constant digital interactions with their followers (Khamis et al., 2017), as opposed to traditional celebrities who use social media as an instrument for their work in the traditional media industry (Jin et al., 2019). To construct and maintain their personal brands, social media influencers always engage in quasipromotional activities that exhibit desirable attributes, such as expertise, enthusiasm, credibility, or social connectivity (Duffy & Hund, 2015). As a result, social media users often regard social media influencers as more credible than traditional celebrities and feel more connected to them (Djafarova & Rushworth, 2017; Tran & Strutton, 2014).

Meiloufeng can be regarded as a social media influencer. In the short videos, Meiloufeng performs as the owner of the account and the leader of the Douyin moderation team. He exhibits his expertise and credibility as a moderator, his enthusiasm to interact with users, and his caring for the safety of his followers' accounts. More importantly, he attracts a group of followers that like him and are interested in learning CM. Since his launch, Meiloufeng's number of followers has increased to 8.1 million, and videos with him also receive more likes. For the 100 videos before Meiloufeng, the average number of likes is 21,000, while for the 100 videos after Meiloufeng, the number of likes is 256,000, which is over 12 times larger. His connectedness with his followers makes him a popular social media influencer.

With strong credibility among their followers, social media influencers become powerful endorsers. An increasing number of companies and advertisers hire influencers to endorse products or services on social media (Hearn & Schoenhoff, 2016), and they become the major source of messages to their followers. Taking celebrity endorsers as the message source, scholars have been using source credibility theory to examine how they influence the attitudes and behaviors of the audience (Pornpitakpan, 2004). Source credibility describes the positive characteristics of a communicator that affects the receiver's acceptance of a message. If an endorser has a high-source credibility, then the receiver would be willing to accept the messages from the endorser. Roobina Ohanian (1990) proposed a framework that includes expertise, trustworthiness, and attractiveness to examine source credibility. Expertise refers to the level of experience, knowledge, and skills. Attractiveness is a multidimensional construct that includes physical attractiveness, chicness, sexiness, likability, and various others (McGuire, 1985). Trustworthiness shows how honest, reliable, and dependable the source is in the eyes of the audience (Ohanian, 1990). If social media influencers display a high level of expertise, attractiveness, or trustworthiness, then they would gain a high degree of source credibility. This framework has become a classic frame to test source credibility and a basis of many models that analyze the endorsement of social media influencers (Munnukka, Uusitalo, & Toivonen, 2016; Weismueller, Harrigan, Wang, & Soutar, 2020). The present study also used source credibility theory to examine Meiloufeng. Our first research question is:

RQ1: How did Meiloufeng construct his expertise, attractiveness, and trustworthiness?

Previous scholarship on source credibility takes social media influencers as human sources of messages, though they can also be institutional sources, especially when they are publicly associated with other institutions. Meiloufeng should be taken as such an institutional source, as he is not an outsider who is temporarily hired to produce videos that endorse some products or services. He is an insider, an integral part of Douyin. Given that Meiloufeng is not a self-made social media celebrity but more like a “mascot” of Douyin, he is not only a human source but an institutional source as well. Therefore, the second research question is:

RQ2: How did being both a human source and an institutional source jointly shape Meiloufeng’s source credibility?

Method

This study employed digital ethnography and qualitative content analysis to analyze the short videos of Douyin Safety Center that feature Meiloufeng.

The authors first conducted a six-month digital ethnography from March to September 2022. Digital ethnography is a mediated form of ethnography that has been used to analyze celebrities, influencers, and short videos of Douyin and TikTok (Abidin, 2020; Wang & Ding, 2022). Both individual videos and the trends on Douyin often fade away or are replaced by new ones in a short period of time. Digital ethnography can help researchers follow the constantly changing cultural background of memes, slangs, and trends that are widely shared by users in the short-video platforms (Schellewald, 2021). During their digital fieldwork, whenever Douyin Safety Center posted a video, the authors were notified by a push notification. They would then watch the videos and use Douyin for an hour in the next three days to (1) observe the interactions between users and Meiloufeng in the comments, (2) follow broader discussions on the topics that were mentioned by Meiloufeng, and (3) search the trendy memes and slangs appropriated by Meiloufeng. During the fieldwork, the authors also watched two live streams of Meiloufeng, whose content was also included in the collected videos.

Second, the authors collected 102 videos with Meiloufeng as the leading figure and used qualitative content analysis to analyze these videos, which allowed them to focus on content features that are related to the construction of source credibility and take the general culture of Douyin into analysis. Qualitative content analysis allowed the researchers to examine the themes and patterns of videos that are related to CM and source credibility. The two authors first familiarized themselves with the gathered data by watching the videos and taking notes, and then generated initial coding schemes based on their notes and the notes taken during the digital ethnography. The item “source credibility” was designed to answer RQ1, and the item “source type” was designed to answer “RQ2.” The two authors coded 20 randomly selected videos, reorganized existing themes, and refined the code book. For instance, video contents with “CM policy” and “CM decision” were merged into “policy and decision.” Next, the first author coded the rest of the videos. The categories of “source credibility” and “content” are mutually exclusive, while the categories of “source type” are non-mutually exclusive because Meiloufeng showed up as both a person and a representative of Douyin in many of the videos. Table 1 shows the result.

