Investigating How University Students in the United States Encounter and Deal With Misinformation in Private WhatsApp Chats During COVID-19

K. J. Kevin Feng
Princeton University

Kevin Song
University of Chicago

Kejing Li
University of Chicago

Oishee Chakrabarti
University of Chicago

Marshini Chetty
University of Chicago

Abstract

Misinformation can spread easily in end-to-end encrypted messaging platforms such as WhatsApp where many groups of people are communicating with each other. Approaches to combat misinformation may also differ amongst younger and older adults. In this paper, we investigate how young adults encountered and dealt with misinformation on WhatsApp in private group chats during the first year of the COVID-19 pandemic. To do so, we conducted a qualitative interview study with 16 WhatsApp users who were university students based in the United States. We uncovered three main findings. First, all participants encountered misinformation multiple times a week in group chats, often attributing the source of misinformation to be well-intentioned family members. Second, although participants were able to identify misinformation and fact-check using diverse methods, they often remained passive to avoid negatively impacting family relations. Third, participants agreed that WhatsApp bears a responsibility to curb misinformation on the platform but expressed concerns about its ability to do so given the platform's steadfast commitment to content privacy. Our findings suggest that conventional content moderation techniques used by open platforms such as Twitter and Facebook are unfit to tackle misinformation on WhatsApp. We offer alternative design suggestions that take into consideration the social nuances and privacy commitments of end-to-end encrypted group chats. Our paper also contributes to discussions between platform designers, researchers, and end users on misinformation in privacy-preserving environments more broadly.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2022.
August 7–9, 2022, Boston, MA, United States.

1 Introduction

WhatsApp is a widely used end-to-end encrypted messaging platform worldwide, with an estimated 74 million users in the United States (U.S.) alone as of 2021 [4]. The platform's widespread usage rose sharply with the global spread of COVID-19. By late March 2020, WhatsApp grew by 40% compared to pre-pandemic months [55]; this growth was likely fueled by its connective capabilities during the pandemic, such as for organizing mutual aid groups [16] and, in the case of millions of immigrants, connecting with family members abroad [42]. WhatsApp's end-to-end encryption [80] means that the platform is unable to easily detect or flag misleading messages, i.e., misinformation [1], which is problematic given its global user base [71]. It has therefore been identified as an effective misinformation pipeline by academics, journalists, and fact-checking organizations [31, 56, 74]. Consequences of this rapid dissemination of misinformation on the platform include the spread of misleading health claims and associated health risks [27, 39], tampering of elections abroad [5], and deaths [10, 34].

Many researchers have studied characteristics of online misinformation including prevalence [1, 22, 38], speed of spread [37], user perceptions [26, 32], and strategic participatory campaigns [67]. However, research on misinformation in WhatsApp specifically has been limited and mainly focuses on users outside of the U.S. [6, 41, 49]. These studies observe user behavior through theoretical frameworks and collect message content from large public WhatsApp groups [31, 41, 46, 49] rather than using empirical user studies of private chats[2] [25, 45, 46, 57, 58]. Private chats yield valuable insights into users' daily communication practices

[1] In this paper, we use the definition of misinformation on social media presented by Wu et al. [85]: an umbrella term that includes all false or inaccurate information that is spread.

[2] A WhatsApp *private* chat can only be joined with an invitation link that is not typically shared publicly or when a group admin adds members to a group chat. A WhatsApp *public* chat can be joined by anyone on the Internet via an invitation link that is usually posted on a public website, making it easier for researchers to study.

since WhatsApp users mainly communicate in small, preselected groups of people [64], notably families. Although misinformation within smaller private group chats may not be broadcasted to large audiences at once, they can still reach high numbers of users through group chats' popularity and frequent forwarding activity between chats [46].

To properly combat misinformation on WhatsApp, we need a better understanding of how WhatsApp users deal with misleading messages, particularly in private chats. Since there is a generally an unreciprocated concern directed towards older family members about health misinformation due to them being perceived as a vulnerable population on the Internet [69], we also need to balance this out with an investigation of the perspectives of younger adults around misinformation on WhatsApp. To address this research gap, we conducted interviews with 16 young adults who were university students in the U.S.—a country with the third most WhatsApp users globally [70]—to better understand their experiences with COVID-19-related misinformation in close-knit private chats. Our study was driven by the following research questions:

- **RQ1**: How do U.S.-based university students currently perceive and encounter misinformation in WhatsApp private chats?
- **RQ2**: How do U.S.-based university students identify misinformation on the platform and respond to it?
- **RQ3**: How aware are U.S.-based university students of current WhatsApp features to combat misinformation and what would improve how the platform handles misinformation?

We uncovered three main findings. First, all participants encountered misinformation multiple times a week in group chats, often attributing the source of misinformation to be well-intentioned family members. Most participants also claimed not to forward information without fact-checking first. Second, although participants were able to identify misinformation using similar indicators seen in previous studies on other social media platforms [26, 32, 47], they often did not confront misinformation senders to avoid negatively impacting family relations. Third, participants were not aware of most existing features to combat misinformation on WhatsApp and agreed that WhatsApp bears a responsibility to curb misinformation on the platform. However, participants expressed concerns about its ability to do so given the platform's commitment to content privacy. Based on our findings, we suggest, assuming users can be made more aware of new features, that empowering users on the platform to better fact-check or flag misinformation for themselves may combat the effects of misleading content. We also suggest that designs that allow users to subtly provide resources for misleading messages within a group could offset the power dynamics in chats that prevent users from confronting misinformation senders. Future work should investigate older adults' role in misinformation on WhatsApp and how to educate users about misinformation leveraging the fact that misinformation is often spread out of care and not malicious intent.

To summarize, our primary contributions are:

- **Findings from a U.S.-based WhatsApp user study**: we contribute novel insights about how U.S.-based WhatsApp university students in our study perceived and reacted to misinformation in private WhatsApp chats. For instance, we found that our participants felt that misinformation was often sent to them from well-intentioned family members out of care for others and that family dynamics make it harder for younger adults to confront older misinformation senders. This contributes to a growing set of studies of public WhatsApp chat data [25, 45, 46, 57, 58].
- We corroborate findings from misinformation studies on other social media platforms such as Facebook and news [26, 32, 47] about the indicators people use to identify misleading content; adding a novel finding about how WhatsApp users weigh the relationship with a misinformation sender to determine if content can be trusted.
- Finally, our paper adds to the literature on how to tackle misinformation in end-to-end encrypted platforms that conventional content moderation techniques used by open platforms such as Twitter and Facebook cannot address, owing to the tradeoff between user-privacy and having to access data for labeling content [43].

Next, we describe related work, our methods, findings, and discussion points before concluding the paper.

2 Background and Related Work

2.1 Misinformation on Social Media

COVID-19 has swept the world, and so has the misinformation associated with it [6, 9, 36, 61, 72]. Kouzy et al. [36] estimates 25% of tweets include misinformation about the pandemic, while 17% include unverifiable information. To date, researchers have studied misinformation and its dissemination through social media extensively [3, 6, 15, 26, 32, 39, 50, 67]. Studies have also shown that misinformation's impact is global, from increasing tensions between neighboring countries [28], to suppressing government-critical voices within borders [52], to interfering with democratic elections [3, 14, 51]. Yet, the scale of social media and the Internet's replacement of expert advice make combating misinformation challenging [3, 39, 67].

To combat misinformation, some studies have explored users' motives for spreading news and misinformation on social media specifically and found that while most participants shared news to inform others, a third share for others' entertainment, with 19% doing so just to upset others [15]. Sharing misinformation can be influenced by culture as shown by Madrid-Morales et al. [50] who found that sharing habits differed by country and age in six sub-Saharan African countries. For example, some users in Kenya only shared tweets by verified Twitter accounts while students in South Africa shared news that was entertaining. Sometimes sharing misinformation depends on the content format. For instance, Singh et al. found that participants were more likely to share questionable claims on Twitter containing Uniform Resource Locators (URLs) with their friends than the same claims without URLs [66]. Often, once misinformation is shared, it is not corrected. For instance, prior works in the United Kingdom suggested that less than 20% of news sharers on social media are informed by others when they have shared dubious information [15] and on Facebook and Twitter, studies show that sometimes users ignore posts they consider misleading with no further action [26].

Other research has focused on the design of combative measures against misinformation. For instance, there have been qualitative experiments and surveys exposing users to 'fake news' on Facebook to see if and how they identified misleading content [22, 26]. Some studies found that lightweight interventions and frictions, such as nudging users to assess information accuracy or even preventing them from accessing known disinformation, helps users identify and avoid disinformation [32, 33]. Companies have also been employing warning labels and other strategies to combat misinformation. For example, Twitter encourages users to add their own commentary to a retweet [24], and Facebook displays a pop-up asking users if they want to share an article they have not yet opened [17]. Our study contributes to this body of knowledge by extending the study of users' encounters and responses to misinformation to WhatsApp private chats.

2.1.1 Generational Challenges With Misinformation

There has been debate in the academic community on whether web-based misinformation can amplify inter-generational gaps. For instance, concerns have been raised around older adults' susceptibility to misinformation due to their lack of experience with technology [48] and higher likelihood of deteriorating memory [60]. Researchers have investigated this phenomenon. Loos and Nihenhuis [40] tracked audience reach with deceptive Facebook ads linking to made-up news articles and found that the ads had higher reach amongst older age groups. Similarly, Madrid-Morales et al. [50] revealed that students and other younger users of social media in sub-Saharan Africa mostly blamed older generations for circulating fake news. Adding to this sentiment, Guess et al. [30] found older Americans more likely to share misinformation during the 2016 presidential election and Tandoc Jr. and Lee [69] found that young Singaporean adults in their 20s were more concerned for parents and older family members about uncertainty around COVID-19 information.

