“Who Knows? Maybe it Really Works”: Analyzing Users’ Perceptions of Health Misinformation on Social Media

Huiyun Tang
University of Luxembourg,
Esch-sur-Alzette, Luxembourg
huiyun.tang@uni.lu

Gabriele Lenzini
SnT,University of Luxembourg,
Luxembourg, Luxembourg
gabriele.lenzini@uni.lu

Samuel Greiff
University of Luxembourg,
Esch-sur-Alzette, Luxembourg
samuel.greiff@uni.lu

Björn Rohles*
Digital Learning Hub; Ministère de
l’Éducation nationale, de l’Enfance et
de la Jeunesse, Luxembourg
Esch-sur-Alzette, Luxembourg
info@rohles.net

Anastasia Sergeeva*
University of Luxembourg,
Esch-sur-Alzette, Luxembourg
anastasia.sergeeva@uni.lu

ABSTRACT

Health misinformation, defined as health-oriented information that contradicts empirically supported scientific findings, has become a significant concern on social media platforms. In response, platforms have implemented diverse design solutions to block such misinformation or alert users about its potential inaccuracies. However, there is limited knowledge about users’ perceptions of this specific type of misinformation and the actions that are necessary from both the platforms and the users themselves to mitigate its proliferation. This paper explores social media users’ (n = 22) perceptions of health misinformation. On the basis of our data, we identify specific types of health misinformation and align them with user-suggested countermeasures. We point to the critical demands for anti-misinformation solutions for health topics, emphasizing the transparency of information sources, immediate presentation of information, and clarity. Building on these findings, we propose a series of design recommendations to aid the future development of solutions aimed at counteracting misinformation.

CCS CONCEPTS

• **Human-centered computing → User studies**; *Web-based interaction*; *Empirical studies in HCI*.

KEYWORDS

Health misinformation, Credibility indicators, Design friction, Reactance, Social media

ACM Reference Format:
Huiyun Tang, Gabriele Lenzini, Samuel Greiff, Björn Rohles, and Anastasia Sergeeva. 2024. “Who Knows? Maybe it Really Works”: Analyzing Users’ Perceptions of Health Misinformation on Social Media. In *Designing Interactive Systems Conference (DIS ’24), July 01–05, 2024, IT University of Copenhagen, Denmark.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3643834.3661510

1 INTRODUCTION

Different types of social media have made it easier for people to communicate and share information about different topics, including health [88]. Social media users often discuss their experiences with health issues, offer remedies, or provide tips for a healthier lifestyle [28, 95]. However, such information is usually based on personal opinions, is not backed by scientific evidence, and is sometimes influenced by commercial interests, such as the promotion of health supplements through native advertising. Such practices can lead to spreading health misinformation. Broadly speaking, health misinformation is defined as any health information that is not currently supported by scientific evidence, including health rumors, misleading advertising, fake news, and conspiracy theories about health [131]. Moreover, misinformation is spread without the intention to cause harm [102, 132] (see Section 2.2). Thus, the landscape of health misinformation varies in its degree of inaccuracy and its potential risks for individuals and society.

There is growing consensus among the scientific community and social media platforms that the dissemination of health misinformation must be curtailed [30, 135], but it is not clear what criteria would justify interventions in user communication. Whereas several studies have analyzed the effects of such interventions on the spread of misinformation, there is a gap in the understanding of users’ perceptions of which actions social networks should take to combat health misinformation. Gaining insight into users’ views on these actions and the harms of health misinformation is crucial for developing more effective and acceptable strategies.

In this qualitative study, we analyzed users’ perceptions of the types and characteristics of health misinformation. We emphasize that it is important for social media platforms and other stakeholders (e.g., users, government agencies, medical societies, and international health organizations) to intervene in order to protect users from this misinformation. Our research offers insights into general strategies that might be implemented (e.g., identifying the factors that should prompt social media platforms to counteract

---

*Both authors contributed equally to this research.

DIS ’24, July 01–05, 2024, IT University of Copenhagen, Denmark

ACM ISBN 979-8-4007-0583-0/24/07
https://doi.org/10.1145/3643834.3661510

specific misinformation) and delves into users' opinions about the design strategies that are currently in place across social media platforms. Thus, our research will help others gauge the adequacy and effectiveness of existing design solutions to prevent misinformation regarding health-related content. Furthermore, we focused on investigating technology in context by applying a scenario-based research approach to fully understand humans' experiences with technology. Our approach allowed us to identify user needs and derive recommendations for the design of interactive technologies.

This paper makes four contributions:

1. We examine users' perceptions of what distinguishes health information from health misinformation on social media;
2. we outline the distinctions between health misinformation and other types of misinformation, highlighting the unique challenges individuals face when trying to verify and debunk health-related claims;
3. we explore users' views on what actions they believe are appropriate for networks to undertake and delve deeper into the factors that may induce resistance to countermeasures against misinformation on these platforms;
4. we suggest potential practices that designers can engage in to create tools for navigating health misinformation.

Our study primarily centers on health misinformation. However, our findings also shed light on the broader discourse concerning how users perceive the ways in which social media platforms address misinformation in general. The results capture users' nuanced views on the balance between freedom of choice and expression versus the wish to protect others from the immediate and long-term impacts of misinformation.

2 RELATED WORK

2.1 Health-related Information in Social Media

Many definitions of social media exist. A systematic literature review found that, in earlier research (until around 2010), social media centered around the idea of connecting people with common interests, whereas in later research, the defining aspect of social media changed to the sharing of user-generated content [4]. Both of these aspects are relevant for the present paper, as health-related information is shared in both private (e.g., messaging among family members) and public settings (e.g., publicly posting one's health experiences). Indeed, social media platforms, such as Facebook, Twitter, [1] TikTok, and YouTube, are frequently utilized to disseminate health-related information. In a study by Marar et al., 85% of hospital patients and their relatives reported seeking health information on social media [83]. Similarly, Bryan et al. found that 68% of parents in pediatric units reported using social media to find health-related information. During the COVID-19 era, a large-scale survey of the U.S. general population revealed that 76% of participants depended on social media for at least "a little information." Furthermore, 59.2% stated they accessed COVID-19-related information on social media at least once a week[96]. Prior to the COVID-19 pandemic, up to 74% of online health information seekers utilized social media [127]. Social media serves a different purpose than search engines. For instance, De Choudhury et al. demonstrated that people often use social media to understand the effects of health conditions on daily life and to gain a broader understanding of medical procedures [36]. This corresponds to Chu et al.'s finding that using search engines can create frustration because the information on medical sites can be too difficult to understand, may offer too much detail, and may lack enough criteria to evaluate it [31]. Personal stories and the opportunity to discuss issues in everyday language can motivate people to use social media for health information [112]. Moreover, social media allows users to seek social and emotional support through peer interactions, especially from those with similar health challenges [141, 143]. Other advantages include more understandable and personalized information [90]. However, the prevalence of content created by non-professionals raises concerns about the reliability of shared health information [20, 32, 83, 90].

2.2 Characteristics of Health Misinformation on Social Media

Many terms exist to describe the topic of the present paper, such as misinformation, disinformation, malinformation, fake news, false news, or clickbait [66, 145]. Zhou and Zafarani suggested three dimensions to define these terms: authenticity (from true to false), intention of sharing (from good to bad), and type of information (news or not) [145]. Based on the framework of information disorder, three distinctions arise: false information deliberately created and disseminated to cause harm (disinformation), genuine information such as leaks spread intentionally to cause harm (malinformation), and false information spread without intention to cause harm (misinformation) [102, 132] . In the current paper, we focus on the dimension of authenticity without intending harm. Therefore, we intentionally adopt the broad term "health misinformation" which has been described as an umbrella term that includes alternative terms and refers to all false health-related claims [130].

Several studies have identified dangerous outcomes from health misinformation [49, 62, 75]. Health misinformation leads to an increase in misleading or incorrect interpretations of evidence, has a detrimental impact on mental health, results in a misallocation of health resources, fosters vaccination hesitancy, and fuels conspiracy beliefs [37]. One example is that during the COVID-19 era, misinformation, disinformation, and conspiracy theories proliferating on social media have eroded trust in government and health systems[78]. The surge in health misinformation not only delays care provision but also contributes to the proliferation of hateful and divisive health rhetoric in public discourse [37]. However, only a few studies have compared health misinformation with other types of misinformation (e.g., in politics). This scarcity of research makes it difficult to protect users specifically from health misinformation [128]. Such health misinformation is particularly important to address, as it might have severe consequences (e.g., when a person is discouraged from seeing a doctor due to health misinformation). Recent studies have indicated that, compared with other types of misinformation, health misinformation tends to have a longer life cycle, possibly because it is linked to ongoing personal health issues rather than short-lived social events [92]. Moreover, these studies have shown that, unlike in the political news domain

---

[1] The network has recently been rebranded as X. However, during our interviews, participants referred to it as "Twitter" based on their past experiences with the platform. For the sake of consistency in this work, we'll continue to use the name "Twitter."

where partisanship is more intense, appeals to scientific organizations and medical authorities can debunk erroneous beliefs in the health domain [89, 118, 124].

Another characteristic of health-related information is that it is often presented through personal health stories, motivated by the need to find people with shared experiences and the desire for social support or, to some extent, for self-presentation purposes [97, 113, 122]. Consequently, health-related information in social media often includes opinions and personal experiences [63]. These are difficult to mark as "true" or "fake," even when any treatments that are mentioned contradict expert consensus. For example, if a social media platform marks such content as "fake," the individuals who shared their experiences might perceive such labeling as a personal attack. Another issue is that health information and misinformation sometimes include genuine medical terms and concepts, thus making it challenging for non-experts to fully understand and make appropriate opinions about the content [26]. In this context, the knowledge gap stops users from understanding the false claims in the text. Alternatively, when presented via social media to nonprofessionals, even genuine medical information can generate claims about misinformation because nonprofessional users cannot fully understand the concepts or the relationships behind the information [73, 142]. These unique features of health misinformation and the severe consequences of incorrect statements made us question whether general methods used by platforms to address misinformation should also be used in the context of health misinformation. Alternatively, are other design solutions needed? In order to contribute to answering this question, we next describe current design solutions before investigating users' experiences with health misinformation.

2.3 Design Solutions for Counteracting Misinformation in Social Media

Several approaches are available for combating the spread of misinformation. Hartwig et al. developed a taxonomy of interventions based on high-level concepts such as intervention design, user interaction, and the timing of the intervention [58]. Aghajari et al. outlined four primary strategies: debunking false information through fact-checkers, signaling content credibility, directing an individual's attention toward content credibility, and reducing the visibility of deceptive content [3]. On the other hand, Goldman delineated five measures for addressing misinformation about individual content: remove content, suspend content, relocate content, edit/redact content with an interstitial warning, add a warning legend, add counterspeech, or disable comments [51].

In this section, we primarily adopt Aghajari et al.'s taxonomy [3]. However, we consolidate different means of indicating credibility, such as straightforward warnings and fact-checker warnings, under the "credibility indicators" category. We also introduce "inducing-reflection-before-sharing nudges" as a separate category.

2.3.1 Credibility Indicators. These strategies reduce false beliefs by offering factual information and explaining why a particular piece of misinformation is erroneous or misleading or by discrediting the source of the misinformation [38, 137]. They are designed to help users make informed decisions about media. These interventions can take the form of labels, provenance indicators, and warnings. Major social media platforms, such as Facebook, Twitter, and TikTok, collaborate with third-party fact-checkers to evaluate online content, subsequently attaching warning labels to content that is deemed false [14, 77, 105]. Similarly, YouTube introduced health source information panels in videos to help viewers identify content from reputable sources [52]. Human-computer interaction(HCI) researchers have explored diverse approaches to identify the credibility of online content. For instance, Bunde et al. incorporated a credibility rating scale and presented the source of credibility [21]. Amin et al. proposed labels indicating credible fact-checking sources and pop-up warning notifications [7]. Several works have proposed that provenance warnings should alert users when the source of information is unverified [44, 110].

2.3.2 Nudges and Frictions. In addition to addressing credibility, researchers have explored design interventions to raise awareness and promote critical news consumption behavior, with the goal of reducing the spread of misinformation on social media [107]. Such interventions can take the form of nudges and friction. Nudges are designed to guide individuals in a specific direction without restricting their freedom of choice [60]. For example, Jahanbakhsh et al. developed behavioral nudges that provide users with assessments and explanations of the accuracy of news, thereby reducing the dissemination of false content [65]. Similarly, Twitter has implemented friction to curtail the rapid sharing of post. When a user attempts to retweet a link without first opening the linked article, Twitter prompts them with a message asking if they would like to open the article before retweeting it [48]. TikTok has also implemented similar mechanisms, encouraging users to pause and reconsider sharing flagged unverified content [59]. Additionally, Andi and Akesson investigated using "social norm-based nudges" to encourage individuals to share more responsibly [8]. These nudges remind users that false news is prevalent online and that responsible individuals think twice before sharing information with friends and followers. Chen et al. focused on nudging and boosting users to resist misinformation by exposing them to the presence of filter bubbles and presenting news beyond those bubbles [27]. Capraro and Celadin examined three types of nudges and discovered that endorsing accuracy reduced the sharing of misinformation while increasing the sharing of genuine news [23].

2.3.3 Reducing Content Visibility. While the removal of content is hotly debated and often viewed as a form of "censorship" [13, 50], and people tend to reject authoritative retraction [125], platforms have implemented various strategies to reduce the visibility of content deemed misleading, for example, by crowd ratings or deep learning algorithms [40, 85, 129]. Instead of eradicating the content entirely, the strategy revolves around excluding it from recommendation systems, thereby reducing the likelihood of general users coming across it. However, research has indicated that while effective in general, this approach does not deter the sharing of conspiracy-themed content within conspiracy-focused communities [22]. Moreover, this method can inadvertently fuel misconceptions about the platform's hidden intentions, heightening uncertainty about the status of the content [50].

2.4 Efficiency of the Correction Measures and Reactance Theory

Whereas most studies have reported a positive effect of existing interventions, such as warning labels, nudges, and reduced content visibility can mitigate misinformation on social media [22, 70, 138], the effectiveness of these measures varies. [98, 116]. Furthermore, findings on the extents to which the interventions can cause backfiring (which means reinforcing misinformation dissemination) have been inconsistent [39, 103]. In the context of health misinformation, certain interventions, such as fact-checking, have demonstrated effectiveness [128]. However, fact-checking can be time-consuming and might not cover all language barriers. As reported, a mere 16% of health misinformation on Facebook came with warning labels flagging false or misleading content [11]. Moreover, approximately 70% of misleading COVID-19-related content in Spanish and Italian did not feature warning labels on Facebook, in contrast to the 29% in English [12]. Studies have also shown that even though fact-checking and real-time corrections may initially help correct misinformation, the impact tends to fade with time. Because misbeliefs can be sticky, fact-checking corrections often fall short of fully dispelling persistent misperceptions [25].

Reactance theory offers one way to understand these mixed outcomes. The term *reactance* refers to the uncomfortable feeling a person has when perceiving that their freedom is being threatened. Reactance involves both cognitive and emotional responses, prompting behaviors aimed at regaining one's sense of autonomy [19]. Previous studies have found that psychological reactance plays a significant role in how people cope with attempts to correct health misinformation on social media [74, 87, 115]. Specifically, it affects the effectiveness of fact-checking corrections and misinformation warning labels [99, 109, 140]. To address these reactance problems, various sources have suggested using narrative-based approaches [29, 33] and leveraging other social media users and health authorities as a means of rebuttal [115].

3 RESEARCH QUESTIONS

Previous research has found that health misinformation on social media often becomes intertwined with genuine health information. Such a mixture of information tends to happen when users want to share their personal experiences, seek community support, or find relatable first-person information about health. Furthermore, health misinformation spans a spectrum, ranging from slightly inaccurate statements to outright falsehoods. It is thus essential to understand how users identify the factors that make them perceive certain health information as "fake". Additionally, to design effective interventions that can address the negative impacts of health misinformation, we must align users' perceptions of the harm that misinformation can cause with their views on suitable solutions. Previous studies have often focused on how social media platforms tackle misinformation in general and have not necessarily zeroed in on health misinformation. As we have shown in the Related Works section, poorly chosen solutions can cause users to push back against interventions [69], making it highly important to choose appropriate measures against specific kinds of misinformation. Our study has two research questions, one about users' perceptions of health misinformation and one about the handling of health misinformation. For reasons of clarity, we split the latter into two subquestions:

- RQ1: What are users' perceptions of health misinformation on social media, and how do they describe its characteristics?
- RQ2.1: How do users view the responsibilities and actions of social networks regarding health misinformation?
- RQ2.2: What are the drawbacks of current design solutions for addressing health misinformation?

4 METHODOLOGY

We conducted a lab-based study with 22 participants (12 female, 10 male; demographics in the Appendix,Table 2). We recruited participants via a university's student forum and by distributing posters in the local community. We had three eligibility criteria: (a) age 18 or over, (b) personal experience with social media platforms (e.g., Facebook, Twitter, Instagram, WhatsApp, TikTok, YouTube), (c) proficiency in English. Considering the increasing tendency for individuals to seek health-related information on social media, particularly in the aftermath of the COVID-19 pandemic, we did not restrict participation to those with a history of seeking health-related information online. Instead, we hypothesized that using multiple social media platforms would increase exposure to both health information and misinformation, as confirmed in our pretest. Therefore, we selected participants who were active on at least two social media platforms. Our analysis showed that 17 participants used more than four platforms, and three used three platforms regularly. All participants were employees or students of a European University but had diverse cultural backgrounds, including East, South, and West Asia, Sub-Saharan Africa, Eastern and Western Europe, and Latin America.

4.1 Research Ethics

The study followed the ethical guidelines of the authors' institution and received approval from its Ethical Review Panel. Before each session, participants were briefed about the study's objectives and data collection. Given that the study pertained to the potentially sensitive subject of health, we highlighted this fact in our recruitment materials and in a prior informational meeting. Participants were advised not to participate if they found discussions about human health too sensitive; also, we deliberately did not inquire about participants' health issues. All participants gave their informed consent and were compensated with a 25-euro gift card.

4.2 Procedure

On the basis of previous qualitative studies of users' perceptions of misinformation [15, 16, 45], we designed and conducted semi-structured interviews focusing on participants' encounters with health misinformation (see Figure 1).

Our study protocol (see the Supplementary material) included questions about participants' online experiences with health information and misinformation and their strategies for distinguishing between the two types of information. We delved into their efforts to counter such misinformation in various social contexts, such as in family group chats or comments on video-sharing platforms.

A second set of questions focused on the steps participants hoped or expected social media platforms would take to combat health

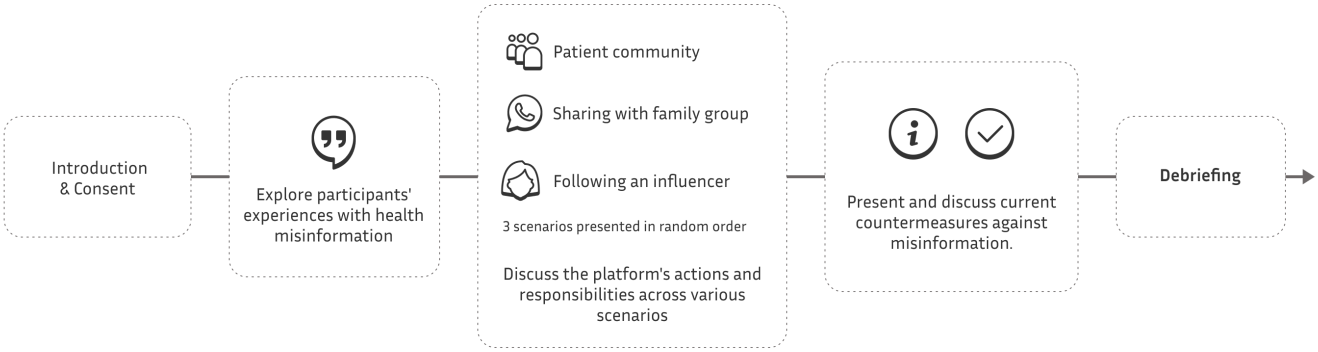


Figure 1: Overview the interview procedure

misinformation. To ensure that our interviews were issue-focused and took into account instances where participants had minimal exposure to the measures implemented by social media, we presented three fictional scenarios (see the Supplementary material). Participants were asked to reflect on the actions of the characters in the scenarios and the fictional social media platform.

Finally, drawing inspiration from the methodology used by Kirchner and Reuter [69], we presented participants with a list of methods currently used by various social media platforms. Contrary to Kirchner and Reuter's approach, we opted to simultaneously show participants a large selection of these solutions. We chose this adapted approach to encourage participants to compare the nuances between various (potentially similar) ideas. Participants were then prompted to identify differences and similarities between the solutions, rank them in terms of perceived efficacy and appropriateness, and pinpoint situations where these solutions might be practically applied to counter health misinformation.

The interviews took from 1 to 2 hours, with a mean time of 1 hour 13 min (SD = 14.8 min). Both the first and the last authors jointly conducted all interviews.

4.3 Scenario Development

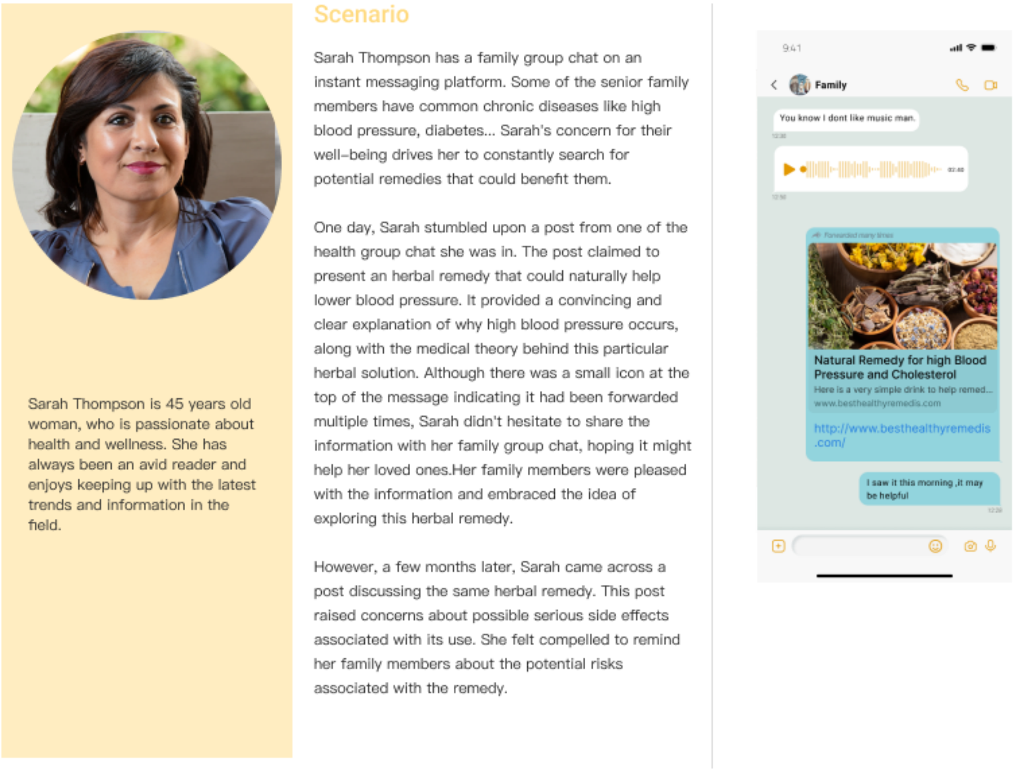
Scenarios describe realistic interactions between humans and interactive systems [133], and personas are realistic (but fictitious) descriptions of people that allow us to immerse ourselves into the minds of others [54]. The two methods are often combined and are well-established in HCI. We used a scenario-based approach that allowed participants to immerse themselves in different situations without referring to their own personal health history. We created three fictional scenarios in which a persona interacted with social media to probe user reactions to potential ways to deal with misinformation on social media. The scenarios described emotionally challenging situations for the persona who faced potential controversial corrective actions from the social media platform. We chose this approach for two reasons: First, the fictional setting ensured that participants would not be pressured into sharing sensitive information. Second, it encouraged them to engage with the issue and draw parallels with their own experiences. This method is consistently used in studies in which participants are asked to reflect on potentially sensitive content [9, 24]. The three scenarios were presented in a random order.

To craft stories and personas, we first researched typical settings involving health misinformation on social media by investigating previous research and conducting searches on social media. We discussed our findings and decided to focus on three areas: patient communities, private sharing within a family group, and following an influencer. Second, we developed a prompt that outlined the primary narrative and crucial story elements. Using the ChatGPT language model, [2] we generated a series of narratives (N = 12). The first and last authors reviewed and refined the resulting narratives to better align them with our research objectives and to protect against biases potentially introduced by ChatGPT. Further corrections were made after two pretests. In the following, we describe the scenarios (see Supplementary material) in detail.

4.3.1 Scenario: Patient Community. Online health communities or patient communities have been a phenomenon since the inception of social media. The primary goal of such communities is to connect people, allowing them to interact with others, seek advice, and provide information and support for those in similar situations. Medical professionals can also use these communities to guide patient groups [18]. Several studies have demonstrated the benefits of such online communities (e.g., [10, 46]). However, studies have also highlighted a considerable level of health misinformation [101, 111]. For instance, a study by Petukhova et al. found that up to 87% of posts offering medical advice in a Facebook skin cancer community contained claims that were not supported by evidence-based medicine [101]. In our fictional scenario, we introduced a character with a chronic illness who is part of a patient community. This person tried a new alternative treatment, which had not yet received clinical validation. After seeing improvement in his condition, he shared his experience on a (fictional) social network. However, the network removed his post, labeling it as "misinformation." Our story highlights the tension between "sharing personal experiences" and potentially "spreading misinformation." We also prepared additional questions to explore instances in which

---

[2] ChatGPT is a conversational AI model developed by OpenAI based on the GPT-4 architecture. https://chat.openai.com



Sarah Thompson is 45 years old woman, who is passionate about health and wellness. She has always been an avid reader and enjoys keeping up with the latest trends and information in the field.

Scenario

Sarah Thompson has a family group chat on an instant messaging platform. Some of the senior family members have common chronic diseases like high blood pressure, diabetes... Sarah's concern for their well-being drives her to constantly search for potential remedies that could benefit them.

One day, Sarah stumbled upon a post from one of the health group chat she was in. The post claimed to present an herbal remedy that could naturally help lower blood pressure. It provided a convincing and clear explanation of why high blood pressure occurs, along with the medical theory behind this particular herbal solution. Although there was a small icon at the top of the message indicating it had been forwarded multiple times, Sarah didn't hesitate to share the information with her family group chat, hoping it might help her loved ones.Her family members were pleased with the information and embraced the idea of exploring this herbal remedy.

However, a few months later, Sarah came across a post discussing the same herbal remedy. This post raised concerns about possible serious side effects associated with its use. She felt compelled to remind her family members about the potential risks associated with the remedy.

You know I dont like music man.

Natural Remedy for high Blood Pressure and Cholesterol
Here is a very simple drink to help remed...
www.besthealthyremedis.com

http://www.besthealthyremedis.com/

I saw it this morning ,it may be helpful

Figure 2: Presentation of a Fictional Scenario to Participants

personal stories with unverified information might pose risks to a broader audience.

4.3.2 Scenario: Sharing Information in a Family Group. Numerous studies have shown that social messaging platforms, such as WhatsApp and Telegram, are significant contributors to the spread of health misinformation (e.g., [6, 79, 86]). For example, a study by Khaja et al. revealed that 59.1% of health claims on WhatsApp were misleading, with 27.3% being entirely false [5]. One distinctive feature involves private groups, often including family members. This can lead to situations where well-intentioned relatives share health misinformation [41]. It can be difficult to challenge such information, especially when it means contradicting respected and older family members [41, 82, 108]. Thus, recent research has indicated that the effectiveness of current attempts by messaging apps to combat misinformation is limited [56]. The scenario centered on a woman who shared a post from a health group with her personal group. Later, she inadvertently realized that the post contained misinformation, even though it had been forwarded multiple times. This narrative examines the shared responsibility of fact-checking between the original poster and social media platforms. We also asked participants whether messaging apps should bear some of the responsibility and discussed potential interventions.

4.3.3 Scenario: Following an Influencer. Social media platforms focusing on photo and video sharing are prominent disseminators of health misinformation [76, 104]. One notable feature of misinformation on these platforms is the influence of popular users, often called "influencers." These influencers can portray themselves as authentic and relatable, fostering relationships built on trust and intimacy with their audience [121]. This authenticity makes their content appear more credible. However, it also complicates the process of debunking false statements, especially when they are framed as personal experiences. Health-related topics, especially well-being and fitness, are highly prevalent on these media platforms [42]. Tiller et al. noted that the health and wellness industry often pushes exaggerated or unfounded claims, sometimes bordering on pseudoscience [120]. These false claims are frequently cloaked in the guise of scientific legitimacy, with dubious evidence supporting their safety and effectiveness. For example, Topham et al. delved into how misinformation finds its way into health and diet videos [121] and showed pseudoscientific claims as one of the presentation strategies. Marocolo et al. concluded that several influential Instagram influencers from Brazil spread mostly low-quality information about exercise and health [84]. Our scenario centered on a young woman who avidly followed a health and well-being influencer. The advice given by the influencer greatly benefited her, leading her to become a dedicated follower. However, controversy arose when one of the influencer's posts was labeled as "spreading misinformation." Through this narrative, we sought to explore users' perceptions of the balance between "useful information" and "not scientifically supported information." We also looked into the network's role in curbing the spread of unverified health content in influencers' promotional and regular videos.

4.4 Analysis

We used a combined inductive-deductive approach to analyze transcripts of the qualitative interviews [71]. Our main goal was to identify the factors that influence why people perceive health-related claims as "misinformation" and understand the measures people believe should be applied to tackle such misinformation. We also analyzed how social media platforms can counteract health misinformation. The process began with inductive coding, in which we identified themes. The first and last authors initiated this process by openly coding a couple of interviews and discussing and refining the codes. As the analysis progressed, we updated our initial coding system when new codes emerged. Given the potential of reactance theory in explaining users' attitudes toward the proposed measures, we applied a deductive approach to connect our initial codes with reactance theory. Finally, the first author coded all the interviews. The last two authors then reviewed this coding, discussing any differences in interpretation until they reached a consensus.

5 RESULTS

5.1 RQ1: What are users' perceptions of health misinformation on social media, and how do they describe its characteristics?

In the following, we first describe users' perceptions of health misinformation based on the interview data. Afterward, we report topics and sources of health misinformation and finally describe the characteristics of health misinformation identified by users.

5.1.1 Users' Perceptions of Health Misinformation. Most of our participants did not perceive health misinformation as entirely false or fabricated. Instead, they defined it as content that lacked scientific support or clear empirical evidence. However, many participants emphasized that this absence of evidence does not necessarily render the health practices ineffective, nor does it mean the information's status cannot change based on new findings. Participants considered it important to approach such information cautiously, given its potential for long-term negative consequences. However, identifying health misinformation can be tricky due to grey areas and ambiguous content. For example, when discussing topics such as "natural health remedies" or "traditional medicine," some participants argued that health-related information should always be scientifically tested and verified, no matter how effective it appears to be for individuals. By contrast, others believed that if the information was based on personal experience, it cannot be called "misinformation," even without official medical verification.

> **P22**: It's a grey area, I'd say. If the outcomes are positive [...], even if the treatment is unverified, and there are no severe consequences, then I'd argue it would be wrong to just call it misinformation.

Many participants mentioned that health misinformation often appears reasonable, logical, and easy to consume, sometimes even accompanied by images, which enhance its comprehensibility. Additionally, the succinct format of social media content was identified as a factor that contributes to the misinterpretation of authentic health information. This problem arises because these short posts typically lack comprehensive details.

Finally, our data provided evidence that religious beliefs might impact users' views of health misinformation because, in some cases, religious figures or texts can be providers of misinformation.

> **P4**: I'm from a country where people always consider religious matters, and for this kind of health-related issue, there are always statements from figures like prophets. People always try to consider what they say. Thus, because this health misinformation has religious origins, people also consider it.

Addressing such misconceptions can be challenging as it can be viewed as an attempt to criticize the faith.

5.1.2 Topics and Sources of Health Misinformation. Prevalent health-related topics on social media mentioned by participants included fitness, food, a healthy lifestyle, mental and general health, and treatments and vaccinations. Three main channels were mentioned: advertising and influencers, dissemination of information within family groups, and self-initiated searches.

Advertising and influencers: Sixteen participants mentioned influencers as a primary source of health information on social media. They stated that they view shared health advice and knowledge as beneficial preventive measures, especially when they are afraid having related health issues and when the information is endorsed by individuals who appear to be doctors. Nonetheless, concerns about the authenticity of these self-proclaimed doctors were noted, as some may be impostors who are just trying to sell products. Our participants mentioned that the only indicator of such content is the "promo" label introduced by social media platforms. For our participants, this label makes such content look suspicious as though it may be a potential source of misinformation. Additionally, repeated exposure to this content will increase the familiarity bias:

> **P1**: This advertisement has claims similar to true information, and it seems reasonable and logical. You then think, "Oh yeah, I've seen this before, so it must be true, and it said the same thing."

Although some influencers have a medical background, which makes them a good source of health information, some *"influencer doctors [. . .] are famous or have a large number of followers" (P7)*, which might make their followers accept their posts as factual.

Dissemination of information within family groups: Fourteen participants mentioned that their family members shared health-related information, sometimes accurately, but more often not, driven by the motivation to protect their loved ones:

> **P7**: She [my mother] wants me to be as healthy as possible. So she shares that [type of advice].

Although older family members sometimes ask younger relatives for assistance in understanding whether information is accurate or false, our participants mostly reported problematic situations in which they had to decide whether to address misinformation shared by older and respected family members, which is in line with findings by [41, 82, 108]. In contrast to [82], our participants were more inclined to embrace direct strategies against misinformation, most likely because they anticipated more dire consequences for their relatives' health (compared with the "polite fighting" in political cases discussed in the study by Malhotra et al. [82]). For instance, P1 blocked his sister's access to a particular website and IP. P4 stated, *"I do this because I want to prevent them from sharing information they should not trust. I sometimes try not to be polite, to show them [family members] something."*

Self-initiated searches driven by personal motivation: Most of our participants stated that they look up general health information or, more specifically, symptom information before visiting a doctor. However, they use Google and medical websites rather than social media. They believe that information on social media is highly subjective and leans more toward "personal experiences" rather than objective facts.

5.1.3 *Identifying Characteristics of Health Misinformation.*

Participants identified four heuristics to quickly assess whether information is likely to be false: content that conflicts with established common knowledge, a lack of credible sources, content that goes against medical expertise, and advertising-in-disguise content. These results are consistent with past studies highlighting the importance of checking the origins of posts to determine their trustworthiness and using common sense to assess the accuracy of information [43, 47]. Eight participants explicitly stated that they are not likely to trust content that is excessively unfamiliar, rare, or offers extreme advice. Additionally, seven participants summarized their skepticism of health misinformation as *"If it is too good to be true."* Furthermore, participants also noted that posts containing misinformation often lack comprehensive details and emphasize only positive aspects. Two participants noticed that the posts did not use proper scientific words and that the claims did not match the reasons that were given. These discrepancies made them look at the posts more carefully.

Five participants were cautious about ads and influencers. They were worried about whether the influencers had researched or even used the advertised products, whether the financial motivation was affecting the truthfulness of the information, and whether the ties between advertisers, influencers, and products had been made clear. In their view, factors that might make people vulnerable to misinformation include age (primarily children and older adults, for which there is some supportive and contradictory evidence in the literature [93, 94]) and level of education (also supported by research [64, 94]). However, participants highlighted that being in need of a cure (and potentially being desperate) reduces critical thinking. This issue was perceived as particularly prevalent for health misinformation as it might in some cases offer a vision of a better life and could be seen as a tool to cope with life crises [136].

> **P15** Sometimes[...] [people who trust health misinformation] realize that this solution is misinformation; it is not really reliable. But they're so desperate that they're just like "OK, let's try it."

5.1.4 *Strategies for Debunking Health Misinformation and the Challenges That Come with Them.*

Many of our participants were willing to speak out against health misinformation in family groups but were hesitant to do so in public spaces (just four participants reported having done this). Participants provided the following reasons:

- Information overload: Participants reported being overwhelmed by the amount of health misinformation on social media and felt that it is not worth addressing each case.
- Lack of confidence in the domain: Many participants doubted their health expertise and thus worried that their comments could influence others in unintended ways.
- Tough and time-consuming online debates: Participants found persuading people online to be difficult. They reported that online discussions often become heated, emotional debates, with individuals sticking to their views.
- Algorithmic impact: Some participants were aware that any interaction with content, even to debunk it, could raise the likelihood of receiving similar content through algorithms, which could amplify the misinformation.

These challenges did not stop our participants from taking action in the family groups. However, they questioned the effectiveness of debunking information within their family groups because the group members could always make a counterargument based on more health misinformation from outside media. Nevertheless, they mentioned the following debunking strategies: providing a reliable source, showing the advertising nature of the reposted information, and using interpersonal relationships and knowledge about certain people to find the most persuasive arguments. These findings are similar to previous results in the area of combatting general [108] and COVID-19-specific [126] misinformation, but in contrast to previous studies [108], our participants did not perceive the family members' personality traits as a barrier, but rather as an opportunity to create tailored interventions.

5.2 RQ2.1: How do users view the responsibilities and actions of social networks regarding health misinformation?

In the following, we first describe where our participants thought the responsibility for preventing the spread of misinformation on social media was located (user- or platform-initiated countermeasures). Afterward, we present several discussed individual solutions.

5.2.1 *Responsibility for Preventing the Spread of Misinformation on Social Media.*

Participants expressed diverse views about the responsibility to curb the spread of health misinformation. Most of our participants argued that the responsibility affects both platforms and users (and sometimes also additional institutional actors, such as governments, health organizations, news sources, and doctors), but they acknowledged that the right balance between the user autonomy and platform accountability is challenging.

When discussing lay users' responsibilities for fact-checking in the health domain, most participants mentioned that they were skeptical about information. They relied on scientific sources and comparing multiple sources to distinguish between "fake" and "true" information. However, participants acknowledged the limitations of this strategy: It takes time, and genuine information can be harder to understand than easily digestible misinformation.

> **P1**: No one is going to read the research paper published in, I don't know, how many pages, and in a certain language form that is not understandable. So, of

> course, the easiest, simple-language-expressed [misleading] advertisement will have much better exposure than research done by some scientists published in Elsevier or somewhere in some journal.

In contrast to assigning responsibility to the platform, a few participants emphasized that a social media platform's purpose is to serve as a venue for interactions, not necessarily to function as a curator of truth:

> **P12**: The main objective of social media or social networks is to connect people from all over the world. This means offering a venue for sharing information. But now, what do you think? While I can provide information, I'm not responsible for what you believe or what you do with the information. So, what should social media do?

Some participants argued that platforms might even benefit from misinformation, as it might spark intense social debates and increase platform engagement:

> **P16**: People voluntarily spread misinformation just to get others to contribute their thoughts in the comments.[...][Social media platforms] know that some content isn't really good, but it spreads, gets likes, and keeps people on the app.[...] I would like them to promote critical thinking and make an effort to ensure that everything is as clear and helpful as possible. However, I also understand that from their business perspective, it's not [working].

Another viewpoint was that the financial relationship between social media platforms and commercial actors might provide benefits to both parties by allowing questionable health content, such as biased or unscientific claims by influencers. In such scenarios, it might be unrealistic to expect any corrective actions from the platforms.

Still, most of our participants believed platforms should intervene to halt the spread of misinformation, thus echoing findings by Saltz et al. [106]. But in addition to these general findings, we discovered that participants believe there should be distinct differences between how political news is handled compared with health news. Participants felt that health misinformation influences not just beliefs but also everyday life decisions, making its impact more immediate and pronounced than the impact of its political counterpart.

> **P9**: I think when discussing political topics, we might be able to afford it [to overlook misinformation]. However, when it comes to health, we can't afford it, because, in health, it is a matter of life or death. If someone is making health decisions or adopting new practices, then it's crucial for them to have accurate information.

This distinction frames health misinformation encounters not merely as misleading but as potential life-threatening events, suggesting that the value of free expression might need to be weighed against safety concerns. At the same time, five participants also mentioned that the same strict regulative measures can be applied to political content, discussing the cases of misinformation that have dire consequences for society.

5.2.2 Attitudes Toward Platforms' Deletion or Hiding of Health Misinformation.

While most of the users mentioned deletion of the content as undesirable for combatting misinformation, many of our participants were positive about blurring out misinformation or algorithmically lowering it in the news feed. This discrepancy might indicate that deleting information does not discredit it: People who have already viewed the information might not become aware of its deletion and might therefore remain unaffected. Thus, deletion becomes the opposite of debunking, as it limits future opportunities to show that the information is wrong. For example, P12 explained that in their country, people often download content for later use because they do not always have access to the internet and cannot always check the content's status when they reconnect again.

> **P12**: Now, this social media has a section that tells us if something is true or not. They can check and recheck if they have a video on their phone or laptop. [People] can verify when they have internet and say, "OK, I want to know if it's correct," and they can check. If information gets deleted without stating that it's been deleted, you can't recheck and find out that it's not true.

This positive attitude toward (reversible) algorithmic downplaying and blurring the content was also perceived as adequate for the "grey" nature of health misinformation, where the status of information can change in light of new scientific (or, for some participants) personal evidence. Finally, participants mentioned the implicit value of health misinformation as an opportunity to gain the expertise needed to assess it (e.g., P19: *"I don't think that deleting everything is a good idea because then you lose your critical reading").*

However, some participants questioned whether any health-related content should be allowed on social media, whether true or misleading. They felt that such information could confuse the general public. Thus, they believed medical topics should be reserved for dedicated alternative media platforms. Similarly, two participants felt that only verified specialists should post health-related content as influencers.

5.2.3 Attitudes Toward Implementing Extra Measures for Vulnerable Populations.

Many of our participants did not believe that protective measures should be taken specifically for vulnerable groups such as the young, elderly, or less literate. Some believed that increasing protection for everyone would be better than extra protection for children. Others suggested that media literacy courses should be provided for both children and vulnerable adults. Twelve participants expressed that, rather than limiting the content, social media platforms should implement general age restrictions, such as setting a minimum age for use and showing the specific content to the users who meet these age requirements. Noteworthy, these findings relate to which groups our users perceived to be vulnerable rather than empirical evidence of whether these groups are actually vulnerable.

5.2.4 Attitudes Toward Using Manipulations to Nudge People Away from Health Misinformation. A common strategy for combating misinformation online involves subtly nudging users toward choices that social media platforms deem optimal, such as discouraging the sharing of questionable information. These practices are aligned with the idea of "good" nudges [119], but they can still be viewed as manipulative because they take advantage of users' inherent "default" biases and employ "interface interference" techniques [53]. Many of our users recognized that these interface designs can be manipulative but also found them acceptable for preventing health misinformation. Some participants felt that these tactics were justifiable for health content but not for political information. Additionally, one participant viewed these design choices as entirely manipulative and unacceptable in any context. In contrast to previous literature, we found minimal evidence that current interface solutions provoke reactance or a backfiring effect when combating health misinformation. The only approach that was associated with strong, negative user emotions was permanently deleting the content. However, we should approach this finding with caution. While our participants reported witnessing some cases of content curation in networks (mostly deletion and labeling), they were not personally subjected to such curation. Therefore, we anticipate that individuals whose content was downgraded or blurred might exhibit a bit more intense reaction to the measures.

5.3 RQ2.2: Drawbacks of Current Solutions for Handling Health Misinformation

When discussing the current solutions, most participants expressed positive feelings about warning labels and pop-ups. While it is unclear whether this finding generalizes to other populations, our participants appreciated measures that ensure credibility, such as "*the video is provided by WHO*." However, they pointed out several issues with both strategies that need to be addressed for more effective interventions.

5.3.1 Actionability of Solutions. Current credibility labels fall short of effectively bridging the gap between detecting misinformation and providing actionable coping strategies. In the context of health information searches, individuals often seek information for practical purposes. Three of our participants noted that without fulfilling these needs, users would continue seeking information from equally questionable sources that may not yet be marked as unreliable. The challenge extends beyond merely flagging misinformation; it involves providing accessible, trustworthy alternatives. As a positive example, some participants mentioned advice such as "*Talk to your doctor about that (symptoms or remedies)*," but this approach cannot be generalized across all the possible issues of health misinformation. Moreover, it might not apply to areas where it is difficult or financially challenging to gain access to healthcare professionals, such as the Global South [123].

5.3.2 The Verification Process is Lengthy and Lacks Transparency. Additional challenges included a lack of transparency in understanding why something was labeled true or false and how to explore the provisional sources of information. For example, most of the participants felt that the "See Why" link provided under the social media platforms' evaluation should lead straight to reasons that pertained directly to the specific piece of information, but it often leads to general rules about judgments made about the content on the platforms. Similarly, if a piece of information was "checked by independent fact-checkers," participants expected the evidence from these fact-checks to appear on the very same page. Several participants mentioned a multistep approach as potentially diminishing the effect of the credibility indicator.

5.3.3 The Lack of Connection Between Scientific Knowledge and Its Representation in Social Media Warnings and Credibility Indicators. While all of our participants mentioned science as the main source of credible health information and for personal fact-checking, they found it difficult to navigate scientific information as nonprofessionals in health. A recurring suggestion to improve this challenge was to either provide direct links to the (limited) number of credible health resources or to present the scientific findings along with credibility indicators as a summary of explanations in common, understandable language.

5.3.4 The Lack of a Global System for Source Verification. Specifically, in the context of health misinformation, several participants acknowledged the absence of mechanisms for verifying the credibility of sources (in contrast to the credibility of information) and proposed various design ideas on how to verify a person's credibility in providing health-related information. In this context, credibility was framed in terms of both expertise credibility (meaning that individuals who post health-related information should provide documents confirming their professional competence) and experience credibility (meaning that a patient who describes their symptoms should provide proof of their specific health condition). Drawing from these suggestions, we can infer that users perceive that current solutions address only the information itself. However, they would also appreciate more solutions that are targeted toward interacting with the source of the information.

6 DISCUSSION

6.1 Changes in Information Status Over Time and the "Grey" Nature of Health Misinformation

Our findings are in line with previous research in the domain of health misinformation in that they point to the usually "grey" nature of such misinformation [117, 144]. In this context, the current binary labeling system can oversimplify the issue and produce measures that can be perceived as too harsh or too weak. Additionally, our participants mentioned that the current practice of labeling information as true or false just once contradicts users' perceptions of knowledge as something that evolves and thus impacts their understanding of truth and falsehood over time. The issue appears to stem from users sometimes being unable or unwilling to revisit information that was previously marked (or not); thus, they are unable to update their perceptions of facts on the basis of new evidence. This issue is aligned with Horne et al.'s position, as discussed in [61], where they argued that the changing nature of evidence should be reflected in the behavior of automated fake-news-detection models. Additionally, fact-checking is a manual, labor-intensive task involving short-term engagement with each claim to determine its accuracy [67]. Given the limited time frame, fact-checkers, once they have assigned a veracity label to a claim, might not have the opportunity to reassess and update their judgments in light of new evidence. However, even with timely updates, such as the addition or removal of "fake news" warning labels on content, users might still unintentionally overlook these changes. This oversight can become even more prominent given the design of social media news feeds, which do not incentivize users to revisit information.

6.2 The Relationship Between Scientific Knowledge and Common Sense in Debunking Misinformation

Our results revealed the coexistence of two primary strategies for debunking misinformation. The first strategy targeted users' heuristics or "gut feelings" about the content (see Section 5.1.3), with discussions about the "too good to be true" or "common sense" approach to evaluating information. Conversely, a more deliberate and thorough process of self-fact-checking was detailed, grounded in a rigorous search for evidence and an exploration of scientific knowledge. This latter approach was also associated with higher personal cognitive costs, as participants recognized that it necessitated medical scientific knowledge (as discussed in 5.3.3) and the comparison of sources (see Section 5.3.4). These two approaches can be interpreted through the lens of Kahneman's System 1 and System 2 theory, wherein System 1 involves fast, intuitive thinking, and System 2 involves slower, more analytical reasoning [34]. Previous studies have demonstrated that social media interventions can effectively employ a comparative approach, engaging both System 1 (intuitive and fast thinking) and System 2 (analytical and slow thinking) processes [91]. Furthermore, interventions that integrate the two approaches tend to have a more pronounced effect [91]. However, in the context of health-related information, designing interventions that appeal to System 2 can be challenging, as they often involve references to specialized domain knowledge. In this context, individuals may more frequently rely on personal heuristics when the complexity of the domain renders elaborate efforts too challenging for most people to deal with everyday encounters with misinformation. At the same time, relying on System 1 thinking is more often associated with susceptibility to misinformation [124]. Therefore, it is important to find the right way to get users to look for more information without unrealistically raising the cognitive costs.

6.3 Interpreting Promotional Labeling as a Proxy for Misinformation

In addition to warning labels, social networks have also implemented promo labels as another labeling system that can provide users with additional information about the content. Interestingly, our participants associated these labels with misinformation. Furthermore, attempts to make the promotional labeling less visually salient were perceived very negatively (as "advertising-in-disguise"). This finding emerged during a discussion of influencers' behavior on social networks. Therefore, our findings are connected to results from [17] showing the audience's negative response to receiving "promoted"-labeled posts from influencers. Our users were also generally suspicious about celebrities who promoted health-related information as one of the main sources of misinformation, discussing both native and open advertising schemes. Previous studies have shown that celebrity promotions of health-related information significantly contribute to the dissemination of misinformation in this domain [139, 142]. However, they have also revealed that the effect of influencer-propagated misinformation can depend on how much the audience trusts the influencer and perceives them as experts in the field [68]. The generally negative attitude we found was strongly connected to the issue of the credibility of the influencers' background. Unlike other types of misinformation, expertise in the health and medical domain is highly institutionalized beyond social networks (in medical education and certification). Therefore, the participants' negative view of promo labels in health content can be interpreted as a call to bring institutional standards into social media practices.

6.4 Institutional Trust and Favorable Views on Content Curation

Our findings showed differences in views about the role that platforms should play to combat misinformation. A few participants viewed platforms as venues for interaction that should focus on enabling users to share information but are not responsible for the truthfulness of this information. Some participants even mentioned that social networks might profit from health misinformation and, therefore, have little incentive to prevent its spread. For most participants, however, platforms and users are both responsible for combatting misinformation. For them, the platforms have the responsibility to limit the impact of health misinformation, especially because health misinformation might have more severe consequences than political misinformation.

Our data also indicate that, at least for health-related content, participants have a rather favorable view of the different kinds of nudges, acknowledging that "manipulations for good" can be acceptable in the health-related context. In the same vein, we did not find solid evidence that users perceived the current warning indicators as useless or as triggering much reactance per se. Our findings are aligned with evidence that a backfiring effect is unlikely to occur in some domains of health misinformation [39]. Moreover, despite concerns that the constantly changing messages from public officials about social distancing and face mask use during the pandemic may have eroded trust in health authorities, our findings support studies that found no negative impact on trust in international health agencies due to COVID-19 management [80]. They were still cited as the primary (and nearly the only) reliable source of health information. When combined, these findings show that the perception of health misinformation as a high-stakes threat makes users delegate the responsibility of filtering and correcting misinformation to trusted organizations, including social network agents. This delegation is facilitated by the perception that these entities are working for the public good, aiming to protect public health and safety. The acceptance of manipulations—in this case, nudging people toward healthier behavior or corrective information—stems from a trust in their good intentions and capabilities. In areas where the stakes are perceived as high and the potential for harm is significant, people are more inclined to accept and even welcome interventions that might otherwise be viewed as paternalistic or manipulative. However, this does not mean that all forms of manipulation, even for beneficial purposes, should be universally accepted. The ethical considerations surrounding these practices require careful scrutiny. Transparency about the nature and intentions behind such manipulations, ensuring they are evidence-based, and allowing individuals the freedom to completely opt out of having manipulative elements in the interface should be the main aspects to consider. Interestingly, our participants were more inclined to delegate the job of judging the veracity of health information to trusted organizations than to their own social networks. This finding might further underscore the idea that users perceive health information as more serious than other kinds of information, making them trust health professionals more than their private contacts.

6.5 Strategies for Debunking Misinformation: Public vs. Private Approaches

Our study reveals that the debunking of misinformation is much more likely to occur in small, family-oriented communications in messaging apps than in open communication (e.g., in the comments under the content), where unfamiliar users can interact with the content. There are two possible explanations for the findings; one of them is that commenting on external platforms can provoke trolling and harsh criticism from the person who posted the misinformation as well as from other commenters sharing their views, creating a situation that can be emotionally draining and generally unpleasant. On the other hand, message-based private communication raises the responsibility of individuals to take debunking into their own hands. As mentioned both by our users and in other studies, messaging apps do not provide any indications that the information could have misleading content by default; the only indicator is the number of shares [2], which can be interpreted in multiple ways [57]. In these circumstances, users cannot rely on the platform's moderators or on some other unfamiliar user to debunk the misinformation. To some extent, this environment fosters a sense of duty and empowerment among users, as they feel personally responsible for the accuracy of the debunking information they share within their private circles. Furthermore, the intimate nature of message-based communication encourages a more thoughtful and careful approach to debunking, as individuals are more likely to consider the impact of their words on their relationships.

However, this emphasis on individual responsibility in private messaging contexts also presents challenges. The debunking of misinformation within closed groups is particularly complex due to cultural nuances. In certain cultures, correcting others, especially elders, may be perceived as disrespectful, and may have the potential to incite conflict rooted in family hierarchies [100, 108]. Therefore, navigating the complexities of relationships and users' motivation to counter health misinformation in their intimate networks requires a collaborative approach from platforms. Moreover, relying on personal networks for fact-checking and debunking can lead to echo chambers, where misinformation is either not challenged or reinforced by similar opinions within the group [30]. This tendency is particularly worrisome given the diverse quality of information sources individuals may access and share. Additionally, the lack of public accountability mechanisms in private messaging systems means that there is no systematic way to correct misinformation that spreads across these networks, unlike on open platforms where public fact-checking and moderator systems can potentially reach a wider audience [81]. While the personal touch of message-based communication can enhance the effectiveness of debunking efforts among close contacts, it may also limit the extent to which such efforts can reach wider communities.

At the same time, if open social media platforms want to facilitate user involvement in this process, individuals might be more inclined to act. For instance, Twitter's introduction of the crowdsource fact-checking platform, Community Notes, has been well-received [1]. This feature allows certain users to contribute additional context, fostering informed decision-making. Although this approach might not guarantee accuracy or rival the trustworthiness of professional fact-checkers, it has proven useful in curbing the spread of low-risk misinformation [134]. Furthermore, users often read comments for more insight, corroborating findings by Guo et al. [55] and Geeng et al. [47]. Hence, platforms could leverage the comments section to address misinformation.

7 DESIGN RECOMMENDATIONS

In the following section, we outline design recommendations that have the potential to balance strategies from the side of users with those from the side of platforms, help to raise users' agency and promote active engagement in the debunking of misinformation. These recommendations were identified in several ideation rounds by the authors based on users' suggestions, acting as foundation for future work. A brief overview of key discussion topics and corresponding design solutions is presented in Table 1. As the work was performed at the European university, we oriented our efforts towards the current European regulatory frameworks related to the liability of social media and the spread of misinformation (namely the "Digital Services Act"[3] and the "EU Code of Practice against Disinformation" [4]) when crafting the recommendations. For example, our recommendations address issues such as transparency, the expert status given to trusted users, and advertising content, which are highlighted in the DSA and relevant in the context of misinformation [114]. However, our recommendations can be adapted to other jurisdictions.

7.1 Tracking Changes in Information Status Over Time Using the Health Information Veracity Scout

In Section 6.1, we discussed users' need for a way to track changes in information status over time without overburdening them. A potential solution might be a system that we call "health information veracity scout," based on a sports scout. In sports such as soccer, a team identifies a position that is in need of reinforcements (e.g., a defender), and then they send their scouts to identify potential candidates for this position. These scouts observe potential candidates over a series of games and report their observations back to

---

[3] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-acten

[4] https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation

Table 1: Overview of Key Discussion Topics and Corresponding Design Solutions

Topic for Discussion	Key Discoveries	Design Recommendations
Changes in Information Status Over Time and the "Grey" Nature of health-related information [6.1]	In light of new findings, the status of the information can be changed from true to false and vice versa [5.1]	Tracking Changes in Information Status Over Time Using the Health Information Veracity Scout [7.1]
	The information can be judged not from a true/false position but to reflect the social component of the network [5.1]	Tailoring Solutions to Diverse Scenarios [7.2]
The Relationship Between Scientific Knowledge and Common Sense in Debunking Misinformation [6.2]	Users apply their set of heuristics to determine the veracity of information; these heuristics help them make fast decision about information [5.1]	Employing User-Provided Heuristics to Detect Misinformation [7.3]
	Users also apply a reflective strategy to the posts, but it usually takes more time and cognitive effort [5.1],[5.2]	Improving Health Information Understanding [7.4]
Users Interpret Promotional Labeling as a Proxy for Misinformation [6.3]	Highly negative attitudes toward the promo labels in health-related content, appeal to have better certification system for both paid and native promoted content in health domain [5.1],[5.2]	Enhancing Access and Trust in Professional Health Care Providers [7.5]
Institutional Trust and Favorable Views on Content Curation [6.4]	Users still acknowledge health-care providers as the ultimate source of health-related information [5.2], [5.3]	
Strategies for Debunking Misinformation: Public vs. Private Approaches [6.5]	Users apply different strategies for debunking privately and publicly [5.1], [5.3]	Increase User Agency [7.6]

the team's managers. Similarly, we imagine a system that saves interesting health information content, tracks changes in its veracity, and reports updates back to the user. A user might be able to flag a particular post as interesting, or the system might automatically flag posts when a user interacts with them (e.g., by sharing). If the post is connected to a system of detailed criteria (e.g., whether it includes scientific evidence or personal stories, see Section 7.6), the system could track changes in these criteria, similar to a changelog or a Git history. It could then inform the user about changes in these criteria, potentially encouraging critical thinking ("Do you want to verify whether these changes impact your opinion?") or debunking ("Do you want to verify a change in the status of the veracity of the post and reshare it?"). We also imagine that users would be able to set preferences such as specific thresholds of changes ("only inform me when the veracity of the information changed at least…") or extend the scouting procedure beyond single posts, for example, based on AI-based extraction of their underlying topic. In order to fight the unintended promotion of filter bubbles, the "health information veracity scout" could also inform users about related but contradictory posts, for example, when these posts meet specified criteria or surpass specific thresholds. Regarding the feasibility of such a "health information veracity scout", we imagine that a post could be referenced by a unique identifier (e.g., a link) and stored in a database, together with the aforementioned criteria. Future work is required to investigate users' reactions to such a "health information veracity scout", in particular regarding its use in a realistic setting.

7.2 Tailoring Solutions to Diverse Scenarios

It is crucial to recognize that not all instances of health misinformation carry the same level of gravity and harm. Consequently, there is a pressing need for collaborative efforts among platforms, public health agencies, scholars in communication, and government authorities to establish clear, transparent criteria (e.g., severity of potential health consequences, sources of genuine information about the topic according to the scientific consensus, expected scale of dissemination, target audience vulnerability, etc.). These criteria should be made readily accessible to the public, serving as a guideline for distinguishing between the boundaries of free speech and safeguarding individuals from health misinformation.

To illustrate, such criteria can delineate the circumstances in which platforms should exercise greater control over information-sharing by diminishing the visibility of misinformation or restricting access to posts. Simultaneously, they can specify situations that warrant the inclusion of disclaimers or merely necessitate the deployment of subtle warning icons. This multifaceted approach ensures that the response to health misinformation is aligned with the specific context, ultimately enhancing the effectiveness of interventions.

7.3 Employing User-Provided Heuristics to Detect Misinformation

In our study, users mentioned a number of heuristics they use to make a decision about information veracity (see Section 5.1.3). Social networks can build on the users' insights and utilize the users' heuristics to assess the questionable content and to communicate the results of the assessment to the users. One possibility might be to build on the above-mentioned characteristics that make users suspicious that a particular piece of content might contain health misinformation. For example, when fact-checkers work on a specific piece of content, they could collect links to related content and tag whether the linked content supports or contradicts the claims made in the currently reviewed content. This information could be made available to users to support their judgment process or as an invitation to learn more about the specific topic. Although building on these heuristics can empower users to verify health information, designers and researchers should not limit themselves to such approaches. Users might not be aware of all the cues that can be used to identify health misinformation, and in addition, they might not be able to apply their heuristics thoroughly if a specific post with health misinformation intentionally mimics characteristics of genuine health information.

7.4 Improving the Understanding of Health Information

Health information often involves complex medical domain knowledge, making it challenging for users to comprehend or consequently debunk misinformation via scientific sources. Thus, we recommend the use of tools that can automatically summarize and simplify text while highlighting key concepts and explaining the main message using nonspecialized language. The results of summarization can be present inside the warning popup or in the extended warning label. Besides summarization, visualization (e.g., using infographics and diagrams) might play an increasingly important role in improving the understanding of health information. We imagine that these services could be provided by specialised journalists on dedicated platforms, financed by users directly or by platforms as part of a premium subscription.

7.5 Enhancing Access and Trust in Professional Health Care Providers

Our results indicate that users are willing to trust individual experts (e.g., fact-checkers) or expert organizations (e.g., WHO), which is in line with the results showing the positive effect of expert debunking on lowering users' reactance towards debunking message [115]; However we found that users need specific and instant factual information about these entities and their sources of expertise. For example, regarding "fact-checkers," users want to know who these fact-checkers are, what their professional background is, and how they decide whether something is true. This information should be easily accessible (possibly right near the credibility claim). Also, incorporating a warning pop-up or label with a link to the trusted source will promote the verification of information, as it makes the necessary information easy to access. We are also in favor of platform-based expert certification in the health domain. If a person claims to be a health expert, social media should encourage them to share their qualifications. A potential solution might be a central cross-platform repository of certifications, for example based on Blockchain or credentials wallet technologies. Furthermore, research is needed on how to comprehensively communicate health expertise to users, for example, with digital badges, endorsements, or gamified elements, as well as into potential differences of users regarding their trust in health experts. The rigorous verification process for health experts could also enhance the quality of promotions and advertisements. Once health influencers receive platform verification, they will likely maintain their positive social profile, thereby choosing their endorsements more carefully. Moreover, transparency is crucial in the promotion of health products. This includes listing potential side effects and adding disclaimers for products that may pose risks.

7.6 Increase User Agency

Most solutions that already exist in social networks passively provide users with final judgments about the integrity of information. At the same time, our participants demonstrated that consumers can also become better protected against misinformation by developing their critical thinking. Previous research has indicated that user involvement is crucial for combating online misinformation [72]. For instance, a system that prompts users to question the characteristics of information can enhance their defense against incorrect details [35]. Possibilities to implement user-agency-increasing functionalities range from small nudges to adapting content algorithms based on user preferences. For example, integrating alerts about content with prompts might allow users to better understand the value of genuine information and promote their engagement. Users should also have the option to ask questions about the specific details of the information in a dialogue format, which can be integrated with or linked to the warning. System design can also encourage critical thinking by inviting users to self-reflect. For example, when sharing information, a system could allow users to specify their level of trust in the information and ask which criteria they are basing their evaluation on. Furthermore, considering the strong role of small private groups in disseminating health information, we hypothesized that users might be interested in delegating the responsibility of fact-checking to trusted members of their social circles. For example, a user might select that they trust a specific friend to judge health information, and this friend could then add a user-generated and verified badge to a post, indicating to their network that they believe it contains genuine information. Therefore, we suggest that future designs incorporate an interactive approach that actively engages users in asking questions about different aspects of health-related information. Also, a fine-grained system of veracity indicators in health information might better align with the differences we observed: Users who prefer empirically validated health information might indicate this in their personal settings. An automatic system could then de-emphasize health posts that are based purely on personal evidence, for example, by adding an additional layer of content blurring.

8 LIMITATIONS AND FUTURE WORK

In our study, we aimed to minimize the Global North bias by recruiting participants from diverse cultures. However, the majority of our participants were university students and staff, which impacts the generalizability of our results. For example, trust in scientific sources might be higher in a university sample than in the general population. Additionally, our sample was not diverse in terms of age, which offers a unique perspective but may not fully represent age-related differences in opinions. In subsequent studies, there should be a more in-depth analysis of the perspectives of individuals without tertiary education and across a broader age range. Such an expansion is crucial because our participants frequently cited education level and age as factors that make certain groups more susceptible to misinformation, particularly health misinformation. Thus, it is necessary to understand misinformation from the viewpoints of these potentially vulnerable groups.

Given the vital role that context plays in HCI, we used a thoroughly validated study design that built on personas and scenarios to collect our data. While this approach allowed our participants to explore diverse design solutions to combat health misinformation in context, this setting is not intended to replace the real personal experience. Likewise, our design recommendations should be validated with users before they are implemented.

Our findings indicate that health misinformation is intricate and has multiple layers. People's reactions vary on the basis of immediate and long-term impacts as well as the nuances in the information they encounter. We used specifically designed misinformation scenarios to understand diverse situations where individuals might come across health misinformation. However, further research is required to develop a comprehensive taxonomy of health misinformation issues prevalent in social media.

In addition, our study focused on users as recipients of health misinformation, not on the perspective of a person who is sharing misinformation. Future studies should complement the users' perspective with research into the motives of the people who share misinformation, both intentionally and unintentionally. Moreover, it is worthwhile to compare experts' views on health misinformation and our proposed design solutions with the users' perspectives.

9 CONCLUSION

In our qualitative study, we explored user perceptions of health misinformation and current social media countermeasures alongside users' expectations for appropriate interventions. We discovered that health misinformation is characterized by its ambiguous nature, an evolving status due to new evidence, and the often required specialized knowledge for comprehension. We found that the current binary true/false labeling strategies are overly simplistic and fail to address the complexity of health information dissemination. We identified four user-provided heuristics for discerning health misinformation, recommending their incorporation by social media platforms to enhance detection accuracy. Additionally, our research underlines the critical need for greater transparency regarding fact-checkers and influencers promoting health content, a step vital for empowering users to make informed decisions. Our observations of varied user behaviors in private versus public settings further illuminate the challenges and limited effectiveness of existing social platform measures. Our findings offer valuable insights for future research aimed at understanding users' reactions to health misinformation. Furthermore, our design recommendations pave the way for next-generation digital tools. These tools should not merely restrict users' interactions with information but should equip users with the means to critically assess and make informed decisions, fostering the growth of critical thinking.

ACKNOWLEDGMENTS

We thank our colleagues from the University of Luxembourg for participating in pretests and giving us detailed feedback on various stages of the project development. We also extend our thanks to our participants who engaged actively in this study and provided us with their valuable insights. Additionally, we thank our anonymous reviewers for their detailed comments, which have significantly helped us improve the work. The research was supported by the Luxembourg National Research Fund (REMEDIS, REgulatory and other solutions to MitigatE online DISinformation (INTER/FNRS/21/16554939)).

REFERENCES

[1] [n. d.]. About Community Notes on X | X Help — help.twitter.com. https://help.twitter.com/en/using-x/community-notes. [Accessed 07-02-2024].

[2] [n. d.]. How to prevent the spread of misinformation | WhatsApp Help Center — faq.whatsapp.com. https://faq.whatsapp.com/431498999157251. [Accessed 01-02-2024].

[3] Zhila Aghajari, Eric PS Baumer, and Dominic DiFranzo. 2023. Reviewing Interventions to Address Misinformation: The Need to Expand Our Vision Beyond an Individualistic Focus. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–34.

[4] Thomas Aichner, Matthias Grünfelder, Oswin Maurer, and Deni Jegeni. 2021. Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019. *Cyberpsychology, Behavior, and Social Networking* 24, 4 (April 2021), 215–222. https://doi.org/10.1089/cyber.2020.0134

[5] Khalid AJ Al Khaja, Alwaleed K AlKhaja, and Reginald P Sequeira. 2018. Drug information, misinformation, and disinformation on social media: a content analysis study. *Journal of public health policy* 39 (2018), 343–357.

[6] Md Sayeed Al-Zaman. 2022. Prevalence and source analysis of COVID-19 misinformation in 138 countries. *IFLA journal* 48, 1 (2022), 189–204.

[7] Zaid Amin, Nazlena Mohamad Ali, and Alan F Smeaton. 2021. Visual Selective Attention System to Intervene User Attention in Sharing COVID-19 Misinformation. *arXiv preprint arXiv:2110.13489* (2021).

[8] Simge Andı and Jesper Akesson. 2020. Nudging away false news: Evidence from a social norms experiment. *Digital Journalism* 9, 1 (2020), 106–125.

[9] Zahra Ashktorab and Jessica Vitak. 2016. Designing cyberbullying mitigation and prevention solutions through participatory design with teenagers. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, USA, 3895–3905.

[10] Dimitrios I Athanasiadis, Alexandra Roper, William Hilgendorf, Adam Voss, Taylor Zike, Marisa Embry, Ambar Banerjee, Don Selzer, and Dimitrios Stefanidis. 2021. Facebook groups provide effective social support to patients after bariatric surgery. *Surgical Endoscopy* 35 (2021), 4595–4601.

[11] AVAAZ. 2020. Facebook's Algorithm: A Major Threat to Public Health — secure.avaaz.org. https://secure.avaaz.org/campaign/en/facebook_threat_health/. [Accessed 11-09-2023].

[12] Avaaz. 2020. How Facebook can Flatten the Curve of the Coronavirus Infodemic. https://secure.avaaz.org/campaign/en/facebook_coronavirus_misinformation/. Accessed: 2023-10-02.

[13] Joseph B Bak-Coleman, Ian Kennedy, Morgan Wack, Andrew Beers, Joseph S Schafer, Emma S Spiro, Kate Starbird, and Jevin D West. 2022. Combining

interventions to reduce the spread of viral misinformation. *Nature Human Behaviour* 6, 10 (2022), 1372–1380.

[14] Arjun Narayan Bettadapur. 2020. TikTok partners with fact-checking experts to combat misinformation. https://newsroom.tiktok.com/en-au/tiktok-partners-with-fact-checking-experts-to-combat-misinformation

[15] Md Momen Bhuiyan, Michael Horning, Sang Won Lee, and Tanushree Mitra. 2021. Nudgecred: Supporting news credibility assessment on social media through nudges. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–30.

[16] Md Momen Bhuiyan, Hayden Whitley, Michael Horning, Sang Won Lee, and Tanushree Mitra. 2021. Designing transparency cues in online news platforms to promote trust: Journalists'& consumers' perspectives. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–31.

[17] Sophie C Boerman, Lotte M Willemsen, and Eva P Van Der Aa. 2017. "This post is sponsored" effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of Facebook. *Journal of Interactive Marketing* 38, 1 (2017), 82–92.

[18] Alexandra Bornkessel, Robert Furberg, and R Craig Lefebvre. 2014. Social media: opportunities for quality improvement and lessons for providers—a networked model for patient-centered care through digital engagement. *Current cardiology reports* 16 (2014), 1–9.

[19] Sharon S Brehm and Jack W Brehm. 2013. *Psychological reactance: A theory of freedom and control*. Academic Press.

[20] Mersine A Bryan, Yolanda Evans, Carlie Morishita, Nikita Midamba, and Megan Moreno. 2020. Parental perceptions of the internet and social media as a source of pediatric health information. *Academic pediatrics* 20, 1 (2020), 31–38.

[21] Enrico Bunde, Niklas Kühl, and Christian Meske. 2021. Fake or Credible? Towards Designing Services to Support Users' Credibility Assessment of News Content. *arXiv preprint arXiv:2109.13336* (2021).

[22] Cody Buntain, Richard Bonneau, Jonathan Nagler, and Joshua A Tucker. 2021. YouTube recommendations and effects on sharing across online social platforms. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–26.

[23] Valerio Capraro and Tatiana Celadin. 2022. "I think this news is accurate": Endorsing accuracy decreases the sharing of fake news and increases the sharing of real news. *Personality and Social Psychology Bulletin* (2022), 01461672221117691.

[24] Hillary Carey, Alexandra To, Jessica Hammer, and Geoff Kaufman. 2020. Fictional, interactive narrative as a foundation to talk about racism. In *Companion publication of the 2020 ACM designing interactive systems conference*. Association for Computing Machinery, New York, USA, 171–177.

[25] John M Carey, Andrew M Guess, Peter J Loewen, Eric Merkley, Brendan Nyhan, Joseph B Phillips, and Jason Reifler. 2022. The ephemeral effects of fact-checks on COVID-19 misperceptions in the United States, Great Britain and Canada. *Nature Human Behaviour* 6, 2 (2022), 236–243.

[26] Canyu Chen, Haoran Wang, Matthew Shapiro, Yunyu Xiao, Fei Wang, and Kai Shu. 2022. Combating Health Misinformation in Social Media: Characterization, Detection, Intervention, and Open Issues. *arXiv preprint arXiv:2211.05289* (2022).

[27] Guangyu Chen, Paolo Ciuccarelli, and Sara Colombo. 2022. VisualBubble: Exploring How Reflection-Oriented User Experiences Affect Users' Awareness of Their Exposure to Misinformation on Social Media. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. Association for Computing Machinery, New York, USA, 1–7.

[28] Junhan Chen and Yuan Wang. 2021. Social media use for health purposes: systematic review. *Journal of medical Internet research* 23, 5 (2021), e17917.

[29] Liang Chen and Hongjie Tang. 2022. Examining the persuasion process of narrative fear appeals on health misinformation correction. *Information, Communication & Society* (2022), 1–19.

[30] Wen-Ying Sylvia Chou, Anna Gaysynsky, and Yuki Lama. 2022. The Health Misinformation Ecosystem on Social Media. *Combating Online Health Misinformation: A Professional's Guide to Helping the Public* (2022), 31.

[31] Joanna TW Chu, Man Ping Wang, Chen Shen, Kasisomayajula Viswanath, Tai Hing Lam, and Sophia Siu Chee Chan. 2017. How, when and why people seek health information online: qualitative study in Hong Kong. *Interactive journal of medical research* 6, 2 (2017), e7000.

[32] Nicole K Dalmer. 2017. Questioning reliability assessments of health information on social media. *Journal of the Medical Library Association: JMLA* 105, 1 (2017), 61.

[33] Viorela Dan and Graham N Dixon. 2021. Fighting the infodemic on two fronts: Reducing false beliefs without increasing polarization. *Science Communication* 43, 5 (2021), 674–682.

[34] Kahneman Daniel. 2017. *Thinking, fast and slow*. Farrar, Straus and Giroux, New York, USA.

[35] Valdemar Danry, Pat Pataranutaporn, Yaoli Mao, and Pattie Maes. 2023. Don't Just Tell Me, Ask Me: AI Systems that Intelligently Frame Explanations as Questions Improve Human Logical Discernment Accuracy over Causal AI explanations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, USA, 1–13.

[36] Munmun De Choudhury, Meredith Ringel Morris, and Ryen W White. 2014. Seeking and sharing health information online: comparing search engines and social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*. Association for Computing Machinery, New York, USA, 1365–1376.

[37] Israel Junior Borges Do Nascimento, Ana Beatriz Pizarro, Jussara M Almeida, Natasha Azzopardi-Muscat, Marcos André Gonçalves, Maria Björklund, and David Novillo-Ortiz. 2022. Infodemics and health misinformation: a systematic review of reviews. *Bulletin of the World Health Organization* 100, 9 (2022), 544.

[38] Ullrich KH Ecker, Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K Fazio, Nadia Brashier, Panayiota Kendeou, Emily K Vraga, and Michelle A Amazeen. 2022. The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology* 1, 1 (2022), 13–29.

[39] Ullrich KH Ecker, Caitlin XM Sharkey, and Briony Swire-Thompson. 2023. Correcting vaccine misinformation: A failure to replicate familiarity or fear-driven backfire effects. *Plos one* 18, 4 (2023), e0281140.

[40] Ziv Epstein, Gordon Pennycook, and David Rand. 2020. Will the crowd game the algorithm? Using layperson judgments to combat misinformation on social media by downranking distrusted sources. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, USA, 1–11.

[41] KJ Kevin Feng, Kevin Song, Kejing Li, Oishee Chakrabarti, and Marshini Chetty. 2022. Investigating How University Students in the United States Encounter and Deal With Misinformation in Private {WhatsApp} Chats During {COVID-19}. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. USENIX, Berkeley, USA, 427–446.

[42] Financesonline.com. 2023. 15 Current Instagram Trends and Forecasts for 2022/2023 According to Experts. https://financesonline.com/instagram-trends/ Last accessed 26 August 2023.

[43] Martin Flintham, Christian Karner, Khaled Bachour, Helen Creswick, Neha Gupta, and Stuart Moran. 2018. Falling for fake news: investigating the consumption of news via social media. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, USA, 1–10.

[44] Frans Folkvord, Freek Snelting, Doeschka Anschutz, Tilo Hartmann, Alexandra Theben, Laura Gunderson, Ivar Vermeulen, and Francisco Lupiáñez-Villanueva. 2022. Effect of source type and protective message on the critical evaluation of news messages on Facebook: Randomized controlled trial in the Netherlands. *Journal of Medical Internet Research* 24, 3 (2022), e27945.

[45] Errol Francis, Ayana Monroe, Emily Sidnam-Mauch, Bernat Ivancsics, Eve Washington, Susan E McGregor, Joseph Bonneau, and Kelly Caine. 2023. Transparency, Trust, and Security Needs for the Design of Digital News Authentication Tools. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–44.

[46] Anna Gaysynsky, Kathryn Romansky-Poulin, and Stephen Arpadi. 2015. "My YAP Family": analysis of a Facebook group for young adults living with HIV. *AIDS and Behavior* 19 (2015), 947–962.

[47] Christine Geeng, Savanna Yee, and Franziska Roesner. 2020. Fake news on Facebook and Twitter: Investigating how people (don't) investigate. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, USA, 1–14.

[48] Thomas Germain. 2020. Twitter Asks Some Users: Have You Read the Article You're Retweeting? — consumerreports.org. https://www.consumerreports.org/social-media/new-twitter-feature-misinformation/. [Accessed 10-09-2023].

[49] Amira Ghenai and Yelena Mejova. 2018. Fake cures: user-centric modeling of health misinformation in social media. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–20.

[50] Tarleton Gillespie. 2022. Do not recommend? Reduction as a form of content moderation. *Social Media+ Society* 8, 3 (2022), 20563051221117552.

[51] Eric Goldman. 2021. Content moderation remedies. *Mich. Tech. L. Rev.* 28 (2021), 1.

[52] Garth Graham. 2021. Introducing new ways to help you find answers to your health questions. https://blog.youtube/news-and-events/introducing-new-ways-help-you-find-answers-your-health-questions/

[53] Colin M Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L Toombs. 2018. The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, USA, 1–14.

[54] Jonathan Grudin. 2006. Why Personas Work: The Psychological Evidence. In *Persona Lifecycle, Keeping People in Mind Throughout Product Design*. Morgan Kaufman, 643–663.

[55] Chen Guo, Nan Zheng, and Chengqi Guo. 2023. Seeing is Not Believing: A Nuanced View of Misinformation Warning Efficacy on Video-Sharing Social Media Platforms. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–35.

[56] Natalie-Anne Hall, Brendan Lawson, Cristian Vaccari, and Andrew Chadwick. [n. d.]. Beyond quick fixes: How users make sense of misinformation warnings on personal messaging. ([n. d.]).

[57] Natalie-Anne Hall, Brendan T Lawson, Cristian Vaccari, and Andrew Chadwick. 2023. WhatsApp's 'forwarded' and 'forwarded many times' tags may not prevent the spread of misinformation as they are often misunderstood by users, new report suggests — lboro.ac.uk. https://www.lboro.ac.uk/news-events/news/

2023/june/whatsapp-forwarded-tags-misunderstood-report/. [Accessed 01-02-2024].

[58] Katrin Hartwig, Frederic Doell, and Christian Reuter. 2023. The Landscape of User-centered Misinformation Interventions–A Systematic Literature Review. *arXiv preprint arXiv:2301.06517* (2023).

[59] Gina Hernandez. 2021. New prompts to help people consider before they share — newsroom.tiktok.com. https://newsroom.tiktok.com/en-us/new-prompts-to-help-people-consider-before-they-share. [Accessed 10-09-2023].

[60] Ralph Hertwig and Till Grüne-Yanoff. 2017. Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science* 12, 6 (2017), 973–986.

[61] Benjamin D Horne, Dorit Nevo, and Susan L Smith. 2023. Ethical and safety considerations in automated fake news detection. *Behaviour & Information Technology* 1, 1 (2023), 1–22.

[62] Shandell Houlden, Jaigris Hodson, George Veletsianos, Darren Reid, and Chris Thompson-Wagner. 2021. The health belief model: How public health can address the misinformation crisis beyond COVID-19. *Public Health in Practice* 2 (2021), 100151.

[63] Mowafa Househ, Elizabeth Borycki, and Andre Kushniruk. 2014. Empowering patients through social media: the benefits and challenges. *Health informatics journal* 20, 1 (2014), 50–58.

[64] Yoori Hwang and Se-Hoon Jeong. 2023. Education-Based Gap in Misinformation Acceptance: Does the Gap Increase as Misinformation Exposure Increases? *Communication Research* 50, 2 (2023), 157–178.

[65] Farnaz Jahanbakhsh, Amy X Zhang, Adam J Berinsky, Gordon Pennycook, David G Rand, and David R Karger. 2021. Exploring lightweight interventions at posting time to reduce the sharing of misinformation on social media. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–42.

[66] Ali Jarrahi and Leila Safari. 2023. Evaluating the effectiveness of publishers' features in fake news detection on social media. *Multimedia Tools and Applications* 82, 2 (Jan. 2023), 2913–2939. https://doi.org/10.1007/s11042-022-12668-8

[67] Prerna Juneja and Tanushree Mitra. 2022. Human and technological infrastructures of fact-checking. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–36.

[68] Alla Keselman, Catherine Arnott Smith, Gondy Leroy, and David R Kaufman. 2021. Factors influencing willingness to share health misinformation videos on the Internet: Web-based survey. *Journal of medical Internet research* 23, 12 (2021), e30323.

[69] Jan Kirchner and Christian Reuter. 2020. Countering fake news: A comparison of possible solutions regarding user acceptance and effectiveness. *Proceedings of the ACM on Human-computer Interaction* 4, CSCW2 (2020), 1–27.

[70] Jasmijn Kruijt, Corine S Meppelink, and Lisa Vandeberg. 2022. Stop and think! Exploring the role of news truth discernment, information literacy, and impulsivity in the effect of critical thinking recommendations on trust in fake COVID-19 news. *European Journal of Health Communication* 3, 2 (2022), 40–63.

[71] Udo Kuckartz and Stefan Rädiker. 2019. *Analyzing qualitative data with MAXQDA*. Springer, Cham, Switzerland.

[72] Eleni A Kyza and Christiana Varda. 2019. NAVIGATING THE POST-TRUTH ERA: TRUST, MISINFORMATION, AND CREDIBILITY ASSESSMENT ON ONLINE SOCIAL MEDIA. *AoIR Selected Papers of Internet Research* 2019, 1 (2019), 1–4.

[73] Jisan Lee, Jongkwan Koh, and Jong-Yeup Kim. 2021. Popularization of medical information. *Healthcare Informatics Research* 27, 2 (2021), 110–115.

[74] Stephan Lewandowsky, Ullrich KH Ecker, Colleen M Seifert, Norbert Schwarz, and John Cook. 2012. Misinformation and its correction: Continued influence and successful debiasing. *Psychological science in the public interest* 13, 3 (2012), 106–131.

[75] Yang-Jun Li, Christy MK Cheung, Xiao-Liang Shen, and Matthew KO Lee. 2019. Health Misinformation on Social Media: A Literature Review. In *23rd Pacific Asia Conference on Information Systems (PACIS 2019): Secure ICT Platform for the 4th Industrial Revolution*. Association for Computing Machinery, New York, USA, 194.

[76] Matheus Lotto, Tamires Sá Menezes, Irfhana Zakir Hussain, Shu-Feng Tsao, Zahid Ahmad Butt, Plinio P Morita, and Thiago Cruvinel. 2022. Characterization of false or misleading fluoride content on Instagram: infodemiology study. *Journal of Medical Internet Research* 24, 5 (2022), e37519.

[77] Tessa Lyons. 2018. Hard Questions: How Is Facebook's Fact-Checking Program Working? https://www.facebook.com/business/help/2593586717571940?id=673052479947730

[78] Shana MacDonald and Wiens Brianna. 2023. Meme-ifying Data: The Rise of Public Health Influencers on Instagram, TikTok, and Twitter during Covid-19. *Journal of Digital Social Research* 5, 3 (2023), 59–84.

[79] Golshan Madraki, Isabella Grasso, Jacqueline M. Otala, Yu Liu, and Jeanna Matthews. 2021. Characterizing and comparing COVID-19 misinformation across languages, countries and platforms. In *Companion proceedings of the web conference 2021*. Association for Computing Machinery, New York, USA, 213–223.

[80] Kristen MC Malecki, Julie A Keating, and Nasia Safdar. 2021. Crisis communication and public perception of COVID-19 risk in the era of social media. *Clinical infectious diseases* 72, 4 (2021), 697–702.

[81] Pranav Malhotra. 2020. <? covid19?> A Relationship-Centered and Culturally Informed Approach to Studying Misinformation on COVID-19. *Social Media+ Society* 6, 3 (2020), 2056305120948224.

[82] Pranav Malhotra and Katy Pearce. 2022. Facing falsehoods: strategies for polite misinformation correction. *International Journal of Communication* 16 (2022), 22.

[83] Sumayyia D Marar, May M Al-Madaney, and Fatimah H Almousawi. 2019. Health information on social media.: Perceptions, attitudes, and practices of patients and their companions. *Saudi medical journal* 40, 12 (2019), 1294.

[84] Moacir Marocolo, Anderson Meireles, Hiago Leandro Rodrigues de Souza, Gustavo Ribeiro Mota, Dustin Jay Oranchuk, Rhaí André Arriel, and Laura Hora Rios Leite. 2021. Is social media spreading misinformation on exercise and health in Brazil? *International journal of environmental research and public health* 18, 22 (2021), 11914.

[85] Cameron Martel, Jennifer Allen, Gordon Pennycook, and David G. Rand. 2024. Crowds can effectively identify misinformation at scale. *Perspectives on Psychological Science* 19, 2 (2024), 477–488.

[86] Antônio Diogo Forte Martins, Lucas Cabral, Pedro Jorge Chaves Mourao, Ivandro Claudino de Sá, José Maria Monteiro, and Javam Machado. 2021. COVID19. br: A dataset of misinformation about COVID-19 in brazilian portuguese whatsapp messages. In *Anais do III Dataset Showcase Workshop*. SBC, Journal of the Brazilian Computer Society, Porto Alegre ,Brazil, 138–147.

[87] Noel H McGuire and Hannah Ball. 2022. Extending psychological reactance theory to include denial of threat and media sharing intentions as freedom restoration behavior. *Communication research reports* 39, 3 (2022), 136–146.

[88] Cristiane Melchior and Mirian Oliveira. 2022. Health-related fake news on social media platforms: A systematic literature review. *new media & society* 24, 6 (2022), 1500–1522.

[89] Miriam J Metzger, Andrew J Flanagin, Paul Mena, Shan Jiang, and Christo Wilson. 2021. From dark to light: The many shades of sharing misinformation online. *Media and Communication* 9, 1 (2021), 134–143.

[90] S Anne Moorhead, Diane E Hazlett, Laura Harrison, Jennifer K Carroll, Anthea Irwin, and Ciska Hoving. 2013. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of medical Internet research* 15, 4 (2013), e1933.

[91] Patricia L Moravec, Antino Kim, and Alan R Dennis. 2020. Appealing to sense and sensibility: System 1 and system 2 interventions for fake news on social media. *Information Systems Research* 31, 3 (2020), 987–1006.

[92] Sadiq Muhammed T and Saji K Mathew. 2022. The disaster of misinformation: a review of research in social media. *International journal of data science and analytics* 13, 4 (2022), 271–285.

[93] Imani Munyaka, Eszter Hargittai, and Elissa Redmiles. 2022. The Misinformation Paradox: Older Adults are Cynical about News Media, but Engage with It Anyway. *Journal of Online Trust and Safety* 1, 4 (2022), 1–16.

[94] Xiaoli Nan, Yuan Wang, and Kathryn Thier. 2022. Why people believe health misinformation and who are at risk? A systematic review of individual differences in susceptibility to health misinformation. *Social Science & Medicine* 314, 2022 (2022), 1–17.

[95] John A Naslund, Kelly A Aschbrenner, Lisa A Marsch, and Stephen J Bartels. 2016. The future of mental health care: peer-to-peer support and social media. *Epidemiology and psychiatric sciences* 25, 2 (2016), 113–122.

[96] Stephen Neely, Christina Eldredge, and Ron Sanders. 2021. Health information seeking behaviors on social media during the COVID-19 pandemic among American social networking site users: survey study. *Journal of medical Internet research* 23, 6 (2021), e29802.

[97] Mark W Newman, Debra Lauterbach, Sean A Munson, Paul Resnick, and Margaret E Morris. 2011. It's not that I don't have problems, I'm just not putting them on Facebook: challenges and opportunities in using online social networks for health. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. Association for Computing Machinery, New York, USA, 341–350.

[98] Anne Oeldorf-Hirsch, Mike Schmierbach, Alyssa Appelman, and Michael P Boyle. 2020. The ineffectiveness of fact-checking labels on news memes and articles. *Mass Communication and Society* 23, 5 (2020), 682–704.

[99] Orestis Papakyriakopoulos and Ellen Goodman. 2022. The impact of Twitter labels on misinformation spread and user engagement: Lessons from Trump's election tweets. In *Proceedings of the ACM web conference 2022*. Association for Computing Machinery, New York, USA, 2541–2551.

[100] Irene V Pasquetto, Eaman Jahani, Shubham Atreja, and Matthew Baum. 2022. Social debunking of misinformation on WhatsApp: The case for strong and in-group ties. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–35.

[101] Tatyana A Petukhova, Britney N Wilson, Mariam Gadjiko, Erica H Lee, Jennifer Wang, Anthony M Rossi, and Kishwer S Nehal. 2020. Utilization of Facebook for support and education by patients with skin cancer. *Dermatology Online Journal* 26, 3 (2020), 1–4.

[102] Whitney Phillips and Ryan M. Milner. 2021. *You are here: a field guide for navigating polarized speech, conspiracy theories, and our polluted media landscape.* The MIT Press, Cambridge, Massachusetts. OCLC: 1202266353.

[103] Toby Prike, Phoebe Blackley, Briony Swire-Thompson, and Ullrich KH Ecker. 2023. Examining the replicability of backfire effects after standalone corrections. *Cognitive Research: Principles and Implications* 8, 1 (2023), 39.

[104] Emma K Quinn, Sajjad S Fazel, and Cheryl E Peters. 2021. The Instagram infodemic: cobranding of conspiracy theories, coronavirus disease 2019 and authority-questioning beliefs. *Cyberpsychology, Behavior, and Social Networking* 24, 8 (2021), 573–577.

[105] Yoel Roth and Nick Pickles. 2020. Updating our approach to misleading information. https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information

[106] Emily Saltz, Claire R Leibowicz, and Claire Wardle. 2021. Encounters with visual misinformation and labels across platforms: An interview and diary study to inform ecosystem approaches to misinformation interventions. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, USA, 1–6.

[107] Emily Saltz, Tommy Shane, Victoria Kwan, Claire Leibowicz, and Claire Wardle. 2020. It matters how platforms label manipulated media. Here are 12 principles designers should follow.-The Partnership on AI.

[108] Lauren Scott, Lynne Coventry, Marta E Cecchinato, and Mark Warner. 2023. "I figured her feeling a little bit bad was worth it to not spread that kind of hate": Exploring how UK families discuss and challenge misinformation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, USA, 1–15.

[109] Gautam Kishore Shahi, Anne Dirkson, and Tim A Majchrzak. 2021. An exploratory study of COVID-19 misinformation on Twitter. *Online social networks and media* 22 (2021), 100104.

[110] Imani N Sherman, Jack W Stokes, and Elissa M Redmiles. 2021. Designing Media Provenance Indicators to Combat Fake Media. In *Proceedings of the 24th International Symposium on Research in Attacks, Intrusions and Defenses.* Association for Computing Machinery, New York, USA, 324–339.

[111] Tristan M Sissung and William D Figg. 2021. Facebook groups for alternative treatments for cancer: advertising masquerading as community support. *The Lancet Oncology* 22, 1 (2021), 25–26.

[112] Edin Smailhodzic, Wyanda Hooijsma, Albert Boonstra, and David J Langley. 2016. Social media use in healthcare: A systematic review of effects on patients and on their relationship with healthcare professionals. *BMC health services research* 16, 1 (2016), 1–14.

[113] H Song, K Komori, J Kim, K Tenzek, J Hawkins, W Lin, and JY Jung. 2016. Trusting social media as a source of health information: Comparison among US, Korea, and Hong Kong. *Journal of Medical Internet Research* 18, 3 (2016), 1–12.

[114] Alain Strowel and Jean De Meyere. 2023. The Digital Services Act: transparency as an efficient tool to curb the spread of disinformation on online platforms? *J. Intell. Prop. Info. Tech. & Elec. Com. L.* 14 (2023), 66.

[115] Yanqing Sun and Fangcao Lu. 2023. How misinformation and rebuttals in online comments affect people's intention to receive COVID-19 vaccines: the roles of psychological reactance and misperceptions. *Journalism & Mass Communication Quarterly* 100, 1 (2023), 145–171.

[116] Briony Swire-Thompson, John Cook, Lucy H Butler, Jasmyne A Sanderson, Stephan Lewandowsky, and Ullrich KH Ecker. 2021. Correction format has a limited role when debunking misinformation. *Cognitive Research: Principles and Implications* 6, 1 (2021), 83.

[117] Briony Swire-Thompson, David Lazer, et al. 2020. Public health and online misinformation: challenges and recommendations. *Annu Rev Public Health* 41, 1 (2020), 433–451.

[118] Wee-Kheng Tan and Chun Yu Hsu. 2023. The application of emotions, sharing motivations, and psychological distance in examining the intention to share COVID-19-related fake news. *Online Information Review* 47, 1 (2023), 59–80.

[119] Richard H Thaler and Cass R Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness.* Penguin, London, UK.

[120] Nicholas B Tiller, John P Sullivan, and Panteleimon Ekkekakis. 2023. Baseless Claims and Pseudo science in Health and Wellness: A Call to Action for the Sports, Exercise, and Nutrition-Science Community. *Sports Medicine* 53, 1 (2023), 1–5.

[121] Justine Topham and Naomi Smith. 2023. One day of eating: Tracing misinformation in 'What I Eat In A Day'videos. *Journal of Sociology* 59, 3 (2023), 682–698.

[122] Sadegh Torabi and Konstantin Beznosov. 2016. Sharing health information on facebook: practices, preferences, and risk perceptions of North American users. In *Symposium on Usable Privacy and Security (SOUPS).* USENIX, Berkeley, USA, 301–320.

[123] Mbih J Tosam, Primus Che Chi, Nchangwi Syntia Munung, Odile Ouwe Missi Oukem-Boyer, and Godfrey B Tangwa. 2018. Global health inequalities and the need for solidarity: a view from the Global South. *Developing World Bioethics* 18, 3 (2018), 241–249.

[124] Sander Van Der Linden. 2022. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine* 28, 3 (2022), 460–467.

[125] Toni GLA Van der Meer and Yan Jin. 2020. Seeking formula for misinformation treatment in public health crises: The effects of corrective information type and source. *Health Communication* 35, 5 (2020), 560–575.

[126] Rama Adithya Varanasi, Joyojeet Pal, and Aditya Vashistha. 2022. Accost, accede, or amplify: attitudes towards COVID-19 misinformation on WhatsApp in India. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, USA, 1–17.

[127] C Lee Ventola. 2014. Social media and health care professionals: benefits, risks, and best practices. *Pharmacy and therapeutics* 39, 7 (2014), 491.

[128] Nathan Walter, John J Brooks, Camille J Saucier, and Sapna Suresh. 2021. Evaluating the impact of attempts to correct health misinformation on social media: A meta-analysis. *Health Communication* 36, 13 (2021), 1776–1784.

[129] Xin Wang, Peng Zhao, and Xi Chen. 2020. Fake news and misinformation detection on headlines of COVID-19 using deep learning algorithms. *International Journal of Data Science* 5, 4 (2020), 316–332.

[130] Yuxi Wang, Martin McKee, Aleksandra Torbica, and David Stuckler. 2019. Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. *Social Science & Medicine* 240 (Nov. 2019), 112552. https://doi.org/10.1016/j.socscimed.2019.112552

[131] Y. Wang, K. Thier, and X. Nan. 2022. *Defining health misinformation.* Rowman & Littlefield, Lanham, Maryland, Chapter 1, 3–16.

[132] Claire Wardle and Hossein Derakhshan. 2017. *Information Disorder: Toward an interdisciplinary framework for research and policymaking.* Council of Europe report DGI, Vol. 27. Council of Europe, Strasbourg. 1–107 pages.

[133] Erik Wende, Gregory King, and Gerhard Schwabe. 2014. Exploring Storytelling as a Knowledge Transfer Technique in Offshore Outsourcing. In *Proceedings of the Thirty Fifth International Conference on Information Systems* (Auckland, New Zealand). Association for Information Systems, Auckland, 1–16.

[134] Stefan Wojcik, Sophie Hilgard, Nick Judd, Delia Mocanu, Stephen Ragain, M. B. Fallin Hunzaker, Keith Coleman, and Jay Baxter. 2022. Birdwatch: Crowd Wisdom and Bridging Algorithms can Inform Understanding and Reduce the Spread of Misinformation. , 25 pages. arXiv:2210.15723 [cs.SI]

[135] David Xiang and Lisa Soleymani Lehmann. 2021. Confronting the misinformation pandemic. *Health Policy and Technology* 10, 3 (2021), 100520.

[136] Jason C Young. 2021. Disinformation as the weaponization of cruel optimism: A critical intervention in misinformation studies. *Emotion, Space and Society* 38 (2021), 100757.

[137] Amy X Zhang, Aditya Ranganathan, Sarah Emlen Metz, Scott Appling, Connie Moon Sehat, Norman Gilmore, Nick B Adams, Emmanuel Vincent, Jennifer Lee, Martin Robbins, et al. 2018. A structured response to misinformation: Defining and annotating credibility indicators in news articles. In *Companion Proceedings of the The Web Conference 2018.* International World Wide Web Conference Committee, Lyon, France, 603–612. https://doi.org/10.1145/3184558.3188731

[138] Jingwen Zhang, Jieyu Ding Featherstone, Christopher Calabrese, and Magdalena Wojcieszak. 2021. Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Preventive Medicine* 145 (2021), 106408.

[139] Shuai Zhang, Feicheng Ma, Yunmei Liu, and Wenjing Pian. 2022. Identifying features of health misinformation on social media sites: an exploratory analysis. *Library Hi Tech* 40, 5 (2022), 1384–1401.

[140] Yuqi Zhang, Bin Guo, Yasan Ding, Jiaqi Liu, Chen Qiu, Sicong Liu, and Zhiwen Yu. 2022. Investigation of the determinants for misinformation correction effectiveness on social media during COVID-19 pandemic. *Information Processing & Management* 59, 3 (2022), 102935.

[141] Junzhang Zhao, Hongxing Han, Bu Zhong, Wenjing Xie, Ying Chen, and Min Zhi. 2021. Health information on social media helps mitigate Crohn's disease symptoms and improves patients' clinical course. *Computers in Human Behavior* 115 (2021), 106588.

[142] Yuehua Zhao, Jingwei Da, and Jiaqi Yan. 2021. Detecting health misinformation in online health communities: Incorporating behavioral features into machine learning based approaches. *Information Processing & Management* 58, 1 (2021), 102390.

[143] Yuehua Zhao and Jin Zhang. 2017. Consumer health information seeking in social media: a literature review. *Health Information & Libraries Journal* 34, 4 (2017), 268–283.

[144] Bu Zhong. 2023. Going beyond fact-checking to fight health misinformation: A multi-level analysis of the Twitter response to health news stories. *International Journal of Information Management* 70 (2023), 102626.

[145] Xinyi Zhou and Reza Zafarani. 2021. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *Comput. Surveys* 53, 5 (Sept. 2021), 1–40. https://doi.org/10.1145/3395046

A APPENDIX

The demographic details of participants and their social media experience are provided in Table 2.

Table 2: Demographic Details and Social Media Experience of Participants

ID	Social Media platform usage	Age	Gender	Professional Background	Origin
P1	Facebook, Twitter, WhatsApp, Youtube, Discord, Snapchat	27	M	Software development	Prefer not to say
P2	WhatsApp, Youtube, Tiktok	25	F	Civil engineering	Latin America& Western Europe
P3	Whatsapp, Instagram, Signal, Discord, Youtube, WeChat	22	M	Physics	Western Europe
P4	Facebook, WhatsApp, Youtube, Instagram	39	M	Doctoral researcher	West Asia
P5	Facebook, Twitter, WhatsApp, Instagram, Youtube, Tiktok, Snapchat	31	M	Accounting and audit	Sub-Saharan Africa
P6	WhatsApp, Youtube, Instagram	26	F	Data science	West Asia
P7	Facebook, WhatsApp, Instagram, Youtube,	33	F	Architecture	South Asia
P8	Facebook, Twitter, WhatsApp, Instagram, Youtube,	27	M	European philosophy	Sub-Saharan Africa
P9	Facebook, Twitter, WhatsApp, Instagram,Youtube, Tiktok,	27	F	Biomathematics	South Asia
P10	Facebook, Twitter, WhatsApp, Instagram, Youtube	29	M	Finance and economics	East Asia
P11	Facebook, Twitter, WhatsApp, Instagram, Youtube, Tiktok,	24	F	Political science	Latin America
P12	Facebook, WhatsApp, Instagram, Youtube,	30	M	Energy efficiency	Sub-Saharan Africa
P13	WhatsApp, Instagram	29	F	Doctoral researcher	Central Europe
P14	Facebook, Twitter, WhatsApp, Instagram, Youtube	29	F	Computer science	Eastern Europe
P15	Facebook, Twitter, WhatsApp, Instagram, Youtube, Tiktok, Snapchat, Pinterest, Discord	21	F	Biology	Western Europe
P16	Facebook, WhatsApp, Instagram, Youtube, Tiktok	29	F	Psychology	Western Europe
P17	Facebook, Twitter, WhatsApp, Instagram, Youtube,	30	M	Wealth management	South Asia
P18	Facebook, WhatsApp, Instagram, Youtube, Reddit	24	M	Economics	West Asia
P19	Facebook, WhatsApp, Instagram, Youtube, Snapchat, Signal, Parfumo	23	F	Medicine	Western Europe
P20	Facebook, WhatsApp, Instagram, Youtube,	35	M	Engineering	West Asia
P21	Facebook, WhatsApp, Youtube	28	F	Computer science	Eastern Europe
P22	WhatsApp, Instagram	31	F	Learning and communication	Latin America