Flagging Controversies: The Effect of Flagging Mechanisms on the Zhihu Platform

Chen Li

Department of Culture Studies, School of Humanities and Digital Sciences, Tilburg University, Tilburg, The Netherlands

c.li_2@tilburguniversity.edu

Abstract: This empirical study explores the formation and configuration of public opinions on Zhihu, a major knowledgesharing social media platform in contemporary China. Though recent studies examined the impact of flags on citizen journalism, how users make meaning of flagged content and how flags affect public opinions largely lack contextualized explanation and investigation. Thus, this research takes the flagging mechanisms of Zhihu as a vantage point to analyze how public opinions are configurated in a flagged controversy. This study focuses on a posted question on Zhihu: "*How to understand Greta Thunberg's advice for Chinese to stop using chopsticks for the environment's sake?*" This study probes the interplay between users, platforms, and public discourses in the ad hoc controversy. The finding suggested the frontpage and backstage of the flagging activity structure an unequal relationship among Zhihu users. Significantly, the flagging mechanisms broaden the circulation of controversy rather than intervening in spreading rumors. Besides, this study found a frame of newsworthiness employed by Zhihu users. In addition to the framing, there is a pattern of prioritizing information sources in public discussions.

Keywords: Chinese platforms, Zhihu, flagging, knowledge, ad hoc public, content moderation

1. Introduction

In April 2020, a speech given by Greta Thunberg about the usage of chopsticks hit the headlines of major Chinese news websites. According to the coverages, Greta Thunberg gave a speech advising Chinese people to stop using chopsticks to save forests and protect the environment. This speech soon became a controversy that provoked a discursive vortex in social media, especially on the Zhihu platform, one of the largest knowledge-sharing communities and question-answering platforms in the Chinese digital spaces. This controversy is contentious from two perspectives: 1) *the factuality of Greta's speech is in doubt* - if Greta gave the speech is not proved, and 2) *the scientific credibility of Greta's advice* for environmental protection are in doubt. As for the latter, it is questionable if Greta's advice is supported by scientific evidence or by political ideologies. This incident is soon flagged as "of doubtful truthfulness" on Zhihu, while it remains viral among Zhihu users. This case (hereafter as the chopsticks controversy) exemplified a discursive phenomenon involving the interplay between the flagging mechanisms of platforms, users' behaviors, and public discourses.

Zhihu is one of the largest social question-answering platforms among Chinese-speaking users (Zhao et al., 2020). Zhihu users can propose and answer questions about specific topics such as breaking news. Like Quora in English social media spaces, Zhihu encourages knowledge exchange and social collaboration among its users (Wang & Zhang, 2016). Recent scholars also observed that large members of Zhihu users are well-educated, middle-class Internet users (Peng, 2020; Peng et al., 2021). Therefore, Zhihu is an appropriate site for exploring the discursive relationship and socio-cultural structure in the global context of climate change. Previous scholars also pointed out that it is crucial to consider the role of platform mechanisms in producing social structure (van Dijck et al., 2018). However, the flagging mechanisms of digital platforms remain insufficient explorations in scholar inquiry (Crawford & Gillespie, 2016). This literature insufficiency is demanding further empirical and theoretical probes in the studies of the Zhihu platform and the effect of its flagging mechanisms on public deliberations and knowledge dissemination. Therefore, this research is motivated to contribute to the literature gap.

The chopstick controversy is positioned in an ad hoc event of detecting fake news. In this case, the public discussions are pluralized and rife with fragmental information and individual expression of attitudes. This phenomenon leads us to the configuration role of the digital platform in the emergence of the ad hoc public. In an ad hoc public, knowledge gathering and information circulation are decentralized and fragmented (Bruns, 2018). Yet, it needs nuances empirical investigation in different cultures and societies to understand how platforms enable and constrain users' efforts of "collecting, assessing, selecting and curating content" about specific news items (Bruns, 2018, pp.104).

Thus, this research takes the flag mechanism of the Zhihu platform as a theoretical angle, aiming to address the main research question - *how public opinions are shaped and structured in an ad hoc flagged controversy*. As to address the research aim, this study is unfolded following **three sub-questions**: 1) how Zhihu users participate in controversy through the flagging mechanisms on Zhihu, 2) what discursive patterns or characteristics can be observed in users' posts, and 3) what kind of information is prioritized in users' responses to the controversy.

