

**Underrepresentation and Biased Representation of Frontline Women Health Workers
in Chinese Social Media**

(中文社交媒體對前綫女醫務工作者低度偏差再現)

Yunwen Wang,¹ Xiaoqin Li,² and Qian Xiao²

¹Annenberg School for Communication and Journalism, University of Macau

²Department of Communication, Faculty of Social Sciences, University of Macau

Correspondence author

Dr. Xiaoqin Li, Department of Communication, University of Macau.

Email: xqli@um.edu.mo

首屆“雙三角”新聞傳播學論壇

2021年12月19日綫上

Abstract

Women constitute the major force in the frontline health sector under the unprecedented pressure of COVID-19. Nevertheless, we found that women health workers were still unrepresented and represented with biased on social media. This study used Latent Dirichlet Allocation (LDA) topic modeling to analyze posts ($N = 199,110$) containing the keyword “援鄂医疗队” (i.e., “aiding-Hubei medical team”) published in the first stage of COVID-19 (January 22, 2020 to June 30, 2020) on a Chinese social media, Weibo. Results revealed that health workers were discussed in relation to serving the “nation,” saving the “patients,” and fulfilling family roles. Among the 13 topics in this online discourse, topics discussing health workers in a non-gendered lens tended to emphasize their professional role. When the female identity of health workers was explicit, depictions tended to emphasize their family roles. Meanwhile, the wave-riding and hashtag-jumping/hijacking of corporates and male celebrities – in celebration of their charitable work toward the COVID-19 medical teams – also deprived women health workers of their voices and visibility. Theoretical and practical implications of women health workers’ social media underrepresentation and misrepresentation are discussed.

Keywords: health worker, gender stereotype, media representation, familism, hashtag-jumping, computational text analysis

中文摘要

新冠時期，女性醫務工作者面對著前所未有的壓力，充當了前綫抗疫的主力。然而，我們的研究發現，女性醫務工作者仍然在社交媒體中被嚴重偏低、且帶有偏差的再現。本研究使用 LDA 模型分析了包含“援鄂医疗队”关键词的 199110 条微博，揭示 COVID-19 首轮爆发期间（2020 年 1 月—2020 年 6 月）与医务工作者相关的社交媒体如何呈现。研究發現，針對這些醫務工作者的話題討論集中在服務國家、救治病患、和履行家庭義務/角色的三個面嚮。在該文本總結出的 13 個語義主題中，當討論角度與性別無關時，微博文本往往強調了專業角色；而如果在文本中涉及到醫務工作者的性別（“她”）時，則會更多地突出其家庭角色。研究同時發現，企業與明星“蹭流量”者也大有人在，尤其是打著為醫療隊伍做慈善的男明星，因此進一步壓制了女性醫務工作者的聲音與呈現，這恐怕是社交媒體時代的新問題。女性醫務工作者的低度與偏差性呈現問題的理論與實踐意義在文中進行了討論。

关键词：医务工作者，性别刻板印象，媒体呈现，家庭主義，蹭流量，CTA

Underrepresentation and Biased Representation of Frontline Women Health Workers in Chinese Social Media

In news media, gender stereotypes of occupations can lead to misrepresentation, as such for nurses (UN Women, 2020). Stereotypes of health workers are deeply ingrained in public discourse and reinforced by the sociodemographic structure (Garg et al., 2018). Among the 8.91 million health technicians in China in 2017 (Min et al., 2018), 98% were female (Li, 2018). Though male nurses in China doubled from 1% in 2010 to 2% in 2017, the percentage was lower than that of developed countries (Li, 2018), e.g. 25% for the United States in 2015 (Garg et al., 2018). Along with such juxtapositions is a common assumption that male physicians suit the professional role better than women because of women's "natural predilection" for family life (Brubaker, 2020). The reality, however, has witnessed women's immeasurable contributions to the combat with COVID-19 at the frontline, away from homes. Of the 42.6 thousand professionals of the 346 aiding-Hubei medical teams, over 60% were female (The Paper, 2020). Against this backdrop, this study shed light on health-worker-related online discourse and social media representation during the first stage of COVID-19, January 2020 to June 2020.

Literature Review

Gender, Gender Identity, and Gender Role

In elucidating media representation of women, it is essential to understand and distinguish the concepts of gender, gender identity, and gender role. The broad concept of gender and gender identity, according to Stoller (2020), included all the characteristics determined by various biological, psychological, social, and cultural factors. More specifically, gender refers to the social and biological aspects of sex, while gender roles generally describe the social regulations or stereotypes related to genders, and the degree to which individuals comply with

social expectations on an assigned gender (Huyck, 1999). By contrast, gender identity is a more intimate relationship between individuals and their gender, namely individuals' perception of their body and the relationship they expect with their gender role (Huyck, 1999). It is worth noting that individuals' self-recognition of gender identity is not static; rather, it develops over time, through continuous conscious and unconscious mental representations in dialogue with others (Tyson, 1982). As Tyson argued (1982), gender identity would eventually evolve into a "mix" of masculinity and femininity, which makes the gender identity non-binary.

In the traditional Chinese consciousness of gender inequality, women are born as a subsidiary of men (Cheung, 1996; Chen, 2004). The stereotyped gender roles stipulated in traditional Chinese society are socialized in the family and reinforced by other social institutions (Chen, 2004), such that the only place where women are given status is in the family, and that they are expected to play an important supporting role in managing the family and raising children (Cheung, 1996; Chen, 2004). In an environment of media information, the media inevitably shapes gender roles (Scharrer, 2013), nurtures the socialization of gender (Steele, & Brown, 1995), and influences the public perception of gender roles (Scharrer, 2013). The social construction of gender is composed of various socially recognized normative attributes and archetypes, and thereafter presented in various social settings, such as the media (Steele, & Brown, 1995). However, the media's worldview is often distorted to a certain extent in the portrayal of gender roles in the real world, especially with respect to the actual contributions of men and women in society (Hust & Brown, 2008).