Yet studies about whether age plays a part in misinformation online are mixed [54]. For example, Trninic et al. [75] concluded that both younger and older populations lack media literacy upon measuring both groups' abilities to recognize, verify, and relate to misinformed content. Additionally, Brosius et al. [13] used survey data across 10 European countries and did not find differing levels of trust in media between generations. On the other hand, Wineburg and McGrew [84] suggest that younger generations of "digital natives" are especially at high risk of being duped by misinformation due to the amount of time spent on social media and the speed at which they consume online media. Some work even investigates younger population's perceptions of misinformation, from feeling frustrated [11], to being under peer pressure to consume certain media [23]. Yet despite previous work, we still lack a detailed empirical understanding of how younger users interact with misinformation-related topics in intergenerational environments such as WhatsApp family chats, particularly during times of crisis such as COVID-19. Our work serves to bridge this gap.

2.2 Misinformation on WhatsApp

The study of misinformation on WhatsApp is not new. Quantitative studies have explored misinformation dissemination on WhatsApp [25, 35, 41, 45, 46, 49, 53, 57, 58]. Using publicly available data from public WhatsApp group chats, researchers have studied the effects of limiting message forwarding on misinformation's spread on the platform [46][3], characteristics of misleading messages [57, 58], and percentages of false information in chats [35]. Studies have shown, for instance, that political and election-based misinformation is prevalent in WhatsApp group chats in Brazil [41], Indonesia [46], India [49], and Nigeria [31], among others. Researchers have typically focused on public WhatsApp group chats in their studies because these chats can be rampant misinformation spreaders and since anyone with an invitation link can join them, it makes data access for research easier. We focus on private WhatsApp chats since existing research lacks insight into misinformation encounters in private, direct messages or group chats with close friends and family. These chats can still be effective conduits for misinformation owing to forwarding on the platform [46].

In other studies of misinformation on WhatsApp, researchers have created tools for detecting misinformation and alerting users to these misleading messages. For instance, some qualitative studies examined public WhatsApp group

[3] WhatsApp introduced new forwarding limits in April 2020 [82]. Messages that are identified as "highly forwarded"—sent through a chain of five or more people—are marked with a double arrow icon and can only be forwarded to a single chat instead of 5. Prior to this change, in 2019, each message could be forwarded to a max of 20 chats [29], regardless of forwarding status.

chat messages [35, 41, 58] for detectable misinformation indicators such as excessively capitalized text and flashy images. In another study by Palomo and Sedano in Spain [53], they created a fact-checking tip line tool so that users could use WhatsApp as to verify claims in local news. Unlike our work, these researchers interviewed a chief editor of a local news publication rather than WhatsApp users themselves to inform design of the tool. Other researchers have developed automated misinformation detection approaches with limited success [25]. In Brazil, researchers also created WhatsApp Monitor, a tool intended to limit the spread of misinformation on WhatsApp in Brazil in public group chats [45]. However, due to WhatsApp's privacy policies and end-to-end encryption, the tool functioned as a window into the prevalence of various content categories (images, videos, audio, text) of misleading content in public WhatsApp chats for researchers rather than a direct intervention on misinformation for users. Finally, some work has looked at the efficacy of family chats in disseminating misinformation in Brazil [58] and Kenya [76].

There are a few studies of COVID-19 misinformation with WhatsApp users but not in the U.S.. Bowles et al. [12] showed from surveying WhatsApp users in Zimbabwe that information sent from trusted authorities have significant impacts on individuals' knowledge and ultimately crowd behavior. In another study of Indian WhatsApp users, Bapaye and Bapaye [8] conducted a web questionnaire survey to better understand the impact of COVID-related misinformation on WhatsApp users in India. They found that users aged over 65 years and those involved in common labor (e.g., street vendors, housekeepers) were found to be the most vulnerable to false information. The study also found that the presence of an attached link can add significant false credibility to a piece of misinformation. Finally, some work has looked at the efficacy of family chats in disseminating misinformation in Brazil [58] and Kenya [76].

While existing research has been focused on analyzing collected messages to *infer* the effect of misinformation dissemination on WhatsApp users, there have been fewer qualitative studies with WhatsApp users to understand their experiences with misinformation and no studies of misinformation encounters in private WhatsApp chats. Finally, prior studies did not investigate U.S.-based experiences with misinformation on the platform; the third most populous user base of WhatsApp users in the world [70]. Since country context affects misinformation encounters, our work serves to fill these gaps.

3 Methods

3.1 Data Collection Process

To answer our research questions, we conducted semi-structured interviews with 16 WhatsApp users who were university students in the U.S. to better understand their experiences with COVID-19 related misinformation on the platform, particularly in their private chats. Interviews were conducted between October and November 2020 and we stopped recruiting upon reaching data saturation i.e., when we encountered repeating themes without detecting new ones from freshly enrolled participants [63]. Our study was approved by the Institutional Review Boards (IRB) of our two institutions. We designed a demographic survey and interview questions based on prior literature discussed in Section 2. For instance, since prior works had investigated the spread of misinformation in different media formats, we asked about text, image-based, and URLs as sources of misinformation. We also investigated how users perceive current measures for combating misinformation online.

Demographic Survey: Participants were asked to provide their demographic information in a Qualtrics survey prior to participating in their interview. We collected their age range, gender, highest level of education completed, estimated annual income, frequency of WhatsApp usage, and the number of years they had been using WhatsApp. Additionally, this survey was used to collect their consent to audio and video recording during the interview.

Interview Guide: We had three main categories of inquiry for our interviews to answer our research questions:

General usage: We asked questions about frequency and duration of WhatsApp usage to confirm participants' answers on the demographic survey, why they used WhatsApp over other messaging platforms, and what relationships they had with their contacts (friends, family, co-workers, etc.).
Misinformation encounters: We asked participants what concerns if any, they had about false, inaccurate, or misleading information on WhatsApp. We also asked how often they encountered this type of content and what factors they considered when deciding to trust information sent to them via WhatsApp. Specifically, we also asked if this content was text-based, an image, or a URL.
Fact checking strategies and technologies: Finally, we asked participants how they fact-checked information they received in WhatsApp. Additionally, we asked participants about current anti-misinformation tools, shown in Figure 1, such as WhatsApp's limitation on message forwarding, their magnifying glass (search) icon (WhatsApp's web-based fact checker [83]) and Health Alert partnership with the World Health Organization (WHO), along with misinformation labels being used on YouTube and Twitter in 2020 [18, 86].

We piloted our interview guide with lab members who were university students and had never been involved in this project. Based on our pilots, we made minor edits to clarify question phrasing and format. Following the pilots, we continued to the main study with the finalized interview script. Our interview questions are available in our Appendix.

Recruiting: We restricted study participation to those over the age of 18, who used WhatsApp at least multiple times a week, and were living in the U.S.. We sent recruiting notices via a university-based survey research center mailing list to undergraduate and graduate students enrolled at that institu-

Code	Explanation
General	
Chat Content	Participant talked about what they usually talked about in the chats, broadly
Foreign (non-U.S.) vs. domestic communication	Participant uses WhatsApp to communicate with people in or out of the U.S.
Relationship with others in the group (with whom they interact with most often)	Participants identified relationships with others in their group chats
Misinformation Encounters	
Most recent misinformation encounter	Participant recounts most recent misinformation counter (info content, who sent it, their reaction, etc.)
Frequency of encountering misinformation	How often does a participant encounter misinformation? (e.g., once a week, month, year, etc.)
Misinformation indicators	Participant describes factors they consider when deciding to trust (and distrust) information
Design Rec.'s & Fact-Checking Strategies	
Fact-checking strategies	Participant describes how they fact-check information (Google search, literature, consulting others, etc.)
Efficacy of current WhatsApp features that combat misinformation	Participant describes the efficacy of WhatsApp features in fact-checking and limiting the spread of misinformation
Concerns about the trade-off between combating misinformation and privacy/security	Participant raises concerns that fact-checking measures (e.g., information censorship) may undermine the privacy and comfort associated with end-to-end encryption

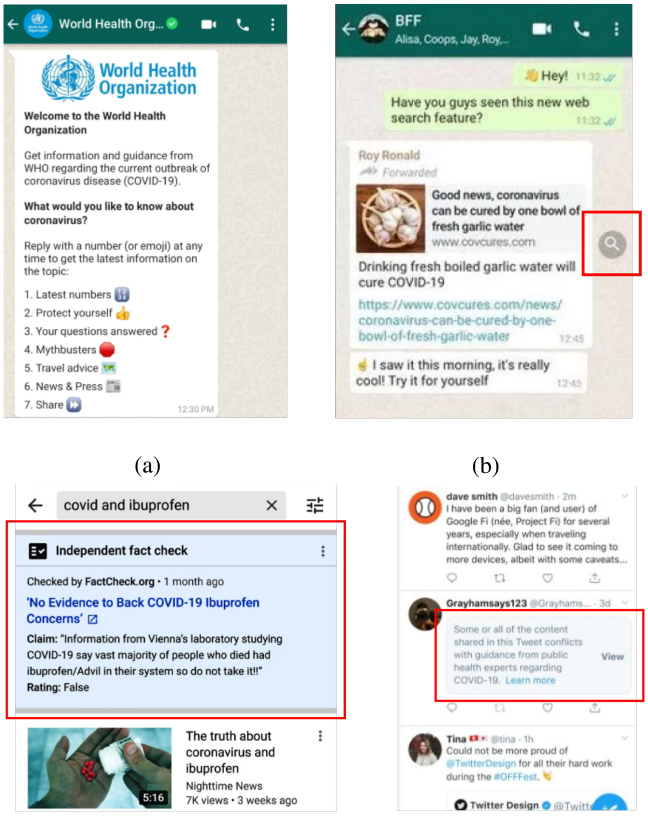
Table 1: A subset of our qualitative code book that is most relevant to the paper with codes and code explanations, organized by topic.

tion, by posting on class Facebook pages at both institutions, and posts on Twitter. The messages did not specifically target users who were aware of misinformation. Note that around 50% of WhatsApp users in the U.S. fall into the typical age range of undergraduate and graduate students in the U.S. [19]. After screening for our filtering criteria, participants completed a demographics survey and were scheduled for interviews. We also used snowball sampling but only recruited one additional participant using this technique. Many participants were in the same geographic region as their university but not necessarily on campus owing to pandemic lockdowns. Each interview lasted 30 minutes to 1 hour and was conducted virtually over Zoom by at least one member of the research team. We interviewed participants in English even though some participants did communicate in other languages. Examining the role of language in the spread of misinformation is beyond the scope of this paper. Note participants were not required to examine their chats during our interviews. Participants were compensated with a $20 Amazon gift card for their time. All interviews were audio-recorded and then transcribed.

Data Analysis: We analyzed our data using deductive coding and thematic analysis [62]. We created a codebook based on our interview guide and our research questions as well as insights from team discussions about emerging points of interest while interviews were being conducted. For instance, we included codes for how participants encounter misinformation and for when they encounter different forms of misinformation such as images or URLs. Our codebook was organized into 3 broad categories, 'General Usage', 'Misinformation Encounters', and 'Design Recommendations and Fact Checking Strategies'. A portion of the codebook is displayed in Table 1, while the full codebook is available in the Appendix. Once we finalized the codebook by consensus in our regular weekly team discussions, each interview transcript was coded by two members of the research team with four coders overall. In total, we ended up with 33 codes and 1183 coded segments across the four coders. Once all the data was coded, we used our weekly research meetings to discuss codes of interest and each of the four coders wrote a detailed summary for a subset of codes resulting in summaries for all of our main codes. These summaries included performing a breakdown of sub-themes within the code and describing each of the sub-themes with representative participant quotes. Each team member then reviewed all the summaries in depth for our thematic analysis [62]. Since we performed coding as input to a thematic analysis, we did not calculate inter-rater reliability as this is not required [44]. However, we still built team consensus through weekly Zoom meetings to decide on the final themes emerging from the data based on the team's reading and discussion of all the thematic summaries.

3.2 Participants

Participants' demographics and WhatsApp usage are summarized in Table 2. Our participants had an almost even gender split with 7/16 participants identifying as male, while 9/16 identified as female. Participants were also younger overall, 14/16 were in the age range of 18-24, while 2/16 were 25-34. Participants were mainly based in the Midwestern U.S. (8/16) and Northeast (6/16) with exceptions of 2/16 based in the West and the Southeast. All participants completed at least high school. The majority (14/16) were students (undergrad-



Welcome to the World Health Organization

Get information and guidance from WHO regarding the current outbreak of coronavirus disease (COVID-19).

What would you like to know about coronavirus?

Reply with a number (or emoji) at any time to get the latest information on the topic:

1. Latest numbers
2. Protect yourself
3. Your questions answered ❓
4. Mythbusters
5. Travel advice
6. News & Press
7. Share

12:30 PM

(a)

Hey! 11:32

Have you guys seen this new web search feature? 11:32

Roy Ronald
Forwarded

Good news, coronavirus can be cured by one bowl of fresh garlic water
www.covcures.com

Drinking fresh boiled garlic water will cure COVID-19

https://www.covcures.com/news/coronavirus-can-be-cured-by-one-bowl-of-fresh-garlic-water 12:45

I saw it this morning, it's really cool! Try it for yourself 12:45

(b)

covid and ibuprofen

Independent fact check

Checked by **FactCheck.org** • 1 month ago

'No Evidence to Back COVID-19 Ibuprofen Concerns'

Claim: "Information from Vienna's laboratory studying COVID-19 say vast majority of people who died had ibuprofen/Advil in their system so do not take it!!"
Rating: False

The truth about coronavirus and ibuprofen
Nighttime News
7K views • 3 weeks ago

(c)

dave smith @davesmith · 2m
I have been a big fan (and user) of Google Fi (née, Project Fi) for several years, especially when traveling internationally. Glad to see it coming to more devices, albeit with some caveats...

Grayhamsays123 @Grayhams... · 3d

Some or all of the content shared in this Tweet conflicts with guidance from public health experts regarding COVID-19. Learn more

Tina @tina · 1h
Could not be more proud of @TwitterDesign for all their hard work during the #OFFFest.

(d)

Figure 1: WhatsApp's WHO Health Alert (a); WhatsApp's search icon fact-checker (b); YouTube's misinformation panel (c); and Twitter's misinformation warning label (d).

uate or graduate) or recent graduates (2/16) including one full-time employee. Seven out of 16 reported annual incomes of <$10,000 per year, 5/16 reported $10,000-$69,999, and 4/16 declined to disclose income. Participants had used WhatsApp for 1-11 years with a median of 7 years.[4] The majority of participants self-reported that they used the app daily.

The number of contacts participants stated they had on WhatsApp varied greatly, ranging from 3 to 1015, with 20-30 being a commonly mentioned range. There was also a significant difference between the total number of contacts a user had and the number of contacts they interacted with on a regular basis. For example, P12 had 1015 total contacts on WhatsApp but was in regular contact with only about 5 of them, while Participants 11 and 15 stated that they had between 100-150 and 20-30 contacts respectively but were in touch regularly with about 20 and 10, respectively. We left the frequency term "regular" up to the definition of the participant. We also asked participants to provide us with the number of people in their chat groups (if they were comfortable doing so) and to estimate the average size of the groups they were in otherwise. Most of the group chats were between 3 and 10 people, which were commonly mentioned sizes for private group chats consisting of family members.

4 Findings

Our analysis of the interviews yielded three main findings: how users are currently using WhatsApp (including their concerns about misinformation on the platform, how often they encountered it, and how it can spread); what misinformation indicators users look for and how they respond to misinformation on the platform; and finally, how users would like the platform to respond to misinformation.

4.1 Misinformation Perceptions And Responses

In research question one, we asked how university students currently perceive and encounter misinformation on WhatsApp. Our participants mostly used WhatsApp to communicate with others abroad, were concerned about frequently encountered misinformation on the platform, and noted that misinformation senders were often well-intentioned relatives.

4.1.1 WhatsApp Usage And Misinformation Encounters

All of our participants stated that they used WhatsApp to communicate with families and/or friends outside of the U.S. as WhatsApp was convenient to stay in touch with people abroad. This is hardly surprising as a significant number of WhatsApp users in the U.S. have non-U.S. family members [42]. Only two of our participants (P6 and P11) used WhatsApp to communicate domestically. Participants told us that they used WhatsApp primarily to share happenings in everyday life with family and friends. Interactions with family groups tended to be more regular than communications with friends.

Although participants praised the pros of WhatsApp, they also expressed concerns towards misinformation and nonsensical content circulating on WhatsApp—the main concern expressed was misleading information on COVID-19 cases and cures. For instance, at least 3/16 participants talked about how easy it is for misleading content to spread on WhatsApp since it was so easy to forward links in general. For example, P6 said that it is also "*almost too easy*" to select many people or groups to send a message to upon tapping the forward button, and that misinformation from families can have a layer of intimacy attached to it that makes it especially harmful:

> *"I know [many] have their families in WhatsApp, and people tend to trust things that come from people close to you. So, I feel like it adds almost a level of genuineness to this misinformation, and then it causes people to panic, which I think is the biggest con [of using WhatsApp]." — P6*

[4] At the time of this study, WhatsApp was more than 11 years old [81].

#	Gender	Age Range	Region	Occupation	Frequency of Use (/week)	Duration of Use (years)
P1	Female	18 – 24	Midwest	Student	Daily	7
P2	Female	18 – 24	Midwest	Student	2 – 3	7
P3	Female	18 – 24	Northeast	Student	2 – 3	1
P4	Male	18 – 24	Midwest	Student	Daily	11
P5	Male	18 – 24	Midwest	Student	Daily	8
P6	Male	25 – 34	Northeast	Student	Daily	8
P7	Female	25 – 34	Midwest	Developer	Daily	8
P8	Male	18 – 24	Southeast	Student Researcher	Daily	6
P9	Female	18 – 24	Midwest	Student	Daily	3
P10	Male	18 – 24	Midwest	Student	Daily	2
P11	Male	18 – 24	Northeast	Student	Daily	8
P12	Male	18 – 24	Northeast	Student	Daily	6
P13	Female	18 – 24	Midwest	Student	4 – 6	4
P14	Female	18 – 24	Midwest	Student	2 – 3	6
P15	Female	18 – 24	Northeast	Student	2 – 3	3
P16	Female	18 – 24	West	Student	Daily	7

Table 2: Participant demographics (gender, age, region, occupation, frequency of WhatsApp use, and duration of use).

Another participant, P5, described how they have gotten so used to skeptical content on the platform that they treat it as a medium for conversation rather than relying on it for news; they also expressed the caveat that older generations trust it more. The majority of the participants (14/16) received misinformation almost every other day or multiple times a week. These participants recognized that false or misleading messages were most frequently seen in group chats possibly because *"people like to keep busy with sending messages."* These false or misleading messages most commonly came in the form of conspiracy theories or potential cures for diseases (particularly when COVID had first entered the U.S.). For instance, P13 recalled an instance of having received a post about how *"juice made out of coriander stems and raw egg and tomato theory helps cure cancer*" in spring of 2020. The 2/16 participants who never encountered misinformation on WhatsApp attributed the lack of encounters to communicating primarily with friends (i.e., in their age range) who they know well—as opposed to family members. We also asked participants about whether or not they forwarded content to their contacts on WhatsApp to better understand how misinformation or any information may travel on the platform. Many participants (8/16) claimed to have either "*rarely*" or "*never*" forwarded any links or posts that they received on one chat to another chat. For instance, participant (P9) shared "*No, I do not because, as I mentioned, I'm guarded when I look at some of these headlines. I feel like we're living in such a weird time.*" The 8/16 participants who did share or forward links told us that they first fact-checked the links and then sent the information only if it seemed reliable to them.

4.1.2 Misinformation Senders

We asked participants about who or what entity was sending them misinformation on WhatsApp. The 14/16 participants who had a high frequency of encountering misinformation (approximately every other day or multiple times a week), revealed that the senders were typically close family members. These family members sent (mis)information in a range of formats (from *"copy pastas*"—long, often joking texts distributed through copy and paste—to texts, images and links). Our participants felt that this information ultimately did not harm them because they were either cognizant of these groundless claims or the information itself did not pose a severe threat to anyone who believed it. In the words of P3:

> *"The sender for me was just my mom, and I did speak to her about it, and she was definitely of a different mindset. She was more of the mindset that we should do whatever we can even if it's not true, even if it's just helping your immune system at this point, we'll do anything. So, I wouldn't say she necessarily believed that it makes you immune to COVID, or protects you or anything, but she also didn't consider it misinformation. She was like "As long as it's helping everyone." She also sent it to people. . . I mean, it's up to you to do whatever you want with it." - P3*

Participants also expressed that these family members were often sending messages without malicious intent of sharing information that could prove dangerous. Another participant (P10), reflecting this sentiment, perceived that:

> *"[her mom and aunts] find it very easy to essentially forward a message from another group chat*

> *to another, essentially spamming the group chat with all sorts of massive, long text messages about something, or a web link that is pretty much misinformation." - P10*

Contrary to having malicious intent, our participants also described how, oftentimes, their family members sent misinformation with the intention of keeping others safe and informed in the midst of a pandemic. For example, P10 also described how half of her family believed "*that we should rinse our noses with saline solution to prevent COVID*" and when asked if she followed this protocol, she would merely respond by saying yes so as to avoid getting into a lengthy argument of whether and why this approach to combating the virus is ineffective.

4.2 Misinformation Indicators and Responses

In our second research question, we asked how users identify whether content is misinformation on the platform and how they respond to misleading content. Participants told us they had four main indicators that a message was misinformation and had developed strategies for fact-checking content. In response to misinformation, not everyone was comfortable with confronting senders, often owing to family dynamics.

4.2.1 Indicators Of Misleading Content

Generally, participants told us about four main indicators that they relied on to decide whether to trust information sent to them via WhatsApp: 1) the credibility of the information source, 2) their relationship with the misinformation sender, 3) the format and framing of the message, and 4) personal politics and values. Many of these strategies, aside from relationship with the sender, echo indicators developed by Jahanbakhsh et al. [32] on reasons people believe or disbelieve claims, as well as textual misinformation indicators for automated detection specified by Resende et al. [57]. These strategies also echo findings on studies of other social media platform users such as Facebook [22, 26, 47], i.e., using the source of a news article to evaluate its credibility.

Source Credibility and Name Recognition. The majority of participants paid attention to the source's credibility when deciding to trust information sent to them (15/16). Participants focused on the reputability of the organization when analyzing information, most often news media content. Established media and news corporations carried greater credibility and legitimacy compared to smaller, more obscure media outlets; e.g., participants mentioned The New York Times and MSNBC. Participants generally expected the source to be linked to an established news platform as opposed to a random individual's social media account. Additionally, participants considered government organizations and links that forwarded to .org and .gov, e.g., www.cdc.gov, as reliable.

Relationship with Sender. Complementary to Geeng et al.'s finding that Facebook and Twitter users may trust certain poster's content because they trust the individual [26], we found that the opposite can be true as well; participants may inherently mistrust content because they have deemed the sender to be unreliable and untrustworthy.

Since participants primarily used WhatsApp to communicate with friends and family, they told us they measured the trustworthiness of information based on their relationship and perception of the sender. If a sender was known to consistently share misleading information, participants were more likely to be skeptical of them. This theme was most prevalent when participants described their relationship with older relatives; 9/16 expressed concern that their older contacts were unable to distinguish between credible and untrustworthy news content and were less prone to fact-checking before sharing on WhatsApp. Over time, P2 felt increasingly suspicious when receiving messages from their grandparents and older relatives in large family group chats:

> *"Just because they are not as able to filter out fake news from real news. I mean, obviously it's presented in a more and more realistic way every single day and they just lap it up and believe in it, and also, they are not as tech savvy to be able to go and Google immediately and do a quick check on what's actually happening" — P2*

Participants described how these contacts would frequently spam family group chats with information they received in other group chats and channels. Five out of 16 participants described ignoring messages from particular senders since they automatically assumed false or misleading content. However, there were a few exceptions where participants trusted their contacts when sharing information on unfamiliar topics. For example, in the midst of school and university closings in response to the early COVID-19 outbreak, P15, a graduate student, said she was bombarded with news stories that contradicted each other. This participant reached out to her sister who told her to expect her school to cancel all in-person activities. Because P15 had a close relationship with her sister, she trusted her sources.

Format and Framing. Six of the 16 participants reported distrusting and avoiding messages that: urged users to spam forwards, shared without context, were overly sensational and attention-seeking, had inflammatory language, and were opinion-based. Three out of 16 participants expressed mistrust of forwarded messages because these messages often followed a template that explicitly asked users to forward the message to their contacts. Further, participants believed if someone did not dedicate time to writing their own messages, they probably did not verify it either. Participants also took the visual layout and format of a message into account as well; two participants avoided messages that displayed excessive use of colors, advertisements, capitalized and bold texts, emoticons, and other eye-catching designs apart from the text itself. Participants also told us they were wary of poorly spliced pictures that may have been edited beforehand or messages framed with inflammatory, opinionated content that were seen as biased and misleading (2/16). In the case of COVID-19 news, these participants trusted sources that presented numerical data (e.g., number of cases, growth rate) in a neutral tone without underlying agendas.

Political ideology. A few participants (4/16) expressed political ideology as an important factor when deciding to trust information. They said they were less likely to trust content, as credible as it may be, from news organizations or their personal contacts with conspicuous political views out of concern of an underlying political agenda. For instance, P9 expressed having conservative political values and criticized left-leaning news sources sent from contacts with opposing political ideologies because they automatically considered them biased and misleading. Likewise, P11, a self-described liberal, disregarded any news articles sent from conservative family members.

4.2.2 Fact-Checking Using Google And Intuition.

Thirteen out of 16 participants were asked about fact-checking strategies, and two main approaches were found as participants' primary fact-checking approaches: 1) searching on Google and 2) relying on personal judgment. Apart from these, reading scientific papers was mentioned once by a graduate student (P11) and directly asking other contacts such as friends by one other participant (P6). It is worth noting that, in reality, these strategies are not mutually exclusive and are often employed together by an individual in a single fact-checking attempt.

Google. 12/14 participants told us their most common way to fact-check information sent to them on WhatsApp was to search on Google to verify its accuracy. When a source's reliability was unknown, P15 stated they usually "*click on the links, maybe read some other articles that have been published by the same website or author and see if those are accurate*". If participants found multiple sources corroborating each other, they felt this was an extra piece of evidence that the information was accurate, therefore trustworthy. Participants told us that their process of verifying the information with other sources, especially those considered authoritative, was not exclusive to Google. They checked the information from any source that they usually consulted for information and trusted.

Prior Knowledge. Eight out of 16 participants relied on their intuition, prior knowledge, and understanding of current affairs to determine whether or not a message, image, text, or URL was intentionally misleading or false. This finding echoes that of Flintham et al. [22], for Facebook users who looked for 'fake news' in an experiment on fake news articles only and sometimes relied on their own judgement for determining veracity. In our study, which occurred in the first year of the COVID-19 pandemic, most participants expressed prior knowledge of COVID-19 cases, precautions, and myths that informed them outside of their WhatsApp channels. For example, myths about COVID-19, such as gargling warm salt water or drinking lemon juice twice a day, sounded completely outlandish to some participants given their understanding of the properties of the virus and the vaccine. In another related example, P10 described a misinformation encounter where their aunt claimed eating ice cream and other cold foods increased the chances of contracting the coronavirus:

> *"If I had to think about basic biology, it's pretty hard to link ice cream to a virus that caused a global pandemic, I would say. I'd say, yes, maybe if you eat ice cream a lot and don't dress up in cold months, your immune system may be more vulnerable to the flu, to the virus. But it wouldn't be a direct cause of COVID" — P10*

4.2.3 Dealing With Misinformation Senders

Out of 15 participants who allegedly encountered misinformation via WhatsApp, 9 people mentioned past experiences of confronting senders of misleading information, 8 people mentioned scenarios where they were passive and didn't challenge the senders—even when they recognized there were something incorrect with the content shared, and 2 others confessed they didn't always stick to one strategy.

Actively Confronting Misinformation Senders. When encountering misinformation, "active" participants confronted the sender, especially if they were on close terms with them. However, most of them recognized that "*there is no point*" in repeatedly resisting and reminding the sender to check the sources of any information they forward, prior to sharing, especially when the sender continues not to do so. In one canonical example, P3 actively confronted their mother by asking a question along the lines of *"Do you also believe this? Do you think it's believable?"* The participant also explained that they were able to confront the sender (in this case their mother) since the participant was a) close with the sender and b) they knew that the sender had no malicious interest in sending incorrect information. Other "active" participants, who fact-checked a topic by doing further research, shared that whenever they received any information that they had not yet encountered, they ventured to ask the sender questions like "*where did you find this?*". In one example, P1's mother sent her sensational and misleading information on COVID cases in the U.S.. Although P1 personally thought that the U.S. could do better in curtailing the virus, she recognized that her mother's sources made the problem worse than it was. Recognizing that she was simply worried and did not purposely share misinformation, P1 confronted her mother to comfort her:

> *"Yes, we did talk about this quite often during the*

> *video chatting. I would just try to assure her, "Oh, Mom. This is okay," and regardless how the numbers surge in America, like myself, at least I can protect myself. I just wear masks and I do hand sanitizing very often, so I'm trying to point out to her, "Mom, this is misinformation. America is actually doing fine." Well, it's not. So, yeah, I don't counter the source directly, but I am trying to comfort her on speaking for my personal level." – P1*

Passively Ignoring Misinformation Senders. While these "active" participants did not let these qualms prevent their confronting of senders, "passive" participants acknowledged that they would simply ignore anything shared via WhatsApp based on the contents and sender of the post (e.g., if the content concerned the 2020 Black Lives Matter protests or COVID-19). At least 2 out of the 6 passive participants expressed explicitly that they did not want to upset any family relations due to a *"trivial"* post shared on social media. Other participants echoed this sentiment and told us they often reacted passively about misinformation, not taking the time to correct others' misaligned opinions or views as it would lead to an "*hour long argument*" which the participants did not want to face. In another anecdote, P2 recalled having received information from her family members regarding unfounded steps of precaution to take against COVID involving gargling with "*warm saltwater every time*" they came back into their home from being outside to "*kill off all COVID particles and be safe.*" This participant did not correct their family members as they did not want to cause any unfriendliness for a harmless piece of information:

> *"I'm not interested in trying to correct people because it's just not going to work, they're going to believe what they want to believe. I had a phase a couple of years ago where I was trying to correct people and I was like, it's not going to happen, it's not going to work. So now I'm just like, 'Sure, you do you and I'm just going to ignore."' – P2*

In another representative example, P8, reported that it was easier to delete group chats which they had flagged as one of main mediums of misinformation without reading any content sent. P8 accepted that *"There was just a point where there was so much going around it was easier to just, honestly, stop reading things."* To summarize, participants often did not want to strain family relationships by correcting misinformation, especially given that, in many cases, they perceived the misinformation to be harmless.

4.3 Views on Existing Mechanisms To Combat Misinformation on WhatsApp

To answer research question three, we asked how aware and confident participants were of current features to combat misinformation on WhatsApp and their opinions on how to improve how the platform handles misinformation, particularly around COVID-19 as shown in Fig. 1. In general, participants showed little to no awareness towards the features probed and expressed varying opinions on efficacy of these features and concerns around the privacy dilemma of combating misinformation in the context of end-to-end encryption.

Of all the existing features shown or discussed with all participants (WhatsApp forwarding limits, WhatsApp search icon, and the WHO health alert), on average only about 4 participants had heard of at least one or more of these features. Generally, participants mentioned that the forwarding limit could be circumvented if a sender manually copied and pasted it or by sending the message one at a time or via another platform. Participants also thought the search icon could link to multiple search engines rather than one and felt the WHO alert did not look professional owing to the use of emojis.

4.3.1 Privacy and Security Concerns

Not only were participants unaware of existing anti-misinformation measures, they also voiced concerns on whether or not WhatsApp should even be responsible for designing preventative measures against misinformation.

Content Moderation Concerns. At least 6/16 participants believed that WhatsApp, as a platform, should not be accountable for curbing any misinformation, arguing that it is up to the user's discretion whether or not they believe what they see. Even if the content is explicitly false, they felt that users are entitled to share anything they want and believe to be true. On the other hand, participants agreed that WhatsApp definitely bears a responsibility in fact-checking and regulating any misleading content, rather than burdening the user to determine what is trustworthy.

Other participants expressed major concerns about the trade-off between users' privacy and WhatsApp's efficacy against misinformation (3/16). They felt these features infringed upon users' privacy and therefore preferred if WhatsApp did not explicitly flag or censor misinformation. Should WhatsApp ever flag or censor direct messages, it would need to clarify any privacy-preserving techniques and the methods used to identify any inflammatory or misleading content.

Misinformation Warnings And Labels. When asked to suggest design recommendations to limit the spread of misinformation, only 5/16 participants thought that WhatsApp should adopt the misinformation warning labels similar to YouTube's and Twitter's warnings [18, 86]. They liked the idea of warning users not to trust certain sources while still giving them the option to share. As P13 said, "*they should be allowed to view it because of free speech, but they should be aware that it is incorrect, it's misinformation.*". An alternative suggestion was for WhatsApp to record known misinformation sources such as websites (4/16) or to generate a credibility rating for websites for when senders share links (2/16).

5 Discussion and Design Suggestions

Our study suggests that WhatsApp is uniquely situated in the misinformation space based on the following three key findings:

- **F1**: Our participants' group-based WhatsApp communications with close family and friends make it especially effective in disseminating misinformation out of good intention. Previous studies observed the efficacy of WhatsApp as a misinformation pipeline in large public chats [31,41,46,49], but our study suggests this may also be the case in private chats. Future studies are needed to confirm if it is mainly older adults spreading content.

- **F2**: The peer-to-peer nature of communication on WhatsApp adds intimacy and complicates users' ability and/or willingness to deal with misinformation they encounter. Because we focused on gathering deep user experiences in private chats over collecting data using automated methods as in prior studies [25,45,57], we were able to surface significant social power dynamics within chats that pose challenges to countering misinformation.

- **F3**: Participants were unaware of current mechanisms on WhatsApp to combat misinformation. Moreover, privacy and information accuracy, both desirable in communication apps, can be seen as conflicting traits on WhatsApp. Such a tradeoff has been a common technical assumption known to experts in the field [43], but our study revealed that everyday users are also well aware of this trade-off.

We think it is particularly important to engage with **F3** when addressing misinformation in end-to-end encrypted environments. While some participants told us they would appreciate more effort on WhatsApp's part to flag misinformation, they also acknowledged that WhatsApp's inability to read messages will hinder its ability to do so. However, no participant mentioned that encryption should be sacrificed to offer more robust fact-checking services, implying that they still hold privacy on the platform in high regard. This tension offers rich avenues for future work.

In addition to privacy, dealing with misinformation in private chats is complicated by social relations. We found that the more personal nature of communication on WhatsApp integrated social dynamics that discouraged a user from actively confronting misinformation senders. Our observed social dynamics include cultural emphases on respect and deference to elders: many of our participants feared correcting older family members' misinformation out of concern for coming across as rude or disrespectful, despite having a justifiable and legitimate reason. Therefore, younger users, who our participants claim to be more adept at identifying misinformation, may not be able to signal the misleading nature of a piece of information to others if it is sent by older family members or relatives. Further, many participants recognized that misinformation often resulted from well-intentioned family members who sent it out of care for others (e.g., bogus COVID-19 cures), supporting preliminary research suggesting that information dissemination on WhatsApp follow familial, communal, and ideological ties [7]. This is worthy of further study in the U.S. as it may be of particular relevance to a rising body of work around digital communication and misinformation within American immigrant diaspora communities [68,78].

These findings point to a need for alternate approaches to combating misinformation in end-to-end encrypted, private group chats, as conventional moderation techniques often rely on examining content and do not take into consideration sociocultural dynamics between group chat members. For example, educational campaigns around misinformation may include tips and suggestions for dealing with relatives but ground this in terms of caring about others.

5.1 Design Suggestions

Our participants were for the most part unaware of anti-misinformation features on WhatsApp, suggesting that even when a platform is actively trying to combat misleading content, users may not know about these measures. Assuming a platform can overcome the hurdle of raising user awareness of new anti-misinformation features, based on the insights above, we propose the following design approaches to improve the ways users can deal with misinformation on end-to-end encrypted platforms. These features may be useful to users within our study demographic, but generalizations to a broader user base cannot be made without additional studies.

5.1.1 Empowering the user to better fact-check or flag misinformation for themselves.

WhatsApp cannot analyze content to identify misinformation due to the platform's encryption policies. Another platform-controlled measure, forwarding limits, has been seen as ineffective by participants in our study as well as previous work [46]. Based on our findings, we suggest designing to *empower the user* with tools to combat misinformation. For misinformation senders, we suggest reminding users of the value of fact-checking before forwarding content. For misinformation receivers, designs should: 1) respect the user's ability to classify misinformation for themselves, and 2) make it easier for the user to organize and track their misinformation encounters so they can later fact-check and better learn from them. This can be translated into features for both the information sender and receiver.

- **Sender**: By adding friction using a popup dialogue box that asks the user whether they have fully read the contents of a link, users can be prompted to reflect on information they are sharing before forwarding content. This kind of friction is already being deployed by other platforms to reduce sharing without context [24] and is shown to be effective in obstructing access to disinformation [33]. However, the friction should not be too high, as it can then be seen as censorship, [59].

- **Receiver**: An option to mark a message as dubious and decrease its visibility in their chat screen may help users mitigate the sight of misleading content. This can protect the user as previous work in psychology indicate that repetition of a message can increase believability in it despite one's initial judgements [20, 21, 77] from believing deal with the constant flow of misinformation. Note that this feature is distinct from WhatsApp's current option to delete a message, which can result in disparate versions of the same chat across different users. [79].
- **Receiver**: To help users track and fact-check messages, users may store messages that have been flagged as dubious in a "quarantine" bin for later inspection.The bin can be equipped with tools to help users surface trends, such as common language or links, across dubious messages. Users can then use these trends better identify misinformation in future messages.

5.1.2 Helping users deal with misinformation in ways that mitigate power dynamics in groups.

Our findings suggest social dynamics in family group chats can make it difficult for users to confront and correct misinformation senders. We propose the following features to allow users to subtly alert others about potential misinformation.

- **Selectively applying the fact checker icon to messages**: We can let users anonymously apply WhatsApp's fact checker[5] to particular messages for everyone in the chat to see and use. This offers resources to group members without accusing anyone of sending misinformation.
- **Anonymous suggestions of alternative resources**: One suggestion is to allow users to anonymously suggest a link to an alternative information resource to the sender. Once the resource is suggested, the sender can receive a notification with the anonymous suggestion and choose whether to accept it. If accepted, the link can be sent into the group as a reply to the original message to update others and gently nudge the group towards discussion.

6 Limitations and Future Work

Our study sample was limited to 16 university students and recent graduates who were mostly in a younger age bracket of 18-35 years. By its nature, our qualitative study is not intended to be generalizable [62, 63]. Future work could expand our study to a broader sample of young users who are not students or to a larger sample of more age-diverse U.S. based participants across the country. Also, while we asked participants about misinformation around topics such as Black Lives Matter protests and U.S. elections, we did not collect sufficient data to report on it. Future work could thus investigate topics beyond COVID-19. Additionally, even though our participants were based in the U.S., we observed that most communication on the app was international. Studies that specifically investigate misinformation within domestic interactions on WhatsApp may also complement our work since the language of communication may affect the perceptions of misinformation. Studying WhatsApp users in other countries would also expand on our study. Finally, future studies could implement and test our design recommendations or study other end-to-end encrypted chat-based platforms, such as Telegram [73], Signal [65], and iMessage [2].

7 Conclusions

We interviewed 16 U.S.-based university students and a recent graduate about their experiences with misinformation related to COVID-19 in private WhatsApp group chats. We were interested in filling in two gaps in previous literature: the lack of qualitative user interviews to understand younger adults' misinformation experiences on end-to-end encrypted messaging platforms such as WhatsApp, and the lack of studies on how WhatsApp is used in the U.S. Our findings suggest that there is a need to differentiate the nature of misinformation on WhatsApp compared to other popular American social media apps such as Twitter and Facebook. Namely, WhatsApp's popularity as an international communication tool used with close family or friends can unknowingly turn good intentions into misinformation-sharing frenzies and hinder the ability of those who identify misinformation to notify others about it. Additionally, WhatsApp's staunch commitment to end-to-end encryption can present limitations to the techniques the platform is able to deploy to combat misinformation. Our findings offer implications for design approaches to both mitigate the sharing of misinformation and improve experiences of users who receive misinformation. These findings and suggestions may help WhatsApp users outside the U.S.—and even users on similar platforms—handle similar issues and spark new discussions around information moderation with privacy-preserving techniques more broadly.

Acknowledgments

We thank our participants. This work was partially supported by the Princeton Council for Science and Technology and a Facebook 'Secure The Internet' award.

[5] WhatsApp has already rolled out to some users its own web-based fact-checker [83]. However, since the platform cannot read message contents, it applies the fact checker to all links, which may not always be desirable.

References

[1] Zara Abrams. Controlling the spread of misinformation. *American Psychological Association*, 52:44, 03 2021.

[2] Apple Inc. Use imessage apps on your iphone, ipad, and ipod touch. https://support.apple.com/en-us/HT206906. Accessed: 2021-07-09.

[3] Ahmer Arif, Leo Graiden Stewart, and Kate Starbird. Acting the part: Examining information operations within #blacklivesmatter discourse. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW), November 2018.

[4] Brooke Auxier and Monica Anderson. Social media use in 2021. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/, 04 2021. Accessed: 2021-06-28.

[5] Daniel Avelar. Whatsapp fake news during brazil election 'favoured bolsonaro'. https://bit.ly/3tqVz61, 10 2019. Accessed: 2021-10-21.

[6] Ahmed Balami and HadizaUmar Meleh. Misinformation on salt water use among nigerians during 2014 ebola outbreak and the role of social media. *Asian Pacific Journal of Tropical Medicine*, 12:175, 01 2019.

[7] Shakuntala Banaji, Ram Bhat, Anushi Agarwal, Nihal Passanha, and Mukti Sadhana Pravin. WhatsApp Vigilantes: An exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India. page 62.

[8] Jay Amol Bapaye and Harsh Amol Bapaye. Demographic factors influencing the impact of coronavirus-related misinformation on whatsapp: Cross-sectional questionnaire study. *JMIR Public Health Surveill*, 7(1):e19858, 01 2021.

[9] Zapan Barua, Sajib Barua, Najma Kabir, and Mingze Li. Effects of misinformation on covid-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. *Progress in Disaster Science*, 8:100119, 07 2020.

[10] Shashank Bengali. How whatsapp is battling misinformation in india, where 'fake news is part of our culture'. https://www.latimes.com/world/la-fg-india-whatsapp-2019-story.html, 02 2019. Accessed: 2021-10-21.

[11] Porismita Borah, Bimbisar Irom, and Ying Chia Hsu. 'it infuriates me': examining young adults' reactions to and recommendations to fight misinformation about covid-19. *Journal of Youth Studies*, pages 1–21, 2021.

[12] Jeremy Bowles, Horacio Larreguy, and Shelley Liu. Countering misinformation via whatsapp: Preliminary evidence from the covid-19 pandemic in zimbabwe. *PLOS ONE*, 15:e0240005, 10 2020.

[13] Anna Brosius, Jakob Ohme, and Claes H de Vreese. Generational gaps in media trust and its antecedents in europe. *The International Journal of Press/Politics*, page 19401612211039440, 2021.

[14] Carole Cadwalladr. The great British Brexit robbery: how our democracy was hijacked. https://bit.ly/3MCpdvE, 2017. Accessed: 2021-06-08.

[15] A. Chadwick and Cristian Vaccari. News sharing on uk social media: misinformation, disinformation, and correction. 2019.

[16] Adélie Chevée. Mutual aid in north london during the covid-19 pandemic. *Social Movement Studies*, pages 1–7, 2021.

[17] Mitchell Clark. Facebook wants to make sure you've read the article you're about to share. https://www.theverge.com/2021/5/10/22429174/facebook-article-popup-read-misinformation, 2021. Accessed: 2021-06-07.

[18] COVID-19 misleading information policy. Covid-19 medical misinformation policy. https://help.twitter.com/en/rules-and-policies/medical-misinformation-policy. Accessed: 2021-07-02.

[19] Statista Research Department. Whatsapp usage penetration in the united states 2020, by age group. https://www.statista.com/statistics/814649/whatsapp-users-in-the-united-states-by-age/, 10 2021. Accessed: 2021-10-26.

[20] Lisa Fazio, Nadia Brashier, B Payne, and Elizabeth Marsh. Knowledge does not protect against illusory truth. *Journal of experimental psychology. General*, 144, 08 2015.

[21] Lisa Fazio and Gordon Pennycook. Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bulletin & Review*, 26, 08 2019.

[22] Martin Flintham, Christian Karner, Khaled Bachour, Helen Creswick, Neha Gupta, and Stuart Moran. *Falling for Fake News: Investigating the Consumption of News via Social Media*, page 1–10. Association for Computing Machinery, New York, NY, USA, 2018.

[23] Fiona Gabbert, Amina Memon, Kevin Allan, and Daniel B Wright. Say it to my face: Examining the

effects of socially encountered misinformation. *Legal and Criminological Psychology*, 9(2):215–227, 2004.

[24] Vijaya Gadde and Kayvon Beykpour. Additional steps we're taking ahead of the 2020 us election. https://blog.twitter.com/en_us/topics/company/2020/2020-election-changes.html, 2021. Accessed: 2021-06-07.

[25] Kiran Garimella and Dean Eckles. Whatsapp and nigeria's 2019 elections: Mobilising the people, protecting the vote. *Harvard Kennedy School (HKS) Misinformation Review*, 07 2020.

[26] Christine Geeng, Savanna Yee, and Franziska Roesner. 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*, CHI '20, page 1–14, New York, NY, USA, 2020. Association for Computing Machinery.

[27] Amira Ghenai and Yelena Mejova. Fake cures: User-centric modeling of health misinformation in social media. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW), November 2018.

[28] Nathaniel Gleicher. Removing Coordinated Inauthentic Behavior and Spam From India and Pakistan. https://about.fb.com/news/2019/04/cib-and-spam-from-india-pakistan/, 2019. Accessed: 2021-06-06.