This article firstly reviews the literature on flagging mechanisms and their reification in the following. Then, this paper reflects on the previous research on ad hoc public and Chinese digital public spheres. In the third section, the methodology for this research is elaborated. The section of result and data analysis presents the findings of this study. The closing section is the conclusion and discussion about the contribution and limitations of this study.

2. Theoretical framework

2.1 Flagging mechanisms in Chinese social media platforms

Plenty of major social media platforms, such as Twitter, Facebook, and Weibo, design the function of the flag for crowdsourcing efforts of fact-checking and misinformation detecting to platform users (Kim et al., 2018). Previous research on flagging demonstrates that flags play an important and specific role in the "techno-cultural construct" (Van Dijck, 2013) of social media platforms (p. 29). They are part of available tools that facilitate, compartmentalize, and quantify user feedback (Gerlitz and Helmond, 2013). In other words, flags function in two directions: 1) it enables users to report and signal controversial content such as misinformation and rumors to other users, and 2) it triggers feedback into the platform's algorithms in the way of mediating users' behaviors. As a result, it provides a practical mechanism for regulating and moderating content on behalf of the platform users. Considering the socio-cultural consequences, such effects of flagging could further cause effects on the shaping of public opinions and community values in public spheres (Crawford & Gillespie, 2016). Thus, the flagging mechanisms could serve as a theoretical prism for unfolding the role of platforms in a controversy. In the case of chopsticks controversy, the conceptualization of ad hoc publics could shed light on exploring the discursive effect in the public discussions in a digital environment.

Studies of flagging on Chinese social media platforms remain few in the previous research. Some scholars pay attention to verifying information sources and debunking rumors (Zeng et al., 2019; Xu & Gutsche, 2021). Resonating to Crawford and Gillespie's findings, these studies also found that the mechanism of debunking rumors on Chinese social media encourages users to participate in fact-checking. However, in terms of the effect of flagging in Chinese contexts, previous studies lack socio-cultural contextualized understanding and interpretation. Specifically, how users make meaning of the flagging as a symbolic expression instead of a practical tool deserves further investigation. For instance, in the case of the chopsticks controversy, how users respond to the flag and how users react to the untruthfulness deserve an in-depth analysis. What effect of flagging mechanism could cause public opinion and power relationships requires more empirical studies and theoretical discussions.

2.2 Ad hoc public and Chinese digital public spheres

An emergence of ad hoc public is occupying the Chinese public sphere due to the vast use of social media. Previous studies referred ad hoc publics to the public assembled rapidly by and on social media platforms, often triggered by breaking news stories. This public is defined by their shared efforts to gather, evaluate, selectively amplify, and curate the picture of a breaking news event. Ad hoc publics are characterized by shared topics, low barriers of participation, fragment information, unverified sources, large audiences, and rapid circulations of information (Bruns, 2018). Accordingly, the public discussions gathered under the question "How to understand that Greta Thunberg advised Chinese not to use chopsticks for environment's sake?" can be seen as the ad hoc publics emerged on Zhihu publics. In this chopstick controversy, Zhihu users can participate in the event by answering, following, commenting, liking, and sharing the question page of Zhihu. As a result, a large amount of information was mobilized and reached the broader public.

In the previous research on major social media such as Facebook and Twitter, the emergence of ad hoc publics in a digital environment is academically encouraged and criticized. On the one hand, digital media's facilitation role helps broaden public engagement in social issues, such as collective citizen journalism (Hermida, 2014, p.360). Also, the agency of users in the production and circulation of information is expanded beyond professional and conventional journalists (Vis et al., 2013). In addition, the collaboration between users and

digital platforms is found contributing to the social media curation of diverse information (Papacharissi & de Fatima Oliveira, 2012). On the other hand, the emergence of ad hoc publics vis-à-vis public deliberations is often challenged and criticized due to the opaqueness of the platform's algorithms and the emotion-charged responses in the digital public spheres. The foremost criticism is that the backstage of a digital platform, i.e., algorithms for information flows and recommendations, remains unpublic. How they affect public attention and journalist practices is untransparent to the public (Crawford& Gillespie, 2016). Besides, the affective connection and personal emotions blended in the public deliberations about news events and public concerns may tradeoff the Habermasian ideal for public spheres (Schäfer, 2015). Moreover, the low barrier of participation and fast spread of voluminous information are obstacles to preventing rumors (Zubiaga et al., 2016).

