medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ORIGINAL PAPER

Tracking private WhatsApp discourse about COVID-19: A longitudinal infodemiology study in Singapore

Edina YQ Tan[1,2], Russell RE Wee[3], Saw Young Ern[1], Kylie JQ Heng [1], Joseph WE Chin[1], Eddie M.W. Tong[3], Jean CJ Liu[1,2,4*]

[1]Division of Social Sciences, Yale-NUS College, 16 College Avenue West, Singapore 138527, Singapore

[2]Centre for Sleep and Cognition, Yong Loo Lin School of Medicine, National University of Singapore, 12 Science Drive 2, Singapore 117549, Singapore

[3]Department of Psychology, National University of Singapore, 9 Arts Link, Singapore 117572, Singapore

[4]Neuroscience and Behavioral Disorders Program, Duke-NUS Medical School, 8 College Road, Singapore 169857, Singapore

*Address correspondence to:

Jean Liu

28 College Avenue West #01-501

Singapore 138533

Email: jeanliu@yale-nus.edu.sg

Phone: (+65) 6601 3694

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Background: Worldwide, social media traffic increased following the onset of the coronavirus disease (COVID-19) pandemic. Although the spread of COVID-19 content has been described for several social media platforms (e.g., Twitter, Facebook), little is known about how content is spread via private messaging platforms such as WhatsApp.

Objective: In this study, we documented: (i) how WhatsApp is used to transmit COVID-19 content; (ii) the characteristics of WhatsApp users based on their usage patterns; and (iii) how usage patterns link to well-being.

Methods: We used the experience sampling method to track day-to-day WhatsApp usage during the COVID-19 pandemic. For one week, participants reported each day the extent to which they had received, forwarded, or discussed COVID-19 content. The final dataset comprised 924 data points collected from 151 participants.

Results: During the week-long monitoring, most participants (143/151, 95%) reported at least one COVID-19-related use of WhatsApp. When a taxonomy was generated based on usage patterns, 1 in 10 participants (21/151, 14%) were found to have received and shared a high volume of forwarded COVID-19 content – akin to 'super spreaders' identified on other social media platforms. Finally, those who engaged with more COVID-19 content in their personal chats were more likely to report having COVID-19 thoughts throughout the day.

Conclusions: These findings provide a rare window into discourse on private messenger platforms. In turn, this can inform risk communication strategies during the pandemic.

Trial Registration: *ClinicalTrials.gov NCT04367363*

Keywords: *Social Media, WhatsApp, Infodemiology, Misinformation, COVID-19*

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction

WhatsApp is the most commonly used messaging application worldwide, with 1.5 billion users across 180 countries [1]. On account of its large user base and near-instant message transmission, the platform has played a critical role in risk communication during the coronavirus disease (COVID-19) pandemic.

On the one hand, WhatsApp has been co-opted by government agencies and the World Health Organization to disseminate official COVID-19 updates [2]. While this showcases the platform's ability to reach a large sector of the population, the same feature has also made it a vessel for misinformation. For example, at the beginning of the pandemic, WhatsApp noted a 40% surge in usage [3]. This was paired with a high volume of message-forwarding activity widely believed to support misinformation, leading the platform to restrict the number of individuals a message could be forwarded to simultaneously [4,5].

Despite these restrictions, a survey in India found that 1 in 2 participants had received COVID-19 misinformation through WhatsApp or Facebook [6]. Likewise, WhatsApp was identified by Hong Kong residents as the foremost source for COVID-19 rumors [7]. As misinformation can jeopardize public health strategies, these findings underscore the need for infodemiological studies documenting how COVID-19 content spreads through WhatsApp.

To date, however, the bulk of infodemiology studies have focused on social media platforms where content is publicly accessible (e.g., Twitter, Facebook) [8,9]. By contrast, research on WhatsApp has proven elusive because of its private nature, with end-to-end encryption ensuring that only senders and recipients have access to messages sent through the platform. Nonetheless, WhatsApp research remains a priority: aside from its popularity and role in carrying crisis-related misinformation [10,11], insights from public posts are also unlikely to generalize to WhatsApp's private

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

messages [12]. It thus remains unclear who sends COVID-19 messages, to whom, and in what manner.