Table 1. Content Analysis of Videos.

Category	Explanation	Count
Source credibility		
Expertise	Meiloufeng shows up as a CM expert.	50 (49.0%)
Attractiveness	Meiloufeng dresses like funny ancient police, playing the role of a CM authority and an entertaining influencer.	36 (35.3%)
Trustworthiness	Meiloufeng displays his work routine and work pressure as an ordinary employee.	16 (15.7%)
Source type		
Institutional source	Meiloufeng behaves as part of the Douyin CM authority.	79 (77.6%)
Human source	Meiloufeng acts like a real person who endorses CM.	58 (56.9%)
Content		
Policy and decision	Introduce and explain CM policy and decision	35 (34.3%)
Comedy	Feature trending meme and hilarious challenges on Douyin	21 (20.6%)
Interactive content	Interact with audience and other influencers	17 (16.7%)
Daily work	Reenact moderation process or display working settings	14 (13.7%)
Personal emotion	Display Meiloufeng's personal feelings	8 (7.8%)

To determine how users perceived Meiloufeng, the authors also collected the top 20 liked comments of each video and coded their attitudes toward Meiloufeng. After comments that could not express a meaningful message were excluded, a total of 1,917 comments were obtained. The authors randomly selected and coded 10% (N = 192) of the comments and tested the reliability. Since the reliability (Krippendorff's alpha = 0.86) was high, each author coded half of the comments. Table 2 shows the result.

Table 2. Attitudes of User Comments.

Attitude	Explanation	Count
Support	Support Meiloufeng or CM; positive expressions	1,198 (62.5%)
Oppose	Question Meiloufeng or CM; negative expressions	270 (14.1%)
Neutral	Seek for help, advice, or information about CM	449 (23.4%)

Results

Meiloufeng is a mediated persona that acts like a real person, as well as a mechanism and a procedural of Douyin. No matter who plays the role or what name it has, the function of Meiloufeng in the CM process remains the same: to publicly promote CM and Douyin. In fact, the staff who had been playing the role of Meiloufeng transferred to another department in November 2022, and Douyin started to find a new Meiloufeng. Meiloufeng should be regarded as a personalized CM institution, a combination of a human source and an institutional source. Accordingly, the following analysis will explore how Meiloufeng constructed his source credibility (RQ1) and how being a human source and an institutional source jointly shaped Meiloufeng's source credibility (RQ2).

The way that Meiloufeng's source credibility was constructed is different from that of many social media influencers, since his credibility is inherently associated with the platform he belongs to. The source credibility of Meiloufeng has a lot more to do with the institutional trust and institutional culture of Douyin. First, the institutional trust in Douyin's CM played an important role in Meiloufeng's expertise. The accountability of CM is generally low among the audience because they have no way to be part of the moderation process. Some Douyin users who have experienced moderation and have not watched the videos are resistant to Meiloufeng. Merely displaying the expertise of CM may not help Meiloufeng to increase his source credibility but may instead push the audience away because it strengthens its role as an institutional source. Instead, Meiloufeng tried to enhance its expertise through being an open-minded expert who not only holds CM knowledge but also interacts with users and expresses the will to learn from users. It is a method that was designed to cope with the low CM institutional trust among audiences.

Second, the key to constructing the attractiveness of Meiloufeng lies in the institutional culture of Douyin. The culture of Douyin and its international sister TikTok emphasizes fun and playfulness (see, e.g., Schellewald, 2021; Zeng, Schäfer, & Allgaier, 2021). In addition to the dominant entertaining videos on the "For You Page," being fun and silly is also a key element of content that aims to promote political campaigns in the United States (Literat & Kligler-Vilenchik, 2019) and China (Chen et al., 2021). Meiloufeng is not only the character that represents the CM of Douyin that frequently deletes user-generated content but also serves as a content creator on Douyin, responsible for creating entertaining videos that fit the cultural trends on the platform. The tension caused by Meiloufeng's dual identity as both an authoritative representation of Douyin's CM and a funny social media influencer to promote CM makes him an attractive figure that audiences can make fun of, which increases his credibility.

Third, Meiloufeng tried to interact with the audience as a real person to enhance his trustworthiness and counterbalance the low institutional accountability of Douyin, which, in turn, lowers Meiloufeng's credibility. Accordingly, he employed a strategy to separate his role as a human source from being an institutional source. In some videos, Meiloufeng attempted to get rid of his role as a moderator and the representation of Douyin CM by performing as an ordinary employee who suffers from the toxic working culture just like his audience. Shared feelings about the pressure from work can increase the homophily between him and other users, which can further enhance the trustworthiness dimension of Meiloufeng's source credibility (Jin, 2018).

Interacting With Users as an Open-Minded Expert

A major complaint from users who have experienced CM is the lack of human interaction in the negotiation process (West, 2018). The launch of Douyin's own social media influencer, Meiloufeng, can be taken as an effort to make human actors stand out in the moderation process. It is a strategy to emphasize the human source aspect of Meiloufeng and build closer connections between the institution and its users. In about half of the collected videos (49%), Meiloufeng shows up with daily wear, playing the role of a Douyin moderator who is an open-minded expert. Meiloufeng interacts with users by answering user questions about CM and asking users to share their ideas about CM to improve the rulemaking.