Particularly, considering the contemporary Chinese digital public spheres, a media-hype manner of discussing public concerns and current affairs is seemingly occupying the Chinese public sphere due to the vast use of social media. Meanwhile, much of English scholarship rarely notice the dynamics of the digital space in China but are often confined within a conceptual framework that merely focuses on the "democratization" potential of the internet and the subversive agency of netizens (Schneider, 2018). It is remarked that many of the English-language scholarship is bogged down in a pervasive "digital orientalism." This paradigm regards the digital process in China as a unique case in which the Internet is ought to be facilitating democracy and the subversive power against the nation-state (Schneider, 2018, p.18). Following this digital orientalism, some scholars argue that the "domestic tug of war between the Party-state and society in China" is playing itself out on the internet (Lagerkvist, 2010, p.20). Therefore, it needs to be pointed out that a large part of the previous studies has generalized the role of social media technology as a catalyst for democratization in Chinese society. They assume confrontational relationships between the state and the people as the sole force underlying and mobilizing the digital public sphere. Therefore, it is demanded to take a close look at features and functions of Chinese platforms to give a situated view on the contemporary Chinese digital public sphere.

To fill the gap between the changing platform mechanisms and the insufficiently contextualized understanding of contemporary China, this research takes the function of flagging on Zhihu as a vantage point. It aims to dissect the socio-cultural consequences of the controversy. This study pays specific attention to how the flagging mechanisms of Zhihu constrain and enable users to participate in a controversy-based ad hoc public. Further, regarding the globalization and digitalization in contemporary societies, this exploration could spotlight the social and political implications of digital communication within and beyond China.

3. Methodology

This research adopts digital ethnography as a research scheme for approaching the research field. This scheme focuses on digital platforms, specific social contexts, and semiotization (Varis, 2016). As such, digital ethnography enables researchers to approach the field as situated and contextualized in the digital environment afforded by platforms. Hence, this study is conducted by following users' discursive activity overtime on the Zhihu platforms, examining the interplay between platforms, users, and public discourses. For this research, I set my fieldwork site on one Q&A site pivoting on the posted question "How to understand that Greta Thunberg advised Chinese not to use chopsticks for the environment's sake?." I chose this specific Q&A site because it was the only query flagged as "doubtful truthfulness" on the incident of Greta's speech. Also, this case aggregates the most answers amongst all questions on the Zhihu platform relating to the incident of Greta's speech. As below, Figure 1 demonstrates the basic information of the flagged Q&A page of the chopstick controversy on Zhihu, providing a context for understanding and analyzing this discursive phenomenon.

As for data collection, this research retrieved posts from a Q&A site on Zhihu. During the digital ethnographic observation and data collection, I collected 1,556 posted answers under the questions. Because the Zhihu platform does not provide APIs for data crawling, I manually collected data until the themes and patterns appeared in data achieve saturation. The forms of data include texts, images, video, and the metrics on each posted answer. Among others, one piece of online video is recorded. The rest is text-and-image based data screenshotted in the form of PDF. In terms of the metrics, it is only possible for collecting the numbers of comments and upvoting of each post. Other metrics, such as downvoting, is not visible due to the feature of the Zhihu platform (Guo & Caine, 2021). This research also recorded the digital classifications designed by the platform: user's verification of expertise and influences. The recording continued from April 2021 to October 2021. Table 2 illustrate the recorded data of the observed answers.