Addressing these gaps in the literature, we designed a study to: (i) describe the base rate of COVID-19 content dissemination; (ii) understand WhatsApp users; and (iii) examine correlates of usage patterns. Specifically, we applied the experience sampling method to track one week of WhatsApp usage amidst everyday routines [13,14], asking participants to report each day their frequency of receiving, forwarding, or discussing COVID-19-related content. Through this method, we generated a taxonomy of participants based on their usage patterns, and examined whether day-to-day variations in WhatsApp usage predicted COVID-19 concerns.

Methods

Recruitment

Between 17 March to 7 May 2020, participants were recruited from the general community via advertisements placed in Facebook and WhatsApp community groups (e.g., residential groups, workplace groups, university groups), posts on popular online forums, and paid Facebook advertisements targeting Singapore-based users. All study activities took place via the online survey platform Qualtrics, and participants were reimbursed SGD $5 upon study completion. The study protocol was approved by the Yale-NUS College Ethics Review Committee (2020-CERC-001) and was pre-registered at ClinicalTrials.gov (NCT04367363).

Participants

Participants were 151 adults who met the following inclusion criteria: (1) aged 21 years or older, (2) had lived in Singapore for at least 2 years, and (3) had a WhatsApp account.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Measures

Following informed consent, participants completed: (i) a baseline questionnaire; (ii) experience sampling responses daily for 7 days; and (iii) a final questionnaire (Figure 1).

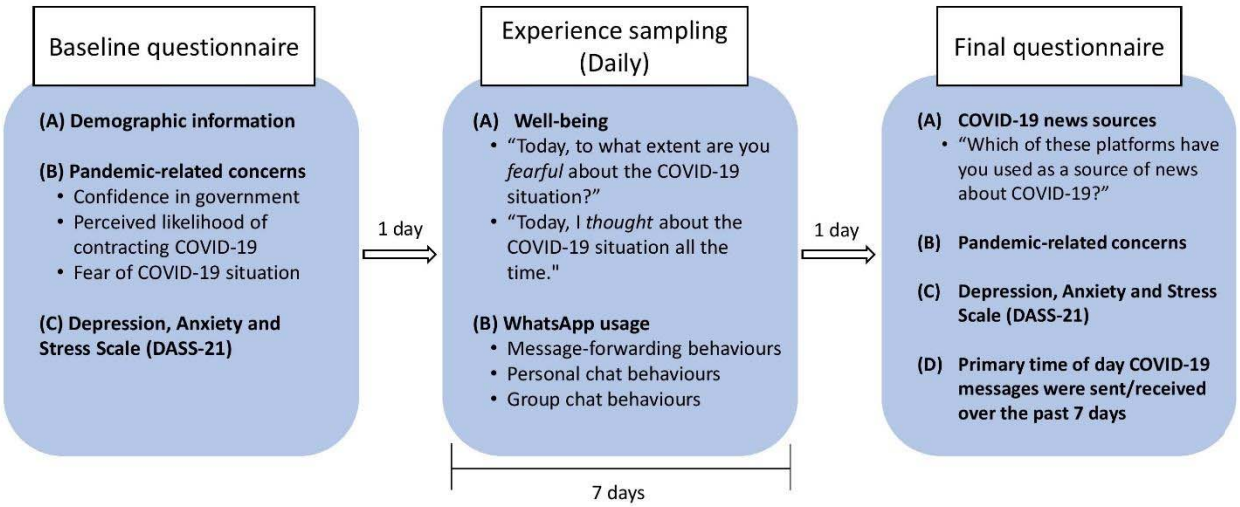

Figure 1. Schematic of study procedures. All participants completed a baseline questionnaire, followed by 7 days of experience sampling where participants addressed questions about well-being and WhatsApp usage daily. Participants completed a final questionnaire one day after the experience sampling protocol ended.