The interactions between Meiloufeng and users are not real interpersonal communications but more like parasocial interactions—interpersonal involvements that give users a sense of friendship or intimacy with the mediated personas (Horton & Wohl, 1956). Although they are described as illusionary experiences of audiences, parasocial interactions can enhance their feelings of connectedness, leading to a perceived close relationship with celebrities. Over time, viewers will see the mediated persona, in our case Meiloufeng, as a trusted source of CM information and seek out their advice (Rubin, Perse, & Powell, 1985). Parasocial interaction, a form of one-to-all communication that has the appearance of personal communication, is a low-cost and controllable way to increase the source credibility of Meiloufeng.

The first type of parasocial interaction is a series of Q&A videos called “How the Official Replies to Comments With Talents,” (e.g., Douyin Safety Center, 2021a) in which Meiloufeng and his colleagues pick and answer some questions from users (see Figure 1). For social media influencers, answering questions is an easy and popular way to interact with fans by giving them a sense of involvement. Meiloufeng has tried to build a sense of equality between users and the moderation team to promote user-platform dialectics, which may become a base for potential platform-user collaborations (Squirrell, 2019). Nevertheless, Meiloufeng has control of the parasocial interactions because he can choose which questions to answer and how to answer them. For instance, an influencer named Feng Xiaoyi was permanently banned from Douyin in August 2021. Many believe that it was because of the firm stance of the Chinese government against “sissy” male celebrities, as Feng resurfaced with a new and more masculine image three months later (Ziyi, 2021). In one Q&A video (see Figure 1), Meiloufeng avoided mentioning the gender issue that may trigger criticism of the party-state’s policies but reinforced that Feng solicited minors to donate money, an illegal conduct that broke the rules of the Cyberspace Affairs Commission (Jiang, 2022). The parasocial interactions enabled Meiloufeng to performatively address user concerns about CM transparency and follow the guidance of the political authority at the same time.



Figure 1. How the official replies to comments with talents (Douyin Safety Center, 2021a, 00:14, 00:20).

Violations of laws and regulations, such as fraud, child enticement, pornography, and extreme speech, are plausible justifications of CM. These are also the proclaimed goals of the Chinese party-state’s

Internet policies that are often used to eliminate dissents. Users may not accept this kind of “safe” explanation from an institutional source. In the Q&A videos, 14.3% of the comments questioned Meiloufeng’s explanations. Some users regarded Meiloufeng as a “puppet expert” who only endorses the rules of the platforms and the government. Others realized that Meiloufeng just enhances the transparency of CM at the surface level because the character is just part of the platform and must follow the institutional rules. To cope with the low institutional accountability, Meiloufeng emphasized his role as a human source through the second type of parasocial interaction.

The second type of parasocial interaction is videos that encourage users to participate in CM discussions and express their thoughts on how to regulate controversial content. These interactions display the human source of CM that positions moderators as open-minded experts who are willing to listen to users. In June 2021, Douyin Safety Center (2021c) launched a campaign called “I Have Something to Say About Community Rules.” The video enacted a routine inner meeting where different moderators raise different opinions over controversial social media content (see Figure 2). As the moderators cannot reach any agreement, they ask the social media users for help. Users can click the link below the videos to the official webpage of the campaign where they can send their suggestions to Douyin Safety Center. In August 2022, a similar campaign that also asked for users’ help, called “My Suggestion for Douyin Safety Center,” was launched to help make CM rules and policies (Douyin Safety Center, 2022c).

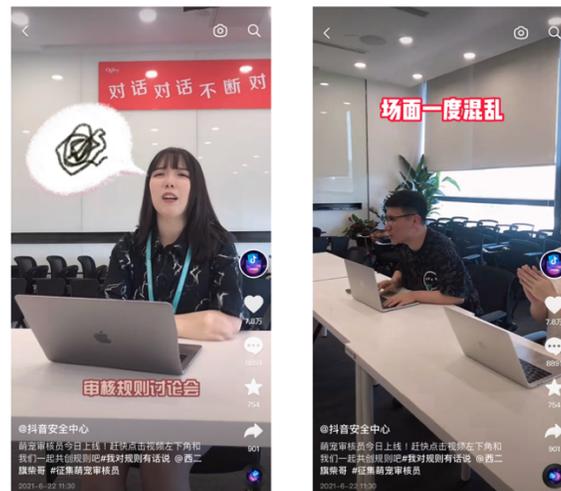


Figure 2. I have something to say about community rules (Douyin Safety Center, 2021c, 00:04, 00:15).

Unlike the top-down interactions where Meiloufeng is an institutional source answering user questions, Meiloufeng and his colleagues in these videos appear as human sources of CM. They are individual moderators who find it very difficult to reach an agreement on moderation decisions, so they seek help from users. Seeking for help is more like a peer-to-peer interaction that aims to build a positive atmosphere for discussion. In the discussion pages of the campaign, users expressed and tried to justify their own opinions over the criteria of CM, some of which were contradictory, but all tried to contribute to the discussion. There

is no high-above institution in the open discussions. Everyone joined in and contributed equally to the CM rulemaking. Although we cannot draw the conclusion that the promise to incorporate user discussions into CM would not be an empty one from existing information, users respond more positively to this second type of interaction where Meiloufeng shows up as a human source (1.7% of comments are negative) compared with the first one where he performs as an institutional source (14.3% are negative).