Chen Li

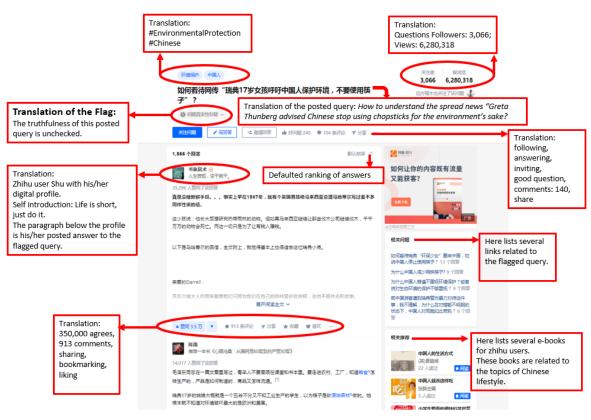


Figure 1: The flagged question-and-answer page of the chopstick controversy on Zhihu platform

Table 1 Summary of observed users' answerers in the chopstick controversy

Username	Labelled in this	Zhihu badge (s)	Likes	Comments	Date
	paper as				
书孰鼠术	Shu	Influential (1,000 followers)	350,000	913	2020-11-30
阵雨	Zhen	Influential (100,000 followers)	140,000	350	2020-05-05
远方青木	Yuan	Influential (100,000 followers) Professional on this question (awarded by another user)	4,433	405	2021-04-17
菲拉克图斯	Fei	none	3,510	270	2020-04-19
王冰	Wang	Influential (1,000 followers)	3,868	235	2021-05-18
我变成了一条 狗	Wo	Influential (100,000 followers)	190,000	531	2020-04-13
Wave	Wave	none	1,795	97	2020-04-25
英国留学君	Ying	Influential (1,000 followers)	8,768	241	2021-07-19
林耳	Lin	Influential (100,000 followers)	2,628	353	2020-04-13
继续者张付	Ji	Influential (100,000 followers)	9,163	478	2020-12-19
小马	Xiao	Influential (1,000 followers)	2,240	324	2020-04-23

This research adopts discourse analysis (Jones et al., 2015) as an analytical instrument to investigate the data. Precisely, analysis is unfolded from perspectives of contexts for meaning making (digital interface and actor's posts), the content of answers (texts and discursive patterns of posted answers), actors (the answerers on the Q&A page), and the power relationships among Zhihu users, and between users and the platform. Discourse analysis in the digital environment considers texts as the "aggregate of semiotic elements that can function as a tool for people to take social action (Jones et al., 2015, p.5)". Thus, this research emphasizes the textual data, including articles, images, audio messages, video clips, and on-screen texts such as metrics, hyperlinks, and digital interfaces. In addition, according to Jones et al., discourse analysis strengthens this research with a critical perspective of probing into ideological agenda and power relationship that often unexpressed on texts (2015). Hence, this research will focus on what is prioritized in constructing public discussions by platform mechanisms and platform users.

4. Results

4.1 Flagging as an act: Frontpage and backstage

Flagging as a technological function of the Zhihu platform provides both front-page information and backstage access for users to understand and respond to a controversy on the platform. Regarding the front page of the chopstick controversy, the flag can be seen under the question, signaling controversial content. On Zhihu, when a user posts a question, other users can follow, answer, comment, bookmark, share, flag, and invite different answerers to answer the question. Regarding these platform-provided social functions, flagging showcases definitions of "controversy" on a question-answering platform.

In this case, Zhihu classified controversial content into eight categories: 1) the posted question is of low quality, 2) the posted question contains advertorial spams, 3) the posted question contains online trolling and violence, 4) the posted question contains harmful information, 5) the posted question involves infringement, 6) the posted question contains harassment content, 7) the posted question contains information against Chinese cyber civilization agenda (footnote), and 8) the posted question implies self-harm and suicidal attempts. These eight types together define the meaning of "controversy" on Zhihu. As such, Zhihu users construct a technocultural understanding of controversy. On the one hand, the Zhihu platform defines the "controversy." The flagged controversy phenomenon on the Zhihu platform can be regarded as a communicative genre. It is recognizable and pluralized by interactions among users and between users and the platform. On the other hand, users following the regulation can flag any content as "controversy." The flagging behavior will algorithmically label the content and feed it into the platform algorithm, recommending it to users who are active in discussing controversies. By such a procedure, a mechanism of platform moderation is accomplished by the collaboration between platform regulation and user labor. In the case of chopstick controversy, the posted question is flagged in the first category, warning that "the posted question is with doubtful truthfulness." Due to this warning signal, content moderation also becomes part of the discussion. The posted question is no longer the only focal topic for users. The flagging becomes an inviting signal for users to pay attention to the rumor detection.