[29] Rachel Greenspan. Whatsapp fights fake news with message forwarding restrictions. https://time.com/5508630/whatsapp-message-restrictions/, 01 2019. Accessed: 2021-07=07.

[30] Andrew Guess, Jonathan Nagler, and Joshua Tucker. Less than you think: Prevalence and predictors of fake news dissemination on facebook. *Science advances*, 5(1):eaau4586, 2019.

[31] Jamie Hitchen, Jonathan Fisher, Nic Cheeseman, and Idayat Hassan. Whatsapp and nigeria's 2019 elections: Mobilising the people, protecting the vote. 07 2019.

[32] Farnaz Jahanbakhsh, Amy X. Zhang, Adam J. Berinsky, Gordon Pennycook, David G. Rand, and David R. Karger. Exploring lightweight interventions at posting time to reduce the sharing of misinformation on social media. *Proc. ACM Hum.-Comput. Interact.*, 5(CSCW1), April 2021.

[33] Ben Kaiser, Jerry Wei, Elena Lucherini, Kevin Lee, J Nathan Matias, and Jonathan Mayer. Adapting security warnings to counter online disinformation. In *30th {USENIX} Security Symposium ({USENIX} Security 21)*, 2021.

[34] Masato Kajimoto, Yenni Kwok, Yvonne Chua, and Ma Labiste. Information disorder in asia and the pacific: Overview of misinformation ecosystem in australia, india, indonesia, japan, the philippines, singapore, south korea, taiwan, and vietnam. *SSRN Electronic Journal*, 03 2018.

[35] Khalid Khaja, Alwaleed Alkhaja, and Reginald Sequeira. Drug information, misinformation, and disinformation on social media: a content analysis study. *Journal of Public Health Policy*, 39, 08 2018.

[36] Ramez Kouzy, Joseph Abi Jaoude, Afif Kraitem, Molly El Alam, Basil Karam, Elio Adib, Jabra Zarka, Cindy Traboulsi, Elie Akl, and Khalil Baddour. Coronavirus goes viral: Quantifying the covid-19 misinformation epidemic on twitter. *Cureus*, 12, 03 2020.

[37] David Lazer, Matthew Baum, Nir Grinberg, Lisa Friedland, Kenneth Joseph, Will Hobbs, and Carolina Mattsson. Combating fake news: An agenda for research and action. https://shorensteincenter.org/combating-fake-news-agenda-for-research/, 05 2017. Accessed: 2021-06-22.

[38] David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. The science of fake news. *Science*, 359(6380):1094–1096, 2018.

[39] Stephan Lewandowsky, Ullrich Ecker, Colleen Seifert, Norbert Schwarz, and John Cook. Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13:106–131, 12 2012.

[40] Eugène Loos and Jordy Nijenhuis. Consuming fake news: A matter of age? the perception of political fake news stories in facebook ads. In *International Conference on Human-Computer Interaction*, pages 69–88. Springer, 2020.

[41] Caio Machado, Beatriz Kira, Vidya Narayanan, Bence Kollanyi, and Philip Howard. A study of misinformation in whatsapp groups with a focus on the brazilian presidential elections. In *Companion Proceedings of The 2019 World Wide Web Conference*, WWW '19, page 1013–1019, New York, NY, USA, 2019. Association for Computing Machinery.

[42] Farhad Manjoo. For millions of immigrants, a common language: Whatsapp. https://nyti.ms/39fwjZv, 12 2016. Accessed: 2022-04-24.

[43] Jonathan Mayer. Content moderation for end-to-end encrypted messaging. https://www.cs.princeton.edu/~jrmayer/papers/Content_Moderation_for_End-to-End_Encrypted_Messaging.pdf, 10 2019. Accessed: 2021-10-22.

[44] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for cscw and hci practice. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), November 2019.

[45] Philipe Melo, Johnnatan Messias, Gustavo Resende, Kiran Garimella, Jussara Almeida, and Fabrício Benevenuto. Whatsapp monitor: A fact-checking system for whatsapp. *Proceedings of the International AAAI Conference on Web and Social Media*, 13(01):676–677, 07 2019.

[46] Philipe Melo, Carolina Vieira, Kiran Garimella, Pedro Vaz de Melo, and Fabrício Benevenuto. *Can WhatsApp Counter Misinformation by Limiting Message Forwarding?*, pages 372–384. 01 2020.

[47] Miriam J Metzger, Andrew J Flanagin, and Ryan B Medders. Social and heuristic approaches to credibility evaluation online. *Journal of communication*, 60(3):413–439, 2010.

[48] Ryan C Moore and Jeffrey T Hancock. Older adults, social technologies, and the coronavirus pandemic: Challenges, strengths, and strategies for support. *Social Media+ Society*, 6(3):2056305120948162, 2020.

[49] Vidya Narayanan, Bence Kollanyi, Ruchi Hajela, Ankita Barthwal, Nahema Marchal, and Philip N. Howard. News and information over facebook and whatsapp during the indian election campaign. *Project on Computational Propaganda*, 02 2019.

[50] Khulekani Ndlovu, Dani Madrid-Morales, Herman Wasserman, Melissa Tully, and Emeka Umejei. Motivations for sharing misinformation: A comparative study in six sub-saharan african countries. *International Journal of Communication*, 15:1200–1219, 02 2021.

[51] Office of the Director of National Intelligence. Assessing Russian activities and intentions in recent US elections. National Intelligence Council. https://www.dni.gov/files/documents/ICA_2017_01.pdf, 2017. Accessed: 2021-06-08.

[52] Jonathan Corpus Ong and Jason Vincent A Cabañes. Architects of Networked Disinformation. The Newton Tech4Dev Network. https://bit.ly/3aIvoRu, 2018. Accessed: 2021-06-08.

[53] Bella Palomo and Jon Sedano. Whatsapp as a verification tool for fake news. the case of 'b de bulo'. *Revista Latina de Comunicacion Social*, 73:1384, 11 2018.

[54] Sora Park, Caroline Fisher, Jee Young Lee, and Kieran McGuinness. Covid-19: Australian news and misinformation. 2020.

[55] Sarah Perez. Report: Whatsapp has seen a 40% increase in usage due to covid-19 pandemic. https://tinyurl.com/bdcw29ct, 03 2020. Accessed: 2021-06-07.

[56] Kunal Purohit. Misinformation, fake news spark india coronavirus fears. https://tinyurl.com/yde9n8sj, 03 2020. Accessed: 2021-10-21.

[57] Gustavo Resende, Philipe Melo, Julio C. S. Reis, Marisa Vasconcelos, Jussara M. Almeida, and Fabrício Benevenuto. Analyzing textual (mis)information shared in whatsapp groups. In *Proceedings of the 10th ACM Conference on Web Science*, WebSci '19, page 225–234, New York, NY, USA, 2019. Association for Computing Machinery.

[58] Gustavo Resende, Philipe Melo, Hugo Sousa, Johnnatan Messias, Marisa Vasconcelos, Jussara Almeida, and Fabrício Benevenuto. (mis)information dissemination in whatsapp: Gathering, analyzing and countermeasures. In *The World Wide Web Conference*, WWW '19, page 818–828, New York, NY, USA, 2019. Association for Computing Machinery.

[59] Margaret Roberts and Margaret E Roberts. *Censored*. Princeton University Press, 2018.

[60] Henry L Roediger III and Lisa Geraci. Aging and the misinformation effect: A neuropsychological analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(2):321, 2007.

[61] Jon Roozenbeek, Claudia R. Schneider, Sarah Dryhurst, John Kerr, Alexandra L. J. Freeman, Gabriel Recchia, Anne Marthe van der Bles, and Sander van der Linden. Susceptibility to misinformation about covid-19 around the world. *Royal Society Open Science*, 7(10):201199, 2020.

[62] Johnny Saldañna. *The Coding Manual for Qualitative Researchers*. SAGE, Los Angeles, 2nd ed edition, 2013.

[63] Irving Seidman. *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*. Teachers College Press, 2013.

[64] Michael Seufert, Tobias Hoßfeld, Anika Schwind, Valentin Burger, and Phuoc Tran-Gia. Group-based communication in whatsapp. pages 536–541, 2016.

[65] Signal. Speak freely. https://signal.org/. Accessed: 2021-07-09.

[66] Lisa Singh, Leticia Bode, Ceren Budak, Kornraphop Kawintiranon, Colton Padden, and Emily Vraga. Understanding high- and low-quality url sharing on covid-19 twitter streams. *Journal of Computational Social Science*, 3:1–24, 11 2020.

[67] Kate Starbird, Ahmer Arif, and Tom Wilson. Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), November 2019.

[68] Wanning Sun. Chinese diaspora and social media: Negotiating transnational space. In *Oxford Research Encyclopedia of Communication*. 2021.

[69] Edson C Tandoc Jr and James Chong Boi Lee. When viruses and misinformation spread: How young singaporeans navigated uncertainty in the early stages of the covid-19 outbreak. *New Media & Society*, page 1461444820968212, 2020.

[70] H. Tankovska. Countries with the most whatsapp users 2019. https://www.statista.com/statistics/289778/countries-with-the-most-facebook-users/, 01 2019. Accessed: 2021-06-23.

[71] H. Tankovska. Most popular global mobile messenger apps as of january 2021, based on number of monthly active users. https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/, 02 2021. Accessed: 2021-06-23.

[72] Mazumder Hoimonty Tasnim Samia, Hossain Md Mahbub. Impact of rumors and misinformation on covid-19 in social media. *J Prev Med Public Health*, 53(3):171–174, 2020.

[73] Telegram. Telegram. a new era of messaging. https://telegram.org/. Accessed: 2021-07-09.

[74] Mayowa Tijani. How to spot covid-19 misinformation on whatsapp. https://factcheck.afp.com/how-spot-covid-19-misinformation-whatsapp, 04 2020. Accessed: 2021-10-21.

[75] Dragana Trninić, Anđela Kuprešanin Vukelić, and Jovana Bokan. Perception of “fake news” and potentially manipulative content in digital media—a generational approach. *Societies*, 12(1):3, 2022.

[76] Melissa Tully. Everyday news use and misinformation in kenya. *Digital Journalism*, pages 1–19, 2021.

[77] Christian Unkelbach and Rainer Greifeneder. Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, 79:78–86, 2018.

[78] Ben Gia Minh Vo. Vietnamese america: On ‘good refugees’, fake news, and historical amnesia. *Asian American Research Journal*, 1(1), 2021.

[79] WhatsApp Help Center. How to delete messages. https://faq.whatsapp.com/android/chats/how-to-delete-messages/?lang=en. Accessed: 2021-10-29.

[80] WhatsApp Help Center. About end-to-end encryption. https://faq.whatsapp.com/general/security-and-privacy/end-to-end-encryption/?lang=en, 2021. Accessed: 2021-06-08.

[81] WhatsApp LLC. Whatsapp 2.0 is submitted. https://blog.whatsapp.com/whats-app-2-0-is-submitted, 2009. Accessed: 2021-03-14.

[82] WhatsApp LLC. Keeping whatsapp personal and private. https://blog.whatsapp.com/Keeping-WhatsApp-Personal-and-Private, 04 2020. Accessed: 2021-06-23.

[83] WhatsApp LLC. Search the web. https://blog.whatsapp.com/search-the-web, 08 2020. Accessed: 2020-09-20.

[84] Sam Wineburg and Sarah McGrew. Evaluating information: The cornerstone of civic online reasoning. 2016.

[85] Liang Wu, Fred Morstatter, Kathleen M. Carley, and Huan Liu. Misinformation in social media: Definition, manipulation, and detection. *SIGKDD Explor. Newsl.*, 21(2):80–90, November 2019.

[86] YouTube Help. Covid-19 medical misinformation policy. https://support.google.com/youtube/answer/9891785?hl=en, 05 2020. Accessed: 2021-07-02.

Appendix A: Interview Questions

General WhatsApp usage

- Why do you use WhatsApp? (vs. other social media or messaging apps like iMessage, Facebook Messenger, etc.)
- Is WhatsApp your primary communication app?
- How often do you use WhatsApp?
- How long have you had WhatsApp?
- What do you think are the pros and cons of WhatsApp?
- How many contacts do you have on WhatsApp?
- What relationship do you have with your contacts? Are they friends? Family? Work colleagues? Acquaintances? Others?
- What do you usually talk about on WhatsApp? Do you share links when you talk?
- Are most of your conversations on WhatsApp direct messages or group chats?
  - Can you give a ballpark percentage of the conversations that happen in private messages vs. in group chats?
  - How large are your group chats? Who are in them?
- Do you know anything about WhatsApp's end-to-end encryption?

Encounters of doubtful information

- What concerns do you have about false, inaccurate, or misleading information in WhatsApp? If none, why?
- Have you ever seen or received any information on WhatsApp that you thought was false or misleading? If so, what happened? What did you do?
  - Who sent it to you?
  - Did you forward it?
  - Did the information consist of images, text, articles, or videos that you thought weren't accurate? Why did you think they were inaccurate?
  - How often do you see this type of content?
  - Has similar content ever appeared on another social media/messaging platform (e.g. Facebook News Feed)?
- What factors do you consider when deciding to trust information sent to you via WhatsApp?
- Do you forward information to your contacts?

Misinformation and recent events (COVID-19, BLM protests, U.S. election etc.)

- What kinds of information on COVID-19 have you received around WhatsApp?
- When was the last time you got a message on WhatsApp about COVID-19? What was it about? Did you think it was accurate? Why/why not?
- Have you seen more information sharing around COVID-19 on WhatsApp compared to before December 2019?
- Have you seen false, inaccurate, or misleading information around COVID-19 on WhatsApp? If so, can you give an example?
  - What did you do?
  - How did the information affect you?
  - Did you talk to the sender about it?
  - Did you fact-check it?
  - Did you ignore it?
- How has the information you've seen on WhatsApp affected your view/opinion on the country's (U.S.) situation with the pandemic (e.g. reopening phases, how COVID-19 affects youth, number of reported cases, conspiracy theories about origins of the virus)?
- How has the information on mask wearing/quarantine/social distancing affected your viewpoint with the COVID-19 information you receive?
  - How has the information on mask wearing + protests affected your viewpoint with the COVID-19 information you receive?
  - What about stay-at-home?
  - What about social distancing?
- What other messages about recent events have you received so far (BLM, elections, schools reopening)?
  - How have they affected your views on these issues?
  - How about your views on COVID-19, if at all?

Technology + fact-checking strategies

App features referenced are shown in Fig 1 (in the main paper).

- Have you used the WHO Health Alert on WhatsApp? If not, why?
  - If yes, what did you think of its helpfulness/usefulness? How easy was it to use?
- The CDC has a bot on WhatsApp you can text to give you information on what to do if you think if you have symptoms. Have you ever used this? If not, why?
  - If yes, what did you think of its helpfulness/usefulness? How easy was it to use?
- Have you seen a new magnifying glass icon pop up beside some of your messages recently?
  - If so, have you tapped on it?
  - What did it lead you to and what did you think of it?
- How do you know what information given to you on WhatsApp can be trusted (or in general)?
  - What do you use to fact-check, if anything at all?
- What's your opinion on WhatsApp limiting the number of forward messages to lessen the spread of false information?
  - What led you to that opinion?
  - The limit is that one can only forward a message to 5 chats at a time.
  - When message is forwarded in a chain 5 times, it can only be forwarded to one chat (indicated with double arrow).
- Do you think WhatsApp can be improved to help address these issues with false, inaccurate, or misleading information? Why or why not?
- With other resources like Twitter's COVID-19 misinformation warnings (Fig. 1(d) in the main paper) and YouTube's information alert boxes (Fig. 1(c) in the main paper), would you want a better way to fact-check information in WhatsApp? Do you think these are enough? Why or why not?

Conclusion

- How has anything you said been vastly different from how you send or receive messages on other social media platforms you use?
- Is there anything else regarding WhatsApp that you want to talk about?
  - Desired technology?
  - False/inaccurate information?

Appendix B: Codebook

Code	Explanation
General	
Reason for using/liking WhatsApp	Participant explained why they like or use WhatsApp
Reason for disliking WhatsApp	Participant explained why they dislike WhatsApp, if they dislike it in any way
Chat Content	Participant talked about what they usually talked about in the chats, broadly
Foreign (non-U.S.) vs domestic communication	Participant uses WhatsApp to communicate with people in or out of the U.S.
Size of groups/chats they're in	Participants estimated the average size of the group chats they are in. They also gave exact numbers if they remember, or if they were in very few groups
Relationship with others in the group (with whom they interact with most often)	Participants identified relationships with others in their group chats
Active contacts/chat groups	Participants estimated the number of WhatsApp contacts they interacted with on a regular basis
Misinformation Encounters	
Information format (image/video/audio/text/links)	Participant describes the format of the information presented to them
Most recent misinformation encounter	Participant recounts most recent misinformation counter (info content, who sent it, their reaction, etc.)
Frequency of encountering misinformation	How often does a participant encounter misinformation? (e.g. once a week, month, year, etc.)
Who sends them misinformation content	Participant describes relationship with the misinformation sender (relative from abroad, immediate family member, etc.)
Frequency of forwarding links	Participant describes how often they forward links to their chats and messages
Misinformation indicators	Participant describes factors they consider when deciding to trust (and distrust) information
Reason for being active (talking with sender, fact-checking) about receiving misinformation	Participant explains how and why they are proactive when receiving misinformation (confronting sender, fact-checking)
Reason for being passive (ignoring) about receiving misinformation	Participant explains how and why they are passive/inactive when receiving misinformation
How WhatsApp content impacted their opinion on how the U.S. handled the pandemic	Participant explains how what they read on WhatsApp has impacted their opinion of how the U.S. handled the pandemic
How WhatsApp content impacted their opinion on BLM, 2020 elections, school reopenings	Participant explains how what they read on WhatsApp has impacted their opinion on other recent events: BLM, U.S. elections, U.S. school reopenings
Design Recommendations and Fact-Checking Strategies	
Willingness to use existing WhatsApp technology from reliable sources	Participants share their awareness of existing resources on WhatsApp from reliable sources designed to combat COVID-19 misinformation, namely the CDC bot
Fact-checking strategies	Participant describes how they fact-check information (Google search, literature, consulting others, etc.)
Efficacy of current WhatsApp features that combat misinformation	Participant describes the efficacy of WhatsApp features in fact-checking and limiting the spread of misinformation
Suggestions for improvement	Participant suggests improvements of current WhatsApp in bettering misinformation prevention/clarification

Concerns about the trade-off between combating misinformation and privacy/security	Participant raises concerns that fact-checking measures (e.g. information censorship) may undermine the privacy and comfort associated with end-to-end encryption
Features of other platforms	Participants share their opinions of existing features on other social media platforms (YouTube, Twitter, etc.) to combat misinformation.

Table 1: Our codes and corresponding explanations, organized by topic.