Experience sampling

As the primary form of data collection, we used the experience sampling method to capture "COVID-19 chatter" on WhatsApp. Through this protocol, we collected 924 data points across 151 participants (compliance rate: 86.2%).

For 7 days, participants accessed an online survey each evening (2130 hrs) to report WhatsApp usage for the day. First, participants indicated whether they had forwarded messages related to COVID-19 (yes/no). We focused on message-forwarding as a proxy indicator for high-risk content, since: (i) large Twitter studies have observed that misinformation is more likely to be shared than posts that are true [15], and (ii) because WhatsApp developers had previously linked forwarded messages to misinformation [4,5]. If participants had forwarded COVID-19 content, they were then

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

asked: (i) how many unique COVID-19 messages they had forwarded, and (ii) how many unique groups and individuals they had forwarded messages to.

Additionally, participants were asked about their personal chats – their one-to-one chats on WhatsApp. They indicated whether COVID-19 messages had been forwarded to them in personal chats (yes/no), and, if so: (i) how many unique messages they had received, and (ii) how many different people they had received messages from. Thereafter, participants recounted whether they had discussed COVID-19 in conversations where either they or the other party generated message(s) related to COVID-19 (yes/no), and – if so – how many unique chats were involved.

Finally, for group chats, participants were asked if COVID-19 had been mentioned in any of their WhatsApp groups by at least one other person (not including themselves) (yes/no). This could have occurred either through others forwarding messages, or through others generating their own comments. Affirmative responses were followed with a question on how many WhatsApp groups had done so.

Aside from WhatsApp metrics, participants also reported their COVID-19 concerns for the day: (i) how afraid they felt about the COVID-19 situation (4-point scale with 1= “Not scared at all” and 4=“Very scared”); and (ii) whether they had thought about the COVID-19 situation all the time (5-point scale with 1=“Not at all true” and 5=“Very true”).

Baseline and final questionnaires

To characterize participants, we included baseline and final questionnaires where participants reported: demographics (age, gender, religion, ethnicity, marital status, education, house type, household size, citizenship, country of birth, and years in Singapore), the time of day they read and sent COVID-19 messages on WhatsApp (mostly in the morning, afternoon, evening, late night, or throughout the day), and

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

sources through which they obtained COVID-19 news (e.g., printed newspapers, radio, WhatsApp, YouTube). Additionally, participants reported their well-being (using the 21-item Depression, Anxiety and Stress Scale (DASS-21; [16]), and responses to the pandemic [2,17]: (1) how confident they were that the government could control the nationwide spread of COVID-19 (1 = “Not confident at all, 4 = “Very confident”); (2) how likely they judged that they (or someone in their immediate household) would be infected with COVID-19 (1 = “Not at all likely”, 4 = “Very likely”); and (3) how fearful they were about the situation in the country (1 = “Not scared at all”, 4 = “Very scared”).

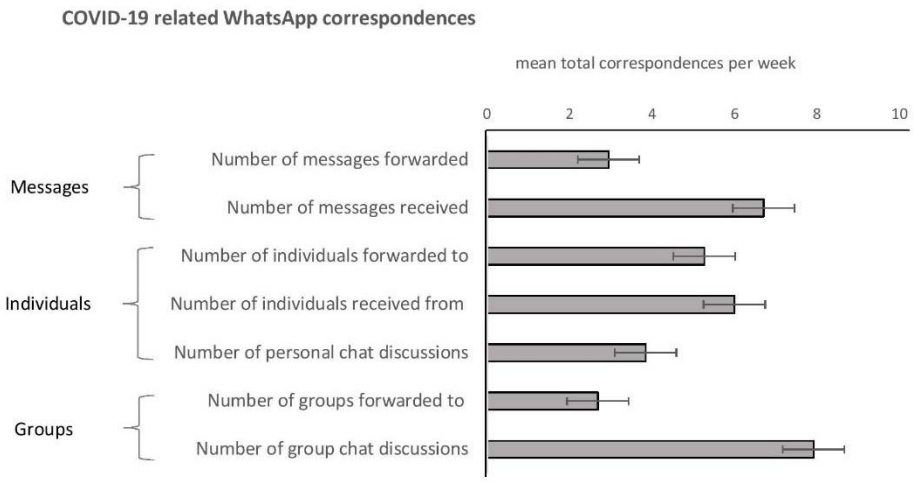

Figure 2. Distribution of COVID-19 related WhatsApp behaviors. In a week-long experience sampling procedure, participants reported the extent to which they engaged in COVID-19 WhatsApp behaviors (either by forwarding or receiving messages, or in conversations). Horizontal bars represent the total amount of each activity captured, averaged across all participants. Horizontal lines represent the 95% confidence interval for the mean.