Besides the front page, users can also participate in the controversy through a backstage-like "question log." This question log shows how different users edited a query on the backstage of the flag over time. As illustrated in figure 2, the question log of the controversy provides a historically contextualized understanding to users. This question log shows two points about the meaning-making of the controversy. Firstly, though Zhihu users can edit the content, the Zhihu platform owns the authority to make the final decision on it. Secondly, only those users with high scores of "yan" ("盐值" in Chinese, a symbolic calculation for classifying Zhihu users) can edit the log. According to the regulation of the Zhihu platform, how to calculate yan scores of users is emphasizing the degree of users' activeness and the impact of users' behaviors on the platform. However, the formula of calculating and weighing is not transparent to users. These two observations suggested a nuanced mechanism of the emergence of ad hoc public proposed by Bruns (2018). Instead of the low barriers to participation, the Zhihu platform has opaque criteria to the public, heightening the threshold for public engagement. In addition, the topics of ad hoc public are not always shared by participating users. In the chopsticks controversy, both the front page and question log show that users are constantly debating on the topics. Especially on the question log, topics are changed over time, from the "Chinese" to "foreigners" at the early stage to "environmental protection" to "fake news" later. However, no matter how high-yan-scored users edit the topics, the Zhihu platform owns the final decision. As we can see in the figure, the Zhihu administrator can lock, close, and resume the discussion and edition on the question.

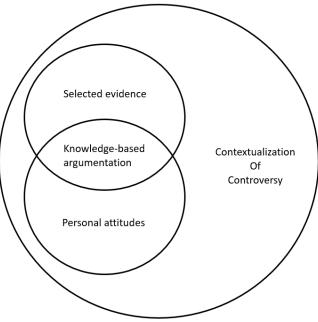
Chen	Li
------	----

	〈 问题日志 - 如何看待网传"瑞典17岁…	•••					
	环境保护 中国人	境保护 中国人					
	如何看待网传"瑞典17岁女孩呼吁中国人保护 境,不要使用筷子"?	环					
	问题编辑日志						
\langle	知乎管理员 设置问题状态为 正常	 The administrator of Zhihu platform set the status of question as normal. 					
	#70785827880 - 2020-12-01 15:13:09						
\langle	知乎管理员 设置问题状态为 关闭:	 The administrator of <u>Zhihu</u> platform set the status of question as closed. 					
	#70785601537 - 2020-12-01 15:11:31						
$\left(\right)$	知乎管理员 锁定问题编辑 问题编辑锁定 #48622993223 * 2020-05-31 16:35:35	The administrator of Zhihu platform set the status of question as locked.					
	#48622993223 * 2020-05-31 16:35:35						
<	知乎用户o4oJE9 移除了话题 网络谣言	A Zhihu user removed the hashtag "online rumors".					
	#48622983977 · 撤销 · 举报恶意编辑 · 2020-05-31 16:35:21						

Figure 2: A screenshot of the "question log". Translation is placed on the right side of the figure.

4.2 Meaning-making of flagging in an ad hoc controversy

The posted answers in the chopsticks controversy showcase a pattern of framing newsworthiness in discursive practices. Three strategies are observed: contextualizing controversies, selecting evidence, and showing personal attitudes. The latter two are often equipped with specific knowledge, such as scientific or journalist expertise (See figure 3).





4.2.1 Framing controversy in specific socio-cultural context

In terms of contextualizing the controversy, the socio-political contentions towards environmentalism are a prominent theme. This is also responding to the topic posted on the question. In the posted answers, the socio-political contentions are specifically exemplified in two directions. One is the challenges against the "western" dominance of environmental discourses. The other is the politicized scientific stances in ecological issues from the "non-western" perspective. In the following analysis, I take Zhihu users Shu and Zhen as examples, presenting how influential answerers contextualized the discussion in the chopsticks controversy.