Last, it is also worth noting that Meiloufeng only responded to a few of the comments that asked for more information about CM. After all, the interactions were one-sided parasocial interactions between an institutional source and users (Labrecque, 2014). These offered a guaranteed way for Douyin to deal with user feedback but cannot build real reciprocal relationships that fully address user concerns about CM.

Enhancing Attractiveness as a Funny Ancient Policeman

Apart from exhibiting his expertise, Meiloufeng also shows up as a funny ancient Chinese policeman to increase his attractiveness (35.3%). He is both an authoritative voice of Douyin's CM and an influencer who explains CM in an entertaining and hilarious way (see Figure 3). Our analysis shows that the attractiveness of Meiloufeng mainly comes from the tension between his authoritative role to explain and promote CM policies and the entertaining role that he plays as a social media influencer in a platform where being funny and silly is the key to its culture.

The character of Meiloufeng as an ancient Chinese police officer first constructs his legitimacy as a CM authority. The major responsibility of a police officer is to capture criminals, while for Meiloufeng, his major task is to find the Douyin account, who has broken the community rules. In most of the videos, he speaks with an authoritative voice to inform the audience why some contents are deleted and some accounts are blocked. The commonly used opening remark in his videos is "Today, let's take a look at who is going to be banned" (Douyin Safety Center, 2022d, 00:03; see left screenshot in Figure 3).



Figure 3. Meiloufeng performs as a funny ancient policeman (Douyin Safety Center, 2022d, 00:03; 2022a, 00:05).

However, being an authority does not make Meiloufeng appealing to the audience but even pushes the audience away. The staff also showed the funny part of Meiloufeng when he played the role. He tried very hard to put on a serious face when playing the role but often failed either because of his own funny behaviors or being teased by his colleagues (Douyin Safety Center, 2022a, 00:05; see the right screenshot in Figure 3). Besides, Meiloufeng has a strong Dongbei accent, which is regarded as one of the funniest dialects in China with “a unique ability to make even the most mundane daily conversation sound comical” (Huang, 2015, para. 3). The strong contrast between the serious official announcement and the entertaining ways to present the announcement to the audience helped make Meiloufeng seem more like a real person who can be made fun of. One comment wrote, “What you have said is so funny, unlike an official staff. I laughed with tears” (personal communication, October 20, 2022). Overall, this type of video has 84.3% positive comments. It is the contrast between being an institutional source and a human source that makes Meiloufeng an entertaining figure, which enhances his attractiveness.

Attractiveness of the social media influencers plays a key role in increasing their source credibility (Lee & Watkins, 2016). In the case of Meiloufeng, being entertaining helps to make him more like a real person who can be made fun of. His attractiveness builds a casual atmosphere where the audience feels free to talk about CM with him. In a live stream that introduced a newly launched CM function, Meiloufeng sent a fraud message to his friends to demonstrate how the new algorithms can automatically detect and block harmful information. As a punishment mechanism, the message was detected and blocked, and the sender’s account, which was the official account of Douyin Safety Center, was suspended, leading to an abrupt stop of the live stream. Audience members commented with laughter and mockeries about Meiloufeng because he screwed up the live stream. Nonetheless, Meiloufeng successfully accomplished his job because the audience was now aware of the efficiency of the algorithm and the outcome of sending misinformation. Soon, users spread the news that the Douyin moderation team moderated themselves, attracting more people to pay attention to how the CM system works. Afterward, when the official account was back, Meiloufeng made a video explaining the whole process. One comment mocked, “Thanks to your performances, I now know how to break the regulations” (personal communication, October 21, 2022).

Building Intimate Relation With Audience as an Ordinary Staff

To enhance his trustworthiness, Meiloufeng showed up as an ordinary employee of Douyin and tried to build intimate connections with his followers as an ordinary staff member (15.7%). In these short videos, he performed the daily routine of his work, a direct way to disclose CM-related information that enhanced the level of transparency. He also complained about the huge pressure at work, through which he personalized CM as someone’s real-life daily work. Notably, this strategy distanced Meiloufeng from being the source of CM, whether the human source or the institutional source. Through personal emotion disclosure, Meiloufeng tried to leave the audience with an impression of an honest and sincere person instead of an endorser of something else, which eventually enhanced his trustworthiness as a source of CM.

First, Meiloufeng put enormous effort in performing the real work of a moderator. A series of videos displays the work of moderators, flagging handlers, and Douyin staff. In these videos, Meiloufeng and his colleagues first showed the working status conceived by the audience and then broke the perception with real work settings. In the four episodes of “What you Thought Were Short-Video Moderators” (Douyin Safety

Center, 2021b), one major theme is that moderators, instead of enjoying the power of deciding the life-or-death of users' contents and accounts (see the left screenshot in Figure 4), follow the rules and regulations of the platforms and are often misunderstood by the users (see the right screenshot in Figure 4). In these videos, Meiloufeng showed up as an employee who just wanted to fulfill his work duties.



Figure 4. What you thought were short-video moderators (Douyin Safety Center, 2021b, 00:03, 00:21).

These presentations of real working conditions are performed by staff members in front of the camera, but they still help to refute the unrealistic guesses from the audience that are caused by the nontransparency of the working conditions and process of moderators (Roberts, 2019; West, 2018). Many users flirted with Meiloufeng with comments that stuck to the imagined working status of moderators and expressed strong desires of working with Meiloufeng so that they can see the forbidden content. One said, "I also want to be a moderator. Salary is not a problem, just because of love." Another said, "This job suits me well. I am impartial and incorruptible since very young [smirk emoji]" (personal communication, October 21, 2022). From these comments, we can see that Meiloufeng has established a close connection with the audiences who would like to make fun of him. These comments also show that users understand the CM rules that are displayed in the videos. They just mocked Meiloufeng by contending that they wanted to see what they are not allowed to see.

Second, Meiloufeng also revealed his personal feelings about performing the role of Meiloufeng and being the source of CM. In these videos, he changed his entertaining style and sat in front of the camera with a serious face, giving a monologue about the huge pressure he felt and his mixed feelings about online curses he received. In one video, he said that he was fully aware that this was because of the nature of this job, but he still lost sleep at night and gained lots of weight (see the left screenshot in Figure 5). In another video, he said that he was trapped by the negative emotions from work and decided to make a change to stop overeating and start to exercise. He encouraged audiences who had similar issues to work out with him to keep their physical and mental health (see the right screenshot in Figure 5).



Figure 5. Videos that reveal Meiloufeng's personal feelings (Douyin Safety Center, 2022a, 00:05; 2022b, 00:05).

The huge work pressure of Meiloufeng resonated with audiences who also suffered from the current Chinese work culture where more than 80% of Chinese employees are overworked and under mental and physical stress (Yang, 2018). Working pressure is a heated topic, as can be seen in the popular catchphrase "996," referring to work from 9:00 a.m. to 9:00 p.m., six days a week. The prevailing 996 work culture in China was famously advocated by the tech tycoon Jack Ma as a blessing for employers but denounced by commentators as "toxic" (Kelly, 2021). The disclosure of personal pressure and emotional sufferings from Meiloufeng makes him part of the oppressed young working laborers, through which the emotional connectedness between users and Meiloufeng can be built. Although the CM of Douyin was not satisfactory for the audience, they could understand the hardship experienced by frontline staff like Meiloufeng. Just like one user said in the comments, "Understand it. He is just a staff member [with] no power at hand. He is also scolded by his boss all day long" (personal communication, October 21, 2022)

The personal emotion disclosure increases the homophily, the likeness between users and Meiloufeng, which tends to develop close relationships (Turner, 1993). The more a person sees him- or herself as similar to others, the more likely that he or she will interact frequently with the person, including the mediated persona on television and celebrities on social media (Lee & Watkins, 2016). The perceived homophily of Meiloufeng was high when he showed up as an employee who suffered from the same toxic working culture just like the audience. The tones of collected comments under these videos also showed support to Meiloufeng as they were all positive.

Nevertheless, the dominant support in the comments is for Meiloufeng, not for the CM of Douyin. From the comments, we can see that many users can tell the difference between Meiloufeng as a real person and Meiloufeng as an endorser of CM. They know that the character has no power and liked and supported the real person but not the platforms that remained powerful and silent behind. After all, Douyin is held by one of the largest corporations in China whose goal is to make profits. The launch of Meiloufeng is meant to improve existing CM practices and provide better moderation service to platform users as well as a performance that addresses user concerns at the surface level. The tension between the need for more

transparent CM from users and the goal to make profits is still there. Meiloufeng cannot solve the problems but just alleviate the tension to a limited extent.

Conclusion and Discussion: Transparency, Personalization, and Source Credibility

Through the analysis of Douyin's moderation influencer Meiloufeng, this study touched on a key issue of CM: the tension between public interest and profits of private sectors. CM, on the one hand, is a matter of public interest because it involves building a healthy and robust online environment for everyone. It is also a service provided by social media corporations whose primary goal is to make profits. To better inform the public, social media platforms should reveal information about their CM policies, decision-making procedures, and algorithms, though such information is also regarded as important company assets that are protected by private sectors. Our analysis shows that Meiloufeng has tried to alleviate the tensions between informing the public and protecting Douyin's own assets. He showed the real work settings and reenacted some moderation procedures but did not record the real working process of Douyin moderators. He displayed how the CM functions stop misinformation but did not involve the logic of algorithms. For many users who supported Meiloufeng, they also saw the difference and acknowledged his efforts. Therefore, instead of addressing the need to enhance the degree of transparency, a more important question for CM studies is to explore the degree of transparency that can better inform the public and be accepted by the platforms.

Second, this study enriched discussions on source credibility by exploring how being a human source and an institutional source jointly shape the construction of source credibility. Analyses of the three dimensions of source credibility displayed the three different roles the two sources played in source credibility construction. For instance, being an institutional source gave Meiloufeng a reputation of high expertise but low accountability. Accordingly, putting more effort into his role as a human source helped him to get away from the Douyin CM institution, whose trust is low. The tension between the image of being an institutional source as an authoritative police officer and that of being a human source as an entertaining influencer in Douyin made Meiloufeng an attractive figure who can be made fun of in an easy atmosphere. Finally, Meiloufeng tried to distance himself from being a source (both a human one and an institutional one) in the eyes of the audience, making himself more like a real person to enhance his trustworthiness.

Third, the popularity of Meiloufeng also manifests how people on social media connect with one another through highly individualized and personalized narratives. Douyin personified CM as an influencer and tried to make the image of the influencer more acceptable to enhance its CM accountability. As we have shown, Meiloufeng in the videos acted as an entertaining and hilarious social media influencer, an authoritative representation of Douyin, an experienced Douyin moderator, and a common employee of the tech giant. He communicated with the audience through personalized expression. Sometimes he was serious about the fake content, and sometimes he was emotional about his sufferings. These personalized messages were significantly different from official announcements such as moderation decision letters and community guidelines. It well suited the patterns of participation and mobilization in social media, where personalized expressions can connect individuals in social media networks.

Social media platforms are increasingly making investments aimed at better communication with their users to improve their CM systems. There are no easy solutions to raise the transparency of CM and the accountability of platforms to an ideal stage, while developing strategies that are based on the affordances and cultures of different platforms can better encourage users to learn CM, police themselves, and contribute their parts to building a better community. It is crucial that future research looks at the ways in which platforms featuring different types and forms of content communicate with users and how users respond to the CM policy and regulation of platforms.

References

- Abidin, C. (2020). Mapping internet celebrity on TikTok: Exploring attention economies and visibility labours. *Cultural Science Journal*, 12(1), 77–103. doi:10.5334/csci.140
- Banchik, A. V. (2021). Disappearing acts: Content moderation and emergent practices to preserve at-risk human rights-related content. *New Media & Society*, 23(6), 1527–1544. doi:10.1177/1461444820912724
- Caplan, R. (2018). *Content or context moderation? Artisanal, community-reliant, and industrial approaches*. Retrieved from https://datasociety.net/wp-content/uploads/2018/11/DS_Content_or_Context_Moderation.pdf
- Chen, X., Valdovinos Kaye, D. B., & Zeng, J. (2021). #PositiveEnergy Douyin: Constructing “playful patriotism” in a Chinese short-video application. *Chinese Journal of Communication*, 14(1), 97–117. doi:10.1080/17544750.2020.1761848
- Crawford, K., & Gillespie, T. (2016). What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society*, 18(3), 410–428. doi:10.1177/146144481454316
- Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities’ Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior*, 68, 1–7. doi:10.1016/j.chb.2016.11.009
- Douyin Safety Center. (2021a). *Di19ji bao yidingyao kandao zuihou yitiao dashen pinglun, wobei qide liangyan mao jinxing~* [The 19th episode Please watch until the last comment, I was angry~] [Video file]. Retrieved from <https://www.douyin.com/video/7002486049255542053?modeFrom=>
- Douyin Safety Center. (2021b). *Di3ji xiwang nimen jiedao women dianhua de shihou buyaogua, you xiwang nimen yongyuan jiebudao women de dianhua~ wuwu, bei women ziji gandong kule* [The third episode Hope you don’t hang up when you receive our calls, but also hope you will never get calls from us~ wuwu, we are moved by ourselves] [Video file]. Retrieved from <https://www.douyin.com/video/6976826558669507876?modeFrom=userPost&secUid=55456703574>

- Douyin Safety Center. (2021c). *Mengchong shenheyuan jinri shangxian! Gankuai dianji shipin zuoxiajiao he women yiqi gongchuang guize ba* [Cute pet moderator is online! Please click the link at the right corner to create rules together] [Video file]. Retrieved from <https://www.douyin.com/video/6976081129401109796?modeFrom=userPost&secUid=MS4wLjABAAAAY0vomT6bkKwbinBMqboF-bWq5RAou4YxOGojm3GS7PY>
- Douyin Safety Center. (2022a). *Dajia doushi zenme biaoda aidene?* [How do you express love?] [Video file]. Retrieved from <https://www.douyin.com/video/7088523323843349768?modeFrom=userPost&secUid=55456703574>
- Douyin Safety Center. (2022b). *Fangda fumian qingxu conglai jiu bushi yige haoxuanze, you he laomei yiqi daka jianfei dema?* [It is not a good choice to exaggerate negative emotion, and is there anyone who wants to lose weight with Meiloufeng?] [Video file]. Retrieved from <https://www.douyin.com/video/7096722441698921735?modeFrom=userPost&secUid=55456703574>
- Douyin Safety Center. (2022c). *Ganjue shuode dou geyou daoli a, tiezimen nimen shi zaxiang dene?* [Everyone has a good point, what do you think?] [Video file]. Retrieved from <https://www.douyin.com/video/7137240918138146085?modeFrom=userPost&secUid=MS4wLjABAAAAY0vomT6bkKwbinBMqboF-bWq5RAou4YxOGojm3GS7PY>
- Douyin Safety Center. (2022d). *Guonian dajia dou chibao hezu le, nalai pinpin zhejiwei doushi zachifan de~* [You must have eaten a lot during the New Year, come and see how these guys ate] [Video file]. Retrieved from <https://www.douyin.com/video/7058519960766729504?modeFrom=userPost&secUid=55456703574>
- Douyin Safety Center. (2022e). *Yuanlai fenghao zhubo yehui shenye Emo ~ ~* [Moderation influencer also gets sad late at night~ ~] [Video file]. Retrieved from <https://www.douyin.com/video/7050072738928643341?modeFrom=userPost&secUid=55456703574>
- Duffy, B. E., & Hund, E. (2015). "Having it all" on social media: Entrepreneurial femininity and self-branding among fashion bloggers. *Social Media + Society*, 1(2), 1–11. doi:10.1177/205630511560433
- Elberse, A., & Verleun, J. (2012). The economic value of celebrity endorsements. *Journal of Advertising Research*, 52(2), 149–165. doi:10.2501/JAR-52-2-149-165
- Fung, A., Graham, M., & Weil, D. (2007). *Full disclosure: The perils and promise of transparency*. Cambridge, UK: Cambridge University Press.
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven, CT: Yale University Press.

- Gorwa, R. (2019a). The platform governance triangle: Conceptualising the informal regulation of online content. *Internet Policy Review*, 8(2), 1–22. doi:10.14763/2019.2.1407
- Gorwa, R. (2019b). What is platform governance? *Information, Communication & Society*, 22(6), 854–871. doi:10.1080/1369118X.2019.1573914
- Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1), 1–15. doi:10.1177/2053951719897945
- Grimmelmann, J. (2015). The virtues of moderation. *Yale Journal of Law and Technology*, 17(42), 42–109.
- Hearn, A., & Schoenhoff, S. (2016). From celebrity to influencer: Tracing the diffusion of celebrity value across the data stream. In P. D. Marshall & S. Redmond (Eds.), *A companion to celebrity* (pp. 194–212). Chichester, UK: John Wiley & Sons.
- Helberger, N., Pierson, J., & Poell, T. (2018). Governing online platforms: From contested to cooperative responsibility. *The Information Society*, 34(1), 1–14. doi:10.1080/01972243.2017.1391913
- Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction. *Psychiatry*, 19(3), 215–229. doi:10.1080/00332747.1956.11023049
- Huang, G. (2015, July 17). It's all smiles until you say it wrong. *China Daily*. Retrieved from https://europe.chinadaily.com.cn/epaper/2015-07/17/content_21307229.htm
- Jiang, Y. (2022, May 9). China restricts minors from live streaming in bid to improve their health. *South China Morning Post*. Retrieved from <https://www.scmp.com/tech/policy/article/3177019/china-restricts-minors-live-streaming-citing-need-improve-their>
- Jin, S. V. (2018). "Celebrity 2.0 and beyond!" Effects of Facebook profile sources on social networking advertising. *Computers in Human Behavior*, 79, 154–168. doi:10.1016/j.chb.2017.10.033
- Jin, S. V., Muqaddam, A., & Ryu, E. (2019). Instafamous and social media influencer marketing. *Marketing Intelligence & Planning*, 37(5), 567–579. doi:10.1108/MIP-09-2018-0375
- Jing, M. (2019, April 8). No smoking, no tattoos, no bikinis: Inside China's war to 'clean up' the Internet. *South China Morning Post*. Retrieved from <https://www.scmp.com/tech/policy/article/3005252/no-smoking-no-tattoos-or-bikinis-inside-chinas-war-clean-internet>
- Kadri, T., & Klonick, K. (2019). Facebook v. Sullivan: Public figures and newsworthiness in online speech. *Southern California Law Review*, 93(1), 37–99. Retrieved from https://digitalcommons.law.uga.edu/fac_artchop/1429

- Kelly, J. (2021, January 12). *China's toxic work culture results in deaths and suicide*. Retrieved from <https://www.forbes.com/sites/jackkelly/2021/01/12/chinas-toxic-work-culture-results-in-deaths-and-suicide/>
- Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, 'micro-celebrity' and the rise of social media influencers. *Celebrity Studies*, 8(2), 191–208. doi:10.1080/19392397.2016.1218292
- Labrecque, L. I. (2014). Fostering consumer–brand relationships in social media environments: The role of parasocial interaction. *Journal of Interactive Marketing*, 28(2), 134–148. doi:10.1016/j.intmar.2013.12.003
- Laloggia, J. (2019, July 11). *U.S. public has little confidence in social media companies to determine offensive content*. Pew Research Center. Retrieved from <https://www.pewresearch.org/fact-tank/2019/07/11/u-s-public-has-little-confidence-in-social-media-companies-to-determine-offensive-content/>
- Lee, J. E., & Watkins, B. (2016). YouTube vloggers' influence on consumer luxury brand perceptions and intentions. *Journal of Business Research*, 69(12), 5753–5760. doi:10.1016/j.jbusres.2016.04.171
- Literat, I., & Kligler-Vilenchik, N. (2019). Youth collective political expression on social media: The role of affordances and memetic dimensions for voicing political views. *New Media & Society*, 21(9), 1988–2009. doi:10.1177/1461444819837571
- Luo, M. (2022, February 25). China's content moderators are overworked and chronically stressed. *Sixth Tone*. Retrieved from <https://www.sixthtone.com/news/1009742>
- McGuire, W. J. (1985). Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 233–346). New York, NY: Random House.
- Munnukka, J., Uusitalo, O., & Toivonen, H. (2016). Credibility of a peer endorser and advertising effectiveness. *Journal of Consumer Marketing*, 33(3), 182–192. doi:10.1108/JCM-11-2014-1221
- Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, 19(3), 39–52. doi:10.1080/00913367.1990.10673191
- Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, 34(2), 243–281. doi:10.1111/j.1559-1816.2004.tb02547.x
- Roberts, M. E. (2018). *Censored: Distraction and diversion inside China's Great Firewall*. Princeton, NJ: Princeton University Press.

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. New Haven, CT: Yale University Press.

Rubin, A. M., Perse, E. M., & Powell, R. A. (1985). Loneliness, parasocial interaction, and local television news viewing. *Human Communication Research, 12*(2), 155–180. doi:10.1111/j.1468-2958.1985.tb00071.x

Schellewald, A. (2021). Communicative forms on TikTok: Perspectives from digital ethnography. *International Journal of Communication, 15*, 1437–1457.

Seering, J., Wang, T., Yoon, J., & Kaufman, G. (2019). Moderator engagement and community development in the age of algorithms. *New Media & Society, 21*(7), 1417–1443. doi:10.1177/1461444818821316

Spry, A., Pappu, R., & Cornwell, T. B. (2011). Celebrity endorsement, brand credibility and brand equity. *European Journal of Marketing, 45*(6), 882–909. doi:10.1108/03090561111119958

Squirrell, T. (2019). Platform dialectics: The relationships between volunteer moderators and end users on reddit. *New Media & Society, 21*(9), 1910–1927. doi:10.1177/1461444819834317

Suzor, N. P. (2018). Digital constitutionalism: Using the rule of law to evaluate the legitimacy of governance by platforms. *Social Media + Society, 4*(3), 1–11. doi:10.1177/2056305118787812

Suzor, N. P., Van Geelen, T., & Myers West, S. (2018). Evaluating the legitimacy of platform governance: A review of research and a shared research agenda. *International Communication Gazette, 80*(4), 385–400. doi:10.1177/1748048518757142

Suzor, N. P., West, S. M., Quodling, A., & York, J. (2019). What do we mean when we talk about transparency? Toward meaningful transparency in commercial content moderation. *International Journal of Communication, 13*, 1526–1543.

Tran, G. A., & Strutton, D. (2014). Has reality television come of age as a promotional platform? Modeling the endorsement effectiveness of celebrealty and reality stars. *Psychology & Marketing, 31*(4), 294–305. doi:10.1002/mar.20695

Turner, J. R. (1993). Interpersonal and psychological predictors of parasocial interaction with different television performers. *Communication Quarterly, 41*(4), 443–453. doi:10.1080/01463379309369904

Tusikov, N. (2017). Transnational non-state regulatory regimes. In P. Drahos (Ed.), *Regulatory theory: Foundations and applications* (pp. 339–353). Canberra, Australia: ANU Press.

- Wang, S., & Ding, R. (2022). "Business inquiries are welcome": Sex influencers and the platformization of non-normative media on Twitter. *Television & New Media*, 23(8), 822–839. doi:10.1177/15274764221077666
- Weismueller, J., Harrigan, P., Wang, S., & Soutar, G. N. (2020). Influencer endorsements: How advertising disclosure and source credibility affect consumer purchase intention on social media. *Australasian Marketing Journal*, 28(4), 160–170. doi:10.1016/j.ausmj.2020.03.002
- West, S. M. (2018). Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, 20(11), 4366–4383. doi:10.1177/1461444818773059
- Yang, J. (2018, November 2). Most Chinese employees overworked: Survey. *China Daily*. Retrieved from <https://www.chinadaily.com.cn/a/201811/02/WS5bdbdf4ca310eff303286339.html>
- Zeng, J., Chan, C., & Fu, K. (2017). How social media construct "truth" around crisis events: Weibo's rumor management strategies after the 2015 Tianjin Blasts. *Policy & Internet*, 9(3), 297–320. doi:10.1002/poi3.155
- Zeng, J., & Kaye, D. B. V. (2022). From content moderation to visibility moderation: A case study of platform governance on TikTok. *Policy & Internet*, 14(1), 79–95. doi:10.1002/poi3.287
- Zeng, J., Schäfer, M. S., & Allgaier, J. (2021). Reposting "till Albert Einstein is TikTok famous": The memetic construction of science on TikTok. *International Journal of Communication*, 15, 3216–3247.
- Ziyi, T. (2021, October 2). This Chinese influencer resurfaced with a 'manlier' look after being banned from Douyin for being too effeminate. *Today*. Retrieved from <https://www.todayonline.com/8days/chinese-influencer-resurfaced-manlier-look-after-being-banned-douyin-being-too-effeminate>