Statistical Analysis

We first summarized the data with counts (%) or means (standard deviations), focusing on the 7 quantitative WhatsApp usage variables (Figure 2): the number of (1) COVID-19 messages participants forwarded; (2) groups forwarded to; (3) individuals forwarded to; (4) forwarded messages received; (5) individuals from whom messages were received; (6) personal chats involving COVID-19 conversations; and (7) group

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

chats discussing COVID-19. (Multimedia Appendix 1 shows the pattern of correlations across these variables).

Next, to understand user profiles, we applied latent profile analysis to create a taxonomy of participants based on their WhatsApp usage (R package: 'mclust' [18]). Latent profile analysis is a bottom-up statistical clustering method that defines classes of people based on common characteristics. Using all observations of a continuous dependent variable, classes are created such that within each class, indicator variables are statistically uncorrelated [19]. We thus applied this technique to cluster participants based on their responses to the 7 WhatsApp usage variables, with values obtained by aggregating the reported frequency of each variable over the week. To uncover clusters, we used Gaussian Mixture Models (GMM) and assigned cluster membership using Bayesian probabilities. The final number of clusters was determined using the Bayesian Information Criterion (BIC), Integrated Completed Likelihood Criterion (ICL), and a Bootstrap Likelihood Test (BLRT).

Finally, we examined whether day-to-day variations in COVID-19 WhatsApp chatter was tracked by variations in COVID-19 concerns. As predictors, we first quantified COVID-19 chatter on personal and group chats. For personal chats, the following variables were summed within each day and for each participant: the number of (i) individuals COVID-19 messages were forwarded to; (ii) individuals from whom forwarded messages were received; and (iii) personal conversations discussing COVID-19. For group chats, the following variables were summed: the number of (i) groups participants forwarded COVID-19 messages to; and (ii) groups where COVID-19 messages were mentioned. Scores were grand-mean centered by subtracting the mean number of chats across subjects and time points from each score (M = 2.47 and M = 1.29 respectively). In addition, we created between- and within-subject versions of each predictor [20]. The final analyses involved linear mixed-effects models for each outcome

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

measure (fear and thoughts about COVID-19), with the following entered as fixed effects: time (centered such that 0 referred to the middle of the week), daily personal chats (between-subjects), daily personal chats (within-subjects), daily group chats (between-subjects), and daily group chats (within-subjects). Random intercepts accounted for correlated data due to repeated measures.

Across all analyses, the type 1 decision-wise error rate was controlled at $\alpha = 0.05$. All statistical analyses were conducted in R 3.5.0 [21] and SPSS 25 [22].

Results

Baseline Participant Characteristics

As shown in Table 1, 69% (104/151) of participants identified as female, with a mean age of 36.35 years (SD 14.7). Participants were predominantly of Asian ethnicity (93% Chinese; 140/151), and had at least post-secondary education (88.1%, 133/151). 40% (60/151) were married, and the majority (69.5%, 105/151) belonged to households of at least 4 members.