Chen Li

From the answer of Shu, the first point we can observe is that the answer neither responds to Greta's speech nor evaluate the factuality of the speech. Instead, it started straightly with an accusation of how Britain destroyed Malaysia's environment. This accusation suggested that the responsibility of Britain to protect forests is historically ignored. This case illustrates two significant characteristics of Zhihu answerers: 1) whether Greta's speech is fake news is not the focus of the controversy, and 2) this controversy is not only about the topic "Chinese" but also related to different non-western societies. Similar to Shu's answer, user Zhen straightly pointed to the anti-green consumption of countries of European and America (see figure 4 below). In addition, the answer provided by Zhen extends the discussion from protecting forests to criticizing Europe-and-Americaled international business of disposable cutlery. These posted answers thus contextualized the chopsticks controversy in a contesting historical and political point where "western dominance" in the climate change discourse is questioned and challenged by non-western societies. Further, this contextualization has two discursive effects. First, it tacitly approves the truthfulness of the news about Greta's speech on the usage of chopsticks. Secondly, it enables the algorithm to identify the controversy with the clustered content of environmental issues instead of encouraging fact-checking and misinformation-detecting. By such interplay between Zhihu users' answers and the platform algorithm, the flagging function of Zhihu is transmuted into an invitation or a venue for broader discussions and debates in the discourse of climate change and environmentalism. In other words, flagging is no longer a warning signal for stopping the circulation of fake news like the vision of Brun's ad hoc public. Instead, ironically, it encourages the spreading of unchecked information and rumors.

く	Translation: "Green" European and American countries will
而"环保"的欧美各国,会在中国沿海血汗工厂订	order large quantities of disposable tableware in
购大批一次性餐具。这些餐具包括餐盒和一次性刀	China's coastal sweatshops. The cutlery includes
叉勺。	lunch boxes and disposable forks and spoons.

Figure 4: Screenshot of the answer posted by Zhihu user Zhen

4.2.2 selecting evidence and personal attitude

Following the contextualization of public concerns, the second component of framing newsworthiness is the knowledge-based argument. Noticeably, it is often blended with selected evidence and the personal attitudes of influential answerers. The chopsticks controversy shows that answers with high ranking and large responses are often rife with knowledge-specific evidence such as historical facts, scientific findings, external information sources, and personal experiences. Blended with the evidence, these answers often employed emotion-charged expressions. In the following, I will take Zhihu user Yuan's answer as an example, presenting how a knowledge-specific answer is blended by evidence and personal attitudes.

In the answer posted by Yuan, the scientific evidence is specifically aimed to contradict the narrative of protecting forests. In Yuan's response, the main argument is that the core of environmental protection is slowing industrialization instead of planting trees. As for supporting the argument, Yuan selected the advice of Ding Zhong-Li as vital evidence. Ding Zhong-Li is an academician of the Chinese Academy of Science. In the realm of climate change, he was well-known for outspokenly questioning the western-dominant narrative. In Yuan's answer, Ding Zhong-Li was not only cited but also hyper-linked in his answer. This linked keyword of Ding Zhong-Li strengthens the credibility of Yuan's answer. In addition, it could aggregate related content with users who are interested in the topics about Ding Zhong-Li and associated discussions about climate change. As an effect, the selection of this evidence supports Yuan's answer and gives it an expert-like position.

Besides the selected evidence, readers can also sense a resisting attitude of Yuan against the oppressing position of America in the global market of carbon emissions. In Yuan's answer stated that "America took its refusal as a matter of course. Make every Chinese enjoy 80% of the carbon emissions of Americans? Don't make a fool of us. That's not going to happen. Chinese don't deserve such portion of carbon emissions." The tone and the ironic rhetoric in Yuan's answer strongly expressed the resentment towards the inequality in the global market of carbon emissions shares. In addition, this emotion-charged message often appears in the topic of "Chinese." The

chopsticks controversy is also tagged with this topic. At this point, we may find a patriotist narrative and antagonistic emotion in Yuan's expert answer.

4.3 Curating ad hoc public: Prioritization of unflagged-concern

Regarding content curation on the Zhihu platform, the chopsticks controversy case suggested two layers of the order of visibility:1) prioritization in public discourses and 2) prioritization in information resources.

Firstly, the chopsticks controversy illustrates those answers, including videos, texts, and images, are presented in a platform-designed order. This order is described as "by default," suggesting a neutral position. However, the order of these answers shows a prioritization in terms of modality and groups value. From the aspect of modality, the highest-ranked answer is a "video answer" (see figure 5 below). What is noticeable is the brief introduction under the video: "Watching the video helps get a quick answer." This short sentence suggested a platform prioritization of high velocity instead of constructing community values or gate-watching public discourses. In addition, compared to other answers, the metrics such as views, up-voting, likes, shares, and comments are not as large as the lower-ranked answer. Yet, the video is still prioritized for readers to read and understand the controversy. Besides the prioritization of modality, the other aspect is groups values. Among the text-and-image based posts in the chopsticks controversy, the highest-ranked answers are majorly emphasizing the discourses of environmental protection instead of fact-checking. In this flagged controversy, though it is flagged as "doubtful truthfulness," substantial amounts of top answers are confronting Greta and the westerncentered narratives of climate change. On the one hand, this prioritization suggests an algorithmic order designed by opaque calculation. On the other hand, it reveals the effect of the flagging mechanisms of datafication in the process of knowledge dissemination. What is debated is no longer what is flagged for. Instead, how to attract more attention and facilitate data traffic drive the changes of public discourses and groups' values in a controversy.



Figure 5: The highest-ranked answer is a "video answer"

Secondly, besides prioritization in public discourses, how users evaluate the countability and trustworthiness of information sources in the controversy showcases a prioritization of external resources. In this chopstick controversy, when it comes to the factuality of Greta's speech, most Zhihu users migrate the information outside Chinese media to Zhihu. For example, one Zhihu user screenshotted related information on Quora.com, an English SQA platform, and posted them on Zhihu. Without detailed translations of the information from English to Chinese, this Zhihu user pointed out that Greta did not give the speech in the case by underlining some sentences in the English resources in red. This pattern of migrating information from English media to Zhihu reveals that though the flagged controversy provides an ad hoc public where users share information, there is a priority of outside information among different information sources. On the one hand, this suggested the inequality of access to information among Zhihu users. On the other hand, it unveiled public spaces where discourses are fragmentized due to different information access and prioritization.

5. Conclusion and discussion

This paper analyzed the flagging mechanism and its effect on an ad hoc controversy on Zhihu platform. It is found that the public discourses in the case are pluralized beyond what was flagged for in the first place. In other words, the research findings revealed that the meaning-making of flagged content is both defined by platform regulation and groups values. As a result, an ad hoc public based on the flag mechanism facilitates the data traffic by spreading contentions instead of intervening in the circulation of misinformation and rumors.

Specifically, the interplay among the backstage access, users and platform operators revealed that flagging as an activity of "gatekeeping" (Crawford & Gillespie, 2016) is obscurely constrained by platform operation. What is noticeable is the interplay between users' agency and platform-designed open backstage area. However, such

backstage is unequally accessible among users. This finding resonates with the "point structure of participation" (Crawford & Gillespie, 2016). Yet, the case of Zhihu is nuanced due to the platform algorithms of how to calculate points of users and who has the final decision on removing or retaining controversial content. Further, platform mechanisms algorithmically designed the inequality for increasing data traffic.

Besides, this research finds that the meaning-making of flagged content in an ad hoc public is pluralized. Platform users do not always value or emphasize what is flagged for. Instead, participants often use a frame of newsworthiness to contextualize and "reproduce" (Blommaert, 2005, p.48) meaning. In this frame, I observe a pattern of blending knowledge-specific evidence and personal attitude. Significantly, this pattern often contributes to the construction of an amateur-expertise narrative. Previous research also pointed out that scientific communication on Zhihu nowadays has deconstructed the boundary between experts and non-expert (Yang, 2021). This finding may contribute to the discursive practice of how the discourse authority of non-scientists is constructed on the platform environment.

In addition, this research further highlights the prioritization of content in terms of velocity and groups values. Previous research found that members of Zhihu are majorly well-educated, middle-class intellectuals (Peng 2020). Since the nature of this chopstick controversy is relevant to scientific knowledge and information access, it encourages the participation of Zhihu users who have access to information from non-Chinese media and who have a knowledge background of environmental protection. What is prioritized in their posts unveiled how today's internet-literate evaluate information and its effects on knowledge dissemination and circulation. This finding may contribute to understanding how "public trust" (Gauchat, 2011; Habermas, 1989; Yang, 2021) of science-related information is conditioned and stratified in a knowledge-specific controversy.

References

Blommaert, J.M.E., 2005. Discourse: A critical introduction. Cambridge University Press

- Bruns, A., 2018. Gatewatching and news curation: Journalism, social media, and the public sphere (Digital Formations, Volume 113). Peter Lang Publishing.
- Crawford, K. and Gillespie, T., 2016. What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society*, *18*(3), pp.410-428.
- Gauchat, G., 2011. The cultural authority of science: Public trust and acceptance of organized science. *Public Understanding of Science*, 20(6), pp.751-770.
- Gerlitz, C. and Helmond, A., 2013. The like economy: Social buttons and the data-intensive web. *New media & society*, *15*(8), pp.1348-1365.
- Guo, C. and Caine, K., 2021. Anonymity, user engagement, quality, and trolling on Q&A sites. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), pp.1-27.
- Habermas, J., 1989. Jurgen Habermas on society and politics: A reader. Beacon Press.
- Hermida, A., 2014. Twitter as an ambient news network. Twitter and society, 89, pp.359-372.
- Jones, R.H., Chik, A. and Hafner, C.A., 2015. Discourse and digital practices. London and New York: Routledge.
- Kim, J., Tabibian, B., Oh, A., Schölkopf, B. and Gomez-Rodriguez, M., 2018, February. Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In *Proceedings of the eleventh ACM international conference on web search and data mining* (pp. 324-332).
- Lagerkvist, J., 2010. After the Internet, before democracy: Competing norms in Chinese media and society. Peter Lang.
- Papacharissi, Z. and de Fatima Oliveira, M., 2012. Affective news and networked publics: The rhythms of news storytelling on Egypt. *Journal of communication*, 62(2), pp.266-282.
- Peng, A.Y. and Peng, 2020. A Feminist Reading of China's Digital Public Sphere. Springer International Publishing.
- Peng, A.Y., Hou, J.Z., KhosraviNik, M. and Zhang, X., 2021. "She uses men to boost her career": Chinese digital cultures and gender stereotypes of female academics in Zhihu discourses. *Social Semiotics*, pp.1-19.
- Schäfer, M.S., 2015. Digital public sphere. The international encyclopedia of political communication, 1(7).
- Schneider, F., 2018. China's digital nationalism. Oxford University Press.
- Van Dijck, J., 2013. The culture of connectivity: A critical history of social media. Oxford University Press.
- Van Dijck, J., Poell, T. and De Waal, M., 2018. *The platform society: Public values in a connective world*. Oxford University Press.
- Varis, P., 2016. Digital ethnography. The Routledge handbook of language and digital communication, pp.55-68.
- Vis, F., Faulkner, S., Parry, K., Manyukhina, Y. and Evans, L., 2013. Twitpic-ing the riots: Analysing images shared on Twitter during the 2011 UK riots.
- Wang, Z. and Zhang, P., 2016. Examining user roles in social Q&A: The case of health topics in Zhihu. com. Proceedings of the Association for Information Science and Technology, 53(1), pp.1-6.
- Xu, N. and Gutsche Jr, R.E., 2021. Going Offline": Social Media, Source Verification, and Chinese Investigative Journalism During "Information Overload. *Journalism Practice*, *15*(8), pp.1146-1162.

Chen Li

- Yang, Z., 2021. Deconstruction of the discourse authority of scientists in Chinese online science communication: Investigation of citizen science communicators on Chinese knowledge sharing networks. *Public Understanding of Science*, 30(8), pp.993-1007.
- Zeng, J., Burgess, J. and Bruns, A., 2019. Is citizen journalism better than professional journalism for fact-checking rumours in China? How Weibo users verified information following the 2015 Tianjin blasts. *Global media and China*, *4*(1), pp.13-35.

Zhao, W., Lu, P., Yu, S. and Lu, L., 2020. Consumer health information needs in China–a case study of depression based on a Social Q&A community. *BMC Medical Informatics and Decision Making*, *20*(3), pp.1-9.

Zubiaga, A., Liakata, M., Procter, R., Wong Sak Hoi, G. and Tolmie, P., 2016. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PloS one*, *11*(3), p.e0150989.