Gender-related Bias and Underrepresentation of Women in the Media

When it comes to gender equality in the news media, progress has virtually ground to a halt. Women constitute only 24% of the persons heard, read about, or seen in radio, newspaper,

and television news across 114 countries over the course of 20 years (UN Women, 2020). Moreover, only 37% of stories in newspaper bylines and newscast reports as of 2015 were reported by women, and the proportion had not increased over a decade (UN Women, 2020). As such, a primary way in which the media distorts reality is underrepresenting women (Wood, 1994). Feminist media research has long been concerned with the reduction and elimination of women's visibility in the media (Gallagher, 2002; Ross, 2010). Research has shown that women's visibility in the media is low (Gallagher, 2002). Under the male-dominated power structure of the newsroom operation mechanism, women's views and voices are marginalized (Ross, 2011). The recognition of women as autonomous and important participants in the economic, social, and political arena is significantly lower than that of the actual positions they hold in the news field (Ross, 2011). On the contrary, when a woman appears in the media, they tend to be first described as a family member, such as a mother and a wife, regardless of their professional and social statuses (Carter, 2011; Niven, 2005). There may be a structural conditions of unbalanced gender representations across different contents in newspapers (Sjøvaag & Pedersen, 2019).

While digital media is promising to promote literalization (Esarey & Xiao, 2011), bias associated with women and the underrepresentation of women found in traditional news media also appears in digital news. According to UN Women (2020), women account for only 26% of Internet news reports and media news tweets. Such unbalanced reporting, or underrepresentation of women in news reports, indicates a form of bias toward women in the digital media. Another form of bias also exists. Even when journalists decide to write about a woman, "journalists are much more likely to consciously decide to criticize a female politician or frame an article about a female athlete in terms of their appearance than they are to explicitly choose not to write about a

woman” (Courtney et al., 2020, p. 1283). Similarly, Lundell and Ekstroumlm’s (2008) demonstrated through a visual communication analysis that a de-emphasizing femininity strategy prevails in the press when they report women in politics. These biases and gender stereotypes often cause women to be limited to traditional feminine topics in media reports (De Swert & Hooghe, 2010).

Investigating the reasons behind this phenomenon, scholars posit that the existing gender inequality in economic participation underlies the gender inequality in media reports (De Swert & Hooghe, 2010). However, Courtney et al. (2020) stress that we should not assume media coverage about women would increase proportionally as women advance in public life. Moreover, the overarching dominance of men’s power has led the women-related agendas to be more scrutinized under the censorship system (Shor et al., 2015).

Gender Stereotypes of Occupations in News Media

Analyzing enormous Google News data (100 billion words), Garg et al. (2018) used word embeddings to find significant associations between the gender percentages of occupations in 2015 and the gender bias of occupations over time. They revealed gender-oriented occupation stereotypes in the news media, especially in associating women with nurses, housekeepers, librarians, and dancers, as opposed to associating men with mechanics, engineers, and carpenters (Garg et al., 2018). At a macro-level, the gender bias of occupations seemed to be consistent with the relative percentage of women in these occupations in the selected year; at a micro-level, however, gender stereotypes attached to certain occupations far exceeded their representation in the real world (Garg et al., 2018). Nurses, in particular, experience arguably the most severe gender stereotypes in news media (Garg et al., 2018).

Women at the COVID-19 Frontline

As mentioned above, women health workers constitute the major force at the COVID-19 frontline (The Paper, 2020). Globally, however, women hold only 25% of leadership positions despite that they comprise 70% of the health workforce, although providing health services to approximately 5 billion people (World Health Organization, n. d.). In the EU, as well, women make up the majority (76%) of healthcare workers (European Institute for Gender Equality, 2020). In China, more than 50% of the doctors fighting against COVID-19 on the front line are women, and the proportion of female nurses has reached more than 90% (21st Century Business Herald, 2020). Despite these statistics, neither the epidemic management decision-making agency nor the propaganda media agency reflects the balance of gender (The Union for the Mediterranean, 2020).

“Countries are not producing enough gender-related data on infection rates, or mortality among health care workers,” a nurse in Baltimore, United States, shared her thoughts with National Geographic on the unrelenting pressure of Coronavirus (Jones, 2020). Particularly, women frontline workers had a higher risk of mental health challenges, including depression, insomnia, and anxiety at the beginning of the COVID-19 pandemic (Lai et al., 2020). Both the gender stereotypes of occupations in news media and the gender inequity in the COVID-19 frontline pose an imperative for updated research to study the gender-specific experience and media representation associated with women health workers during a global public health crisis.

Women’s (in)visibility in the media has been under scrutiny in the context of COVID-19. In a total of 56 U.K. government’ daily press briefings, 43% featured an all-male lineup without female politicians or experts, and only 7% reported two briefings led by a female politician (Smith, 2020). Similarly, a photo of the U.S. Coronavirus Taskforce led by the then Vice

President Mike Pence in February 2020 showed only men figures (Jones, 2020). Recent studies about the representation of women in COVID-19 discourses primarily focused on women's (in)visibility among politicians (Smith, 2020) or among medical experts who have the authority of medical science to speak about COVID-19 (Tezel et al., 2021; Triana et al., 2021). The majority of health workers who were neither politicians (Smith, 2020) nor expert speakers appearing in cable news (Tezel et al., 2021) were further marginalized in journalism research. As of October 2021, there was only one content analysis that examined female nurses' media visibility and voices in China (along with the U.S. and India) during the COVID-19 pandemic (DeWees & Miller, 2020) -- to the best of our knowledge. The content analysis (DeWees & Miller, 2020) indicated women health workers' visibility in the news but also country-specific differences in how their experiences were presented. Given the limited research on the media representation of the whole community of women health workers during COVID-19, especially on social media which entails more user-generated content, we asked the following research questions in scopes of familism and gender roles (Basow, 1992) and in the context of China.

RQ1. Based on the distributions and peaks of social media postings about aiding-Hubei medical teams during the first stage of the COVID-19 pandemic, how were health workers (particularly women health workers) portrayed on social media?

RQ2. In this online discourse about aiding-Hubei medical teams, did the representation of women health workers accurately reflect their actual professional participation?

Methods

Sample and Procedure

Sample of this study are unique microblogs (i.e. weibo, $N = 199,110$) from Sina Weibo after deduplication. We searched for all microblogs posted between January 1, 2020 and June 30,

2020 that contained the Chinese keyword “援鄂医疗队” (i.e. “aiding-Hubei medical team”) in its main text including the hashtags; Wuhan is the capital city of Hubei Province in China. Data distributed between January 22, 2020 and June 30, 2020, as the keyword did not occur on Weibo before then. Starting January 22, 2020, medical professionals (or health workers) totaling 42.6 thousand (The Paper, 2020) from all over China traveled to Hubei to help meet the dire need of medical professionals suddenly imposed by the COVID-19 pandemic.

Data Preprocessing

Data was processed and analyzed using self-designed as well as published algorithms using Python. We preprocessed the dataset prior to analyses by firstly removing punctuations. We then performed word segmentation using the pkuseg toolkit that was built for multi-domain Chinese word segmentation (Luo et al., 2019). In Chinese Natural Language Processing (NLP), word segmentation is necessary because there are by norm no spaces between Chinese words in each sentence. Lastly, we removed stop words with a self-defined stop-word dictionary.¹ The remaining words in each weibo post are stored in bag-of-words after further dropping the words that appeared in less than 15 posts (for limiting effects of rare domain vocabularies on model training) and in more than 50% of the posts (e.g. “Wuhan” and “aiding-Hubei”). We used the bag-of-words for topic modeling.

Topic Modeling

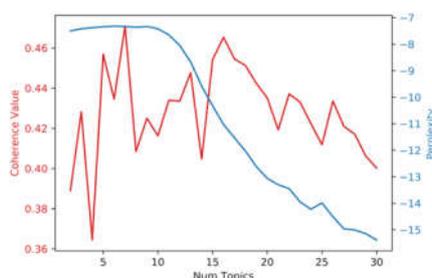
Using Latent Dirichlet Allocation (LDA) topic modeling, a machine-assisted statistical tool for identifying latent topics in a collection of documents (Blei et al., 2001), we sought to summarize an array of semantic themes in this online discourse about health workers amidst the public health crisis. Human interpretation, in conjunction with the machine-generated topics,

¹ Our stop-word dictionary is available upon request.

helped gain insights into the overall patterns of female health workers' social media representation during the COVID-19 pandemic.

To determine the appropriate number of topics for extraction, we calculated the coherence value and perplexity of each candidate LDA topic model with the number of topics ranging from 2 to 30 (see Figure 1). A higher coherence value indicates a better quality of topics judging the level of semantic coherence in topical concepts (Mimno et al., 2011). A lower perplexity indicates a better predictive power (Blei et al., 2001) when classifying an unseen document to the most suited topic. Judging both evaluation measurements in Figure 1, having 5-7 or 11-27 topics would be ideal for this dataset. Among these 20 models, the 13-topic model made the most sense to a human understanding and interpretation perspective, which is an essential part of the decision process. It also yielded a relatively high coherence value at 0.448 and a relatively low perplexity at -8.672.

Figure 1. *Topic Model Evaluation: Coherence Value and Perplexity by the Number of Topics*



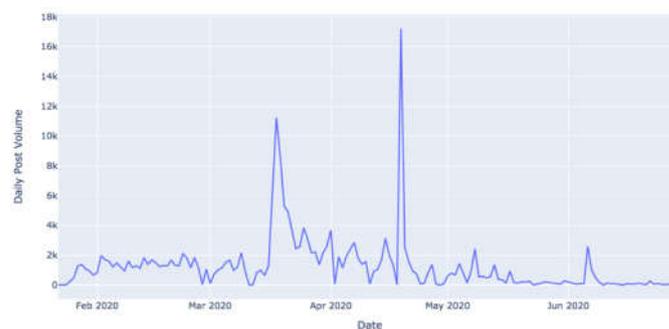
Results

Descriptive statistics

On the dataset ($n = 199,110$), we plotted the daily weibo post volume by date. See Figure 2. There are two posting activity peaks. On March 18, 2020, many microblogs ($n = 11,225$) were posted on Sina Weibo when aiding-Hubei medical teams started to leave Wuhan for their homes in other provinces of China. More posts ($n = 17,204$) surged on April 19, 2020 as the last batch of

aiding-Hubei medical teams started to pull out. Both peaks reflected respect, gratitude, and eulogy. However, posts of the first peak emphasized health workers' professionalism and bravery, as well as the nation's united efforts. Posts of the second peak, in comparison, paid more attention to families, sacrifice, and touching personal stories.

Figure 2. *Daily Weibo Post Volume by Date (January 22, 2020 and June 30, 2020)*



Here are a few exemplar posts from the first posting peak.

“The epidem steering group from the central government set off the national emergency medical support team. They said, ‘we need to ensure the work of aiding-Hubei medical teams carry out in an orderly, safe, and smooth manner.’”

“Today my social media feeds are ‘flooded’ with the good news that aiding-Hubei medical teams are gradually pulling out. Grateful to you all! It is your selfless devotion and efforts that enable our nation to control the pandemic. Thank you!”

“I’ve watched quite a few videos about aiding-Hubei medical teams pulling out. Cried to death! It is so touching that you couldn’t help but cry! China is really awesome.”

“#A Wuhan citizen dressed in a Pikachu costume seeing off aiding-Hubei medical teams# It is so touching that I couldn’t refrain myself from crying when I saw this. Love has always been two-directional. Your peaceful and happy life is possible because someone else is fighting and going through hardship for you. Grateful to them (note: ‘them’ was in a female pronoun)!”

Below are another two example posts from the second posting peak.

“After two months of separation, the nurse who is a mom burst into tears the moment she heard her daughter calling ‘mama’. Today the last batch of aiding-Hubei medical teams are returning home for a family reunion.”

“Hunan Province’s aiding-Hubei medical teams are returning home in triumph. Tens of thousands of people are seeing them off. The scene was touching.”

dimension of health workers. Overall, topic 5 (“ICU stories”) involved a slightly more personal tone in reporting on the frontline healthcare during the first surge of COVID-19, but many of its posts were still about hospital admission, patient cases, and hospital bed occupancy rates.

Table 1. Proportions, Labels, and Keywords of the 13 LDA Topics.

Topic	Proportion	Label	Top 10 Words
1	17.3%	professionalism	hospital, first, severe case, Wuhan, pneumonia, crew, team leader, team, process, date
2	12%	triumphant return	deliver support, salute, hero, white coat, medical professional, return home, personnel, Wuhan, weibo, thank
3	11.5%	homecoming gourmet	video, hotpot, weibo, crew, miaopai, Wuhan, Hubei Province, news, Anhui Province, date
4	10.6%	leadership (solicitude)	pandemic, give play to, effect, science, work, resist the epidemic, prevention and control, frontline, young people, spirit
5	10.2%	ICU stories	patient (“患者”), Wuhan, hospital, discharge, doctor, patient (“病人”), finally, ward, treat and cure, medical professional
6	7.4%	pickpocketing	personnel, medical professional, pickpocket, China, thief, activity, ceremony, notice, journalist, Shanxi Province
7	6.3%	Coronavirus’ physiological effects	UN, COVID-19, Beijing, nation, expert, infect, confirmed, virus, research, case
8	6%	gender role	nurse, daughter, husband, mom, child, Guangxi Province, wife, video, family, quarantine
9	4.5%	feminism	hospital, university, Shanghai, West China, Shandong Province, first, date, Dr. Jingjing Zhang, doctor, youth
10	3.9%	leadership (order)	nation, press conference, mechanism, joint defense, joint control, the State Council, central government, smoothly, module hospital, medicine
11	3.8%	celebrity solicitude (general)	present/gift, Jack Ma, everyday, medical professional, Meidi, video, Weibo, netizen, nurse, mask
12	3.8%	celebrity solicitude (Yibo Wang)	Wang Yibo, doctor, Peking University, Beijing, IQ, positive energy, bWang ^a , celebrity, up, bo-send ^a
13	2.6%	news coverage	link, webpage, Guangdong Province, resist the epidemic, return in triumph, crew, Jiangxi Province, Jilin Province, Tianjin City, headline

Note. Topics are ranked in a descending order by topic proportions (%).

^a These are errors from computational word segmentation; the original words were “wyb 王一博”, “王一博送.”

In contrast, when the female identity was explicit in posts, depictions tended to first emphasize the family role of women health workers, as parents, wives, daughters, and the persons who are expected to manage the household. This type of posts converged in topic 8 (“gender role”), the 8th largest topic with 6% tokens. Topic 8 posts depicted health workers, particularly nurses, in a family context. The top 10 words in the topic were “nurse”, “daughter”, “husband”, “mama”, “child”, “Guangxi”, “wife”, “video”, “family member”, and “quarantine” (i.e., “护士,” “女儿,” “丈夫,” “妈妈,” “孩子,” “广西,” “妻子,” “视频,” “家人,” and “隔离” in Chinese characters). As this group of posts revealed, women health workers’ professional role was discussed as secondary to their family role. One example message of topic 8 is as below, where the nurse’s role as the “mama” to her daughter was highlighted. This sample of public discourse celebrated her contribution to COVID-19 healthcare by celebrating her sacrifice of family life. The assumption of this logic implied a normative bias toward women that working mothers are eventually returning home for family life.

“After months of separation, the nurse who is a *mama* burst into tears the moment she heard her daughter calling ‘*mama*.’ Today the last aiding-Hubei medical teams are *returning home* for *family* reunions.”

In juxtaposition to the gender stereotypes associated with women health workers, posts in topic 9 (“feminism”) commented on the lack of women’s visibility in public discourse, citing city-specific data to show that women constitute the major forces in the frontline healthcare teams across different cities in China. These posts were primarily reposting with no commentaries, or short commentaries to a long chain of reposts. They called for more recognition and opportunities for women health workers and women in general. Topic 9 posts constituted 4.5% of all tokens, entailing typical posts such as “see them (the female pronoun),”

“Shanxi girls are awesome,” “ironic,” and “please never again deprive women of their workplace in the excuse of ‘women cannot endure hardship.’”

In addition to health workers, the topic modeling results also showed that during this period of the pandemic, some corporates and celebrities increased their public exposure on social media. This form of increased social media exposure was primarily made through posts celebrating the corporates’ and celebrities’ charitable activities such as donating goods and materials as well as food to the aid-Hubei medical teams (see topic 11 and 12). Specifically, topic 11 posts entailed corporates such as Alibaba Group and Midea Group and male celebrities such as Jack Ma, Yibo Wang, Bei’er Bao, and Wallace Chung Hon-Leung. Interestingly, the posts about Wallace Chung Hon-Leung, as included in this dataset, appeared to be relevant to topic 11 only because they included hashtags about Jack Ma’s charitable work instead of Chung’s. Much of these contents about Chung appeared to be generated by bot accounts upon our further inspection. Furthermore, topic 12 posts were almost exclusively about the charitable work of Yibo Wang, a male dancer, singer, and actor who are among the most popular celebrities in China as of this writing.

Discussion

The health sector has been faced with unprecedented pressure due to COVID-19. Numerous frontline health workers are making immeasurable contributions to the combat with this public health crisis, especially women (The Paper, 2020). According to findings of our study, we found a notable amount of visibility with women in this online discourse about the “aiding-Hubei medical teams.” However, this visibility takes place within a paradox. Specifically, topic 8 suggested normative gender stereotypes of women health workers that link them to their family roles, while topic 9 reflected on this public discourse through a critical lens,

advocating women rights. Both topics pushed the boundaries of women's agenda on social media. They challenged the dominant discourse we identified (i.e., topics 1, 2, 3, 4, 5, 10, 11, 12, and 13), which overlooked women through covering them in a collective identity. Still, in this sample of Weibo posts ($N = 199,110$), women were not only not representative enough but also more likely to be represented in a family role than a professional role. Their social media representation was disproportionate to their social participation as they constitute the major force among the frontline healthcare professionals (topic 9).

In addition, inspecting the two posting peaks in the time series, we found a large number of posts from the first peak emphasizing health workers' professionalism and bravery, as well as the nation's united efforts. Meanwhile, posts of the second peak delineated narratives around families, sacrifice, and touching personal stories. Posts in the observation period were underpinned by a collectivism- and marital-oriented metaphor as health workers fought against COVID-19. Echoing the topic modeling results, posting peaks implied two different ways of representing frontline health workers, one being the collective identity of "teams" and the other being "individuals." There was an overall tendency for Weibo posts to mention teams more often than individuals. Such a value-laden tendency may be rooted in a home-country-isomorphism ideology, which is the culture tradition in China (Jiang, 2012). With the influence of a home-country-isomorphism ideology, family serves as the micro unit of the nation, reinforcing the power of the country or the central (Jiang, 2012).

The normative gender stereotypes of occupations, as well, may be intertwined with this home-country-isomorphism ideology, feeding to the common assumption that men suit the professional role of health providers better than women because of women's "natural predilection" for family life (Brubaker, 2020). Further unpacking the effects of a home-country-

isomorphism ideology on women's professional recognition will be essential to gender equality in news reporting. In particular, findings of our study corroborated previous research in illustrating ways in which media report women with underrepresentation and biases (Tuchman, 2000). Female names and female-related topics are often censored and silenced except for such special cases as the passing-away of the female physician Jingjing Zhang during her time serving at the COVID-19 frontline (see topic 9). Some media scholars argued that such "symbolic annihilation" (Tuchman, 2000), where women are demeaned or canceled, would pose serious problems by playing an important role in maintaining the gendered imbalance in the power structure (De Swert & Hooghe, 2010).

Echoing Courtney et al. (2020), this study suggested that media coverage about women would not necessarily increase proportionally as women advance in public life, such as the medical field. News reports on social media, as well as the corresponding social media discussions, still tended to demeaning women if not completely silencing or cancelling them in the public sphere (Lundell & Ekstroumlm, 2008). Gender disparities in both economic participation and censorship can lead to unequal treatment in occupations (Blumell, 2020), including gender discrimination in positions in different industries (Kilbourne, 2012) and inequality in salary distribution (Lips, 2013). The legacy media have played a critical role in society as the "fourth estate" in the power system (Powe, 1992). On legacy media, especially printed news media, women are substantially underrepresented (Shor et al., 2015), for that "the media focuses nearly exclusively on the highest strata of occupational and social hierarchies, in which women's representation has remained poor" (Shor et. al, 2015, p. 960). Differential media visibility may reinforce long-standing gender status beliefs and serve as the manifestations of these beliefs. Therefore, legacy media and digital media both have the obligation to reflect on

their professional conduct in news reporting and ensure accurate representation of women to foster inclusive and equal media and social environments.

Moreover, as mentioned in the results section, we found that corporates and celebrities used Weibo as a means to increase their online presence during the first stage of the COVID-19 pandemic. A large number of posts in our sample reported on the charitable work of celebrities, especially male celebrities, for the aid-Hubei medical teams. A close inspection of this phenomena revealed the online participation of fans and we media in populating such content. In reflection of Bourdieu's theory of economic, social and cultural capital (Bourdieu 1986), this phenomena may be an example of the conversion of social capitals to material capitals in an online public sphere, where "social capitals" mean the "aggregate of the actual or potential resources which are linked to the possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition" (Bourdieu 1986, p. 248).

On the one hand, celebrities with a large fan base may potentially help the disaster-relieving actions to gain more public attention and encourage more charitable activities to mobilize social resources for frontline health workers in face of COVID-19. In this process, social media helped public relationships to evolve from one-way information dissemination to two-way interactions that involve multiple stakeholder voices (Sanderson et al., 2016). On the other hand, previous research has pointed out that even the well-intentioned public relations efforts can quickly go wrong as they spread on social media, as these efforts can deviate from the expected goal (Chewning, 2015).

In this study, for example, the mass populating of corporate- and celebrity-related content appeared to be no more than performative in essence. The celebratory nature of these posts toward corporates and celebrities – rather than toward the frontline health workers – distracted

the public attention, took away the voices of frontline health workers, and contributed to a hyper-reality of the pandemic that distorted what was really going on at the COVID-19 frontline and which nonetheless came to constitute reality. Therefore, this type of corporate and celebrity behavior constitutes social media wave-riding and hashtag-jumping/hijacking (Christensen, 2013), and it can be detrimental to the information management during a public health crisis. Even in the social media wave-riding and hashtag-jumping/hijacking in our study, we observed almost exclusive male figures; the women health workers are further marginalized. The role of Weibo in exacerbating the underrepresentation and misrepresentation of women has to be further critically examined through future research.

Limitations

This study used only one keyword, “援鄂医疗队” (i.e. “aiding-Hubei medical team”), which resulted in a limited sample of 199,110 Weibo posts. Although it was not a small sample, a larger corpus using more diverse keywords may provide richer insights into the public discourse about frontline health workers during the COVID-19 pandemic. Also, the observation period was limited to the first stage of the COVID-19 pandemic between January 1, 2020 and June 30, 2020. Findings of our research can only be interpreted within the context of this observation period; extrapolation of findings should be made with caution.

Conclusion

Women health workers were underrepresented and represented with biases on the Chinese social media of Weibo during the first stage of the COVID-19 pandemic. Despite that women health workers comprised over 60% of the task forces at the COVID-19 frontline, they received no sufficient recognition nor fair reporting for their actual professional participation. Meanwhile, the identity of women health workers were often in the shadow of a collective

identify. Due to the value of familism and normative gender roles, gender and occupation stereotypes of women health workers persisted despite their major contributions to the healthcare sector. Not only the family and the nation but also the corporate and the celebrities participated the deprivation of women's voices and visibility in this particular online public sphere. From "a paper ceiling" (Shor et al., 2015) to "a web ceiling," media representation of working women was still limited and biased. Actions, through research and practices, must be taken to confront the factors associated with the underrepresentation and misrepresentation of women in media.

References

- 21st Century Business Herald. (2020, March 8). Raise hygiene products for frontline female health workers in Hubei.
<https://m.21jingji.com/article/20200308/herald/44b1c4c43cfb19ea0a61049c00979325.html?from=weibo>.
- Basow, S. A. (1992). *Gender stereotypes and roles*. Pacific Grove, CA: Brooks.
- Blei, D., Ng, A., & Jordan, M. (2001). Latent dirichlet allocation. *Advances in Neural Information Processing Systems*, 14, 601-608. Retrieved November 18, 2020 from <https://papers.nips.cc/paper/2001/file/296472c9542ad4d4788d543508116cbc-Paper.pdf>
- Blumell, L. E. (2020). Bro, foe, or ally? Measuring ambivalent sexism in political online reporters. *Feminist Media Studies*, 20(1), 53-69.
<https://doi.org/10.1080/14680777.2018.1546211>
- Bourdieu, P. (1986). The Forms of Capital. In J. Richardson (ed), *Handbook of Theory and Research for the Sociology of Education*. New York: Greenwood.
- Brubaker, L. (2020). Women physicians and the COVID-19 pandemic. *JAMA*, 324(9), 835-836.
<https://www.doi.org/10.1001/jama.2020.14797>
- Carter, C. (2011). Sex/gender and the media. In K. Ross (Ed.), *The Handbook of Gender, Sex, and Media* (pp. 363–382). Malden, MA: Wiley-Blackwell.
- Chen, F. (2004). The division of labor between generations of women in rural China. *Social Science Research*, 33(4), 557-580. <https://doi.org/10.1016/j.ssresearch.2003.09.005>
- Cheung, F. M. (1996). Gender role development. *Growing up the Chinese way: Chinese child and adolescent development*, 45-67.

- Chewning, L. V. (2015). Multiple voices and multiple media: Co-constructing BP's crisis response. *Public Relations Review*, *41*(1), 72-79.
<https://doi.org/10.1016/j.pubrev.2014.10.012>
- Christensen, C. (2013). Wave-riding and hashtag-jumping: Twitter, minority 'third parties' and the 2012 US elections. *Information, Communication & Society*, *16*(5), 646-666.
<https://doi.org/10.1080/1369118X.2013.783609>
- Courtney, M., Breen, M., McGing, C., McMenemy, I., O'Malley, E., & Rafter, K. (2020). Underrepresenting reality? Media coverage of women in politics and sport. *Social Science Quarterly*, *101*(4), 1282-1302. <https://doi.org/10.1111/ssqu.12826>
- De Swert, K., & Hooghe, M. (2010). When do women get a voice? Explaining the presence of female news sources in Belgian news broadcasts (2003—5). *European Journal of Communication*, *25*(1), 69-84. <https://doi.org/10.1177/0267323109354229>
- DeWees, M. A., & Miller, A. C. (2020). The costs of care: A content analysis of female nurses' media visibility and voices in the United States, China, and India during the COVID-19 pandemic. In *COVID-19* (pp. 221-233). Routledge.
- Esarey, A., & Xiao, Q. (2011). Digital communication and political change in China. *International Journal of Communication*, *5*, 22.
<https://ijoc.org/index.php/ijoc/article/view/688/525>
- European Institute for Gender Equality. (2020). *Essential workers*. Retrieved Dec 30, 2020 from <https://eige.europa.eu/covid-19-and-gender-equality/essential-workers>
- Gallagher, M. (2002, November). Women, media and democratic society: In pursuit of rights and freedoms. In *Expert group meeting on "participation and access of woman to the media,*

and the impact of media on, and its use as an instrument for the advancement and empowerment of women.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, *115*(16), E3635-E3644. <https://doi.org/10.1073/pnas.1720347115>

Hust, S. J., & Brown, J. D. (2008). Gender, media use, and effects. *The handbook of children, media, and development*, 98-120.

Huyck, M. H. (1999). Gender Roles and Gender Identity in Midlife. In Sherry L. Willis & James D. Reid (Eds.), *Life in the middle: Psychological and social development in middle age* (pp. 209-232). Academic Press. <https://doi.org/10.1016/B978-0-12-757230-7.X5017-2>

Jones, R. (2020, July 24). Women are on the frontlines in the fight against COVID-19. *National Geographic*. <https://www.nationalgeographic.com/history/2020/07/women-on-frontlines-fight-against-covid-19/>

Kilbourne, J. (2012). *Can't buy my love: How advertising changes the way we think and feel*. Simon and Schuster.

Lai, J., Ma, S., Wang, Y., Cai, Z., Hu, J., Wei, N., ... & Hu, S. (2020). Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. *JAMA Network Open*, *3*(3), e203976-e203976. <https://doi.org/10.1001/jamanetworkopen.2020.3976>

Li, Q. (2018). Numbers of men in nursing are rising in China. *John Hopkins Nursing*. Retrieved on February 26, 2020 from <https://magazine.nursing.jhu.edu/2018/06/numbers-of-men-in-nursing-are-rising-in-china/>

Lips, H. M. (2013). Acknowledging discrimination as a key to the gender pay gap. *Sex Roles*, *68*(3), 223-230. <https://doi.org/10.1007/s11199-012-0245-0>

- Lundell, Å. K., & Ekström, M. (2008). The complex visual gendering of political women in the press. *Journalism Studies*, 9(6), 891-910. <https://doi.org/10.1080/14616700802227845>
- Luo, R., Xu, J., Zhang, Y., Ren, X., & Sun, X. (2019). *Pkuseg: A toolkit for multi-domain chinese word segmentation*. arXiv preprint arXiv:1906.11455.
- Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011, July). Optimizing semantic coherence in topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 262-272). <https://www.aclweb.org/anthology/D11-1024.pdf>.
- Niven, D. (2005). Gender bias? Media coverage of women and men in Congress. *Gender and American politics: Women, men, and the political process*, 264-283.
- Powe, L. A. (1992). *The fourth estate and the constitution: Freedom of the Press in America*. University of California Press.
- Ross, K. (2011). Silent Witness: News sources, the local press and the disappeared woman. In C. Alvares, T. Krijnen and S. Van Bauwel (Ed.), *Gendered transformations. Theory and practices on gender and media* (pp. 20). Brussels: European Consortium of Political Research (ECREA).
- Sanderson, J., Barnes, K., Williamson, C., & Kian, E. T. (2016). ‘How could anyone have predicted that# AskJameis would go horribly wrong?’ Public relations, social media, and hashtag hijacking. *Public Relations Review*, 42(1), 31-37. <https://doi.org/10.1016/j.pubrev.2015.11.005>
- Scharrer, E. L. (2013). Representations of gender in the media. In K. E. Dill (Ed.), *The Oxford handbook of media psychology* (pp. 267–284). Oxford University Press.

- Shor, E., Van De Rijt, A., Miltsov, A., Kulkarni, V., & Skiena, S. (2015). A paper ceiling: Explaining the persistent underrepresentation of women in printed news. *American Sociological Review*, *80*(5), 960-984. <https://doi.org/10.1177/0003122415596999>
- Sjøvaag, H., & Pedersen, T. A. (2019). Female voices in the news: Structural conditions of gender representations in Norwegian newspapers. *Journalism & Mass Communication Quarterly*, *96*(1), 215–238. <https://doi.org/10.1177/1077699018789885>
- Smith, J. C. (2020). Where are the women? Descriptive representation and covid-19 in UK Daily Press Briefings. *Politics & Gender*, *16*(4), 991-1000. <https://doi.org/10.1017/S1743923X20000513>
- Steele, J. R., & Brown, J. D. (1995). Adolescent room culture: Studying media in the context of everyday life. *Journal of Youth and Adolescence*, *24*(5), 551-576. <https://doi.org/10.1007/BF01537056>
- Stoller, R. J. (2020). Sex and gender: The development of masculinity and femininity. Routledge.
- Tezel, A., Griffith, K. A., Jones, R. D., & Jagsi, R. (2021). Diversity and Representation of Physicians During the COVID-19 News Cycle. *JAMA Internal Medicine*, *181*(1), 124-127. <https://doi.org/10.1001/jamainternmed.2020.6285>
- The Paper. (2020, September 26). *28 thousand female medical professionals: Women power in combating COVID-19 should not be forgotten* [2.8 万名女性医务人员 : 抗疫中的女子力不应被遗忘]. https://www.thepaper.cn/newsDetail_forward_9359612
- The Union for the Mediterranean. (2020). *Women on the frontline of Covid-19*. <https://ufmsecretariat.org/woco-women-frontline-covid/#:~:text=Women%20are%20at%20the%20forefront,their%20lives%20to%20save>

%20others.&text=Today%20women%20account%20for%2070,to%20around%205%20billion%20people.

Triana, H. W., Kustati, M., Yusuf, Y. Q., & Reflinaldi, R. (2021). The representation of women in COVID-19 discourses: The analysis of Sara Mills' critical discourse on media coverage. *Journal of Language and Linguistic Studies*, 17.

<https://dergipark.org.tr/en/download/article-file/1663181>

Tuchman, G. (2000). The symbolic annihilation of women by the mass media. In *Culture and politics* (pp. 150-174). Palgrave Macmillan, New York.

Tyson, P. (1982). A developmental line of gender identity, gender role, and choice of love object. *Journal of the American Psychoanalytic Association*, 30(1), 61-86.

<https://doi.org/10.1177/000306518203000103>

UN Women. (2020). *Guidelines for gender and conflict-sensitive reporting*.

<https://www2.unwomen.org/->

[/media/field%20office%20eca/attachments/publications/2019/07/guidelines%20for%20media%20and%20gender%20ukraine/guidelines_eng_prew_40719_compressed.pdf?la=en&vs=3351](https://www2.unwomen.org/-/media/field%20office%20eca/attachments/publications/2019/07/guidelines%20for%20media%20and%20gender%20ukraine/guidelines_eng_prew_40719_compressed.pdf?la=en&vs=3351)

Wood, J. T. (1994). Gendered Media: The Influence of Media on Views of Gender. In Julia T. Wood, *Gendered Lives: Communication, Gender, and Culture* (pp. 231-244). Wadsworth Publishing. <http://www1.udel.edu/comm245/readings/GenderedMedia.pdf>

World Health Organization. (n.d.). *Value gender and equity in the global health workforce*.

<https://www.who.int/activities/value-gender-and-equity-in-the-global-health-workforce>

Jiang, Y. (2012). "Home-country-isomorphism" and the double construction of women's gender role – The nation and the women during the planned economy period of China. *Journal*

of Shandong Women's University, 1, 1-6. [蒋永萍。(2012)。“家国同构”与妇女性别角色的双重建构——计划经济时期中国社会的国家与妇女。《山东女子学院学报》2012年第1期，1-6页。]

Appendix A. Word Frequency and Word Cloud

Table A1. Top 50 Frequent Words of the Whole Topic-related Microblogs (N = 199,110)

Rank	Word	Frequency	English Translation
1	医疗队 [†]	320797	medical team
2	援鄂 [†]	262614	aiding-Hubei
3	武汉	129137	Wuhan
4	医院	125718	hospital
5	视频	97164	video
6	人员	85723	personnel
7	医护	74542	medical professional
8	湖北	72356	Hubei
9	队员	71303	crew
10	疫情	69038	pandemic
11	微博	56617	weibo
12	月日	54825	month/date
13	工作	44715	work
14	国家	43812	nation
15	医生	42419	doctor
16	患者	42334	patient
17	人民	41500	people
18	回家	39523	return home
19	抗疫	39273	resist the epidemic

20	医疗	38902	medical treatment
21	王一博	37538	Wang Yibo
22	支援	33626	aid
23	英雄	32941	hero
24	肺炎	32567	pneumonia
25	护士	32179	nurse
26	大学	30612	university
27	第一	30034	first
28	防控	25360	prevention and control
29	感谢	25322	thank
30	一线	25170	front line
31	致敬	25102	salute
32	凯旋	23856	return in triumph
33	新闻	23337	news
34	重症	23172	severe case
35	战疫	22576	fight the epidemic
36	礼物	22273	present/gift
37	北京	21741	Beijing
38	医务	21704	medical
39	撤离	21412	evacuate
40	救治	20810	treat and cure

41	链接	20792	link
42	医学	20593	medicine
43	上海	20355	Shanghai
44	平安	19923	safe/safety
45	秒拍	19663	Miaopai (a video app)
46	网页	18959	webpage
47	白衣	18933	white coat
48	紧急	18880	urgent
49	送行	18524	see off
50	中国	18349	China
† Words that were search terms and hence were excluded from topic modeling.			

Appendix B. Topic Keywords in the Original Language

Table B1. Top 10 Frequent Words in Chinese in Each Topic of the 13-Topic LDA Model[†]

Topic (%)	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Word 10
T1 (17.3%)	医院	第一	重症	武汉	肺炎	队员	队长	团队	过程	月日
T2 (12%)	送援	致敬	英雄	白衣	医护	回家	人员	武汉	微博	感谢
T3 (11.5%)	视频	火锅	微博	队员	秒拍	武汉	湖北	新闻	安徽	月日
T4 (10.6%)	疫情	发挥	作用	科学	工作	抗疫	防控	一线	青年	精神
T5 (10.2%)	患者	武汉	医院	出院	医生	病人	终于	病房	救治	医护

T6 (7.4%)	人员	医护	扒窃	中国	扒手	活动	仪式	通报	记者	山西
T7 (6.3%)	UN	新冠	北京	国家	专家	感染	确诊	病毒	研究	病例
T8 (6%)	护士	女儿	丈夫	妈妈	孩子	广西	妻子	视频	家人	隔离
T9 (4.5%)	医院	大学	上海	华西	山东	第一	月日	张静静	医生	青春
T10 (3.9%)	国家	发布会	机制	联防	联控	国务院	中央	顺利	方舱	医学
T11 (3.8%)	礼物	马云	天天	医护	美的	视频	微博	网友	护士	口罩
T12 (3.8%)	王一博	医生	北大	北京	IQ	正能量	b 王 ^a	艺人	向上	博送 ^a
T13 (2.6%)	链接	网页	广东	抗疫	凯旋	队员	江西	吉林	天津	头条
<p>† Vertically, topics are ranked in a descending order by topic sizes. Each row lists the top 10 words of the given topic.</p> <p>^a These are two errors from computational word segmentation; the original words were “wyb 王一博” and “王一博送*.”</p>										