Base rate of COVID-19 WhatsApp usage

Participants' self-reports revealed that WhatsApp was the second-most common source for COVID-19 news, after news website or apps (Figure 3). Quantifying this through one week of experience sampling, we found that nearly all participants (95%, 143/151; 95% CI: 90-98%) reported at least one COVID-19 related use of WhatsApp. Namely, 1 in 2 participants (52%, 79/151; 95% CI: 44-60%) forwarded at least one COVID-19 message (to either individuals or groups), 78% (118/151; 95% CI: 71 - 84%) received at least one forwarded message in personal chats, 66% (100/151; 95% CI: 58-74%) engaged in personal chat conversations about COVID-19, and 88% (133/151; 95% CI: 82% - 93%) had been in groups where COVID-19 was mentioned.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 2 shows the extent to which participants engaged in each of these activities. On average, participants: (i) received two times more messages (2.3) than they forwarded on, and (ii) were more likely to forward messages to individuals than to groups (average of 5.3 versus 2.7 messages a week, respectively). Beyond passive engagement, participants also took part in an average of 3.8 one-to-one conversations about COVID-19 during the week; however, these interactions occurred less frequently than the sending or receiving of forwarded messages in group chats.

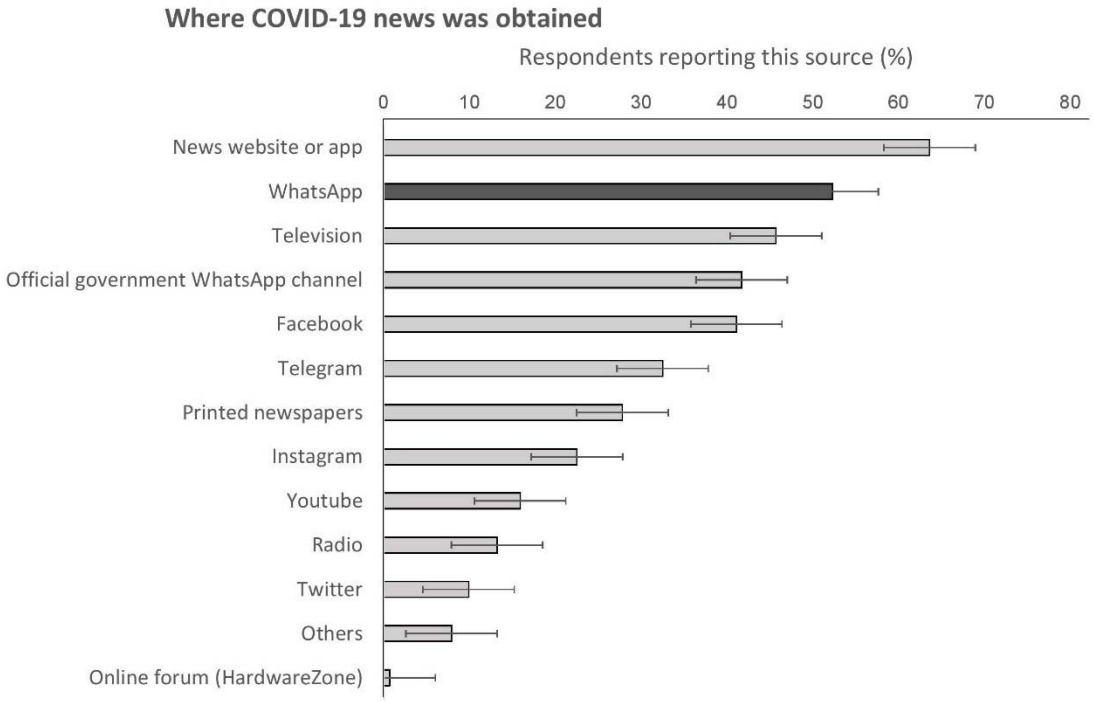

Figure 3. Sources of COVID-19 news. In a questionnaire, participants self-reported where they received COVID-19 news from.

Characterizing participants based on COVID-19 WhatsApp usage

Latent profile analysis: Generating a taxonomy of WhatsApp usage

Although most participants received and shared COVID-19 content on WhatsApp, there were individual differences in usage patterns (Figure 4).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Correspondingly, we conducted a latent profile analysis to understand how usage patterns clustered.

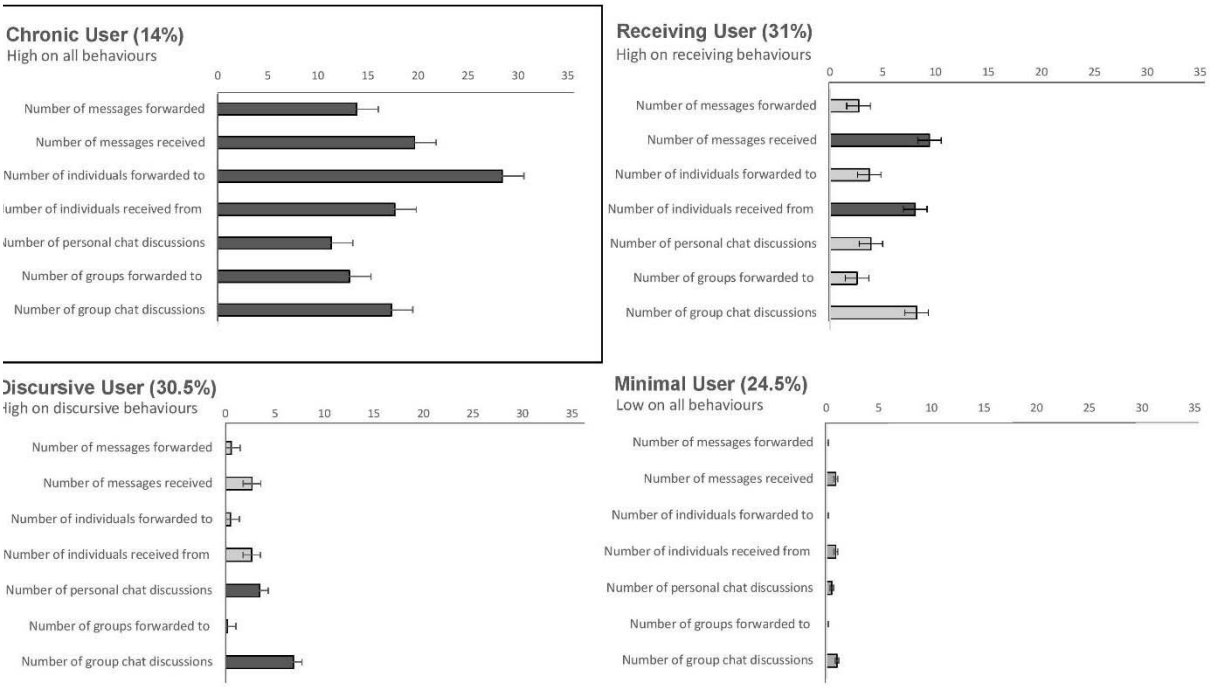

Figure 4. Taxonomy of COVID-19 related WhatsApp usage. Using latent profile analysis, participants were classified based on how they had used WhatsApp for COVID-19 content during a week of monitoring. The figure depicts WhatsApp usage activities for the chronic user (top left), receiving user (top right), discursive user (bottom left), and minimal user (bottom right). Horizontal lines represent the 95% confidence interval for the mean.

A 4-cluster solution yielded the lowest absolute BIC values (Multimedia Appendix 2), resulting in the following taxonomy (Figure 4). First, 14% (21/151) of participants were “chronic users” with high activity in each of the WhatsApp usage variables. Correspondingly, this group of participants was responsible for receiving and transmitting a large volume of forwarded COVID-19 messages, sending the messages both to individual contacts and to groups. Second, 31% (47/151) of participants were “receiving users” distinguished by their receipt of multiple COVID-19 forwarded messages. Although this group discussed COVID-19 frequently in group chats, they rarely passed along forwarded COVID-19 messages they had received. A third group,

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

"discursive users" (31%, 46/151), had low exposure to forwarded COVID-19 messages and primarily engaged with COVID-19 content through personal and group chats. Finally, 25% (37/151) of participants were "minimal users" with low engagement with COVID-19 content overall.

Understanding user characteristics

As an exploratory analysis, we performed a classification tree analysis to predict WhatsApp user type based on: demographics, COVID-19 concerns, depression and anxiety scores (DASS-21), and the time of day for WhatsApp usage. We applied recursive partitioning ("rpart"), a machine learning technique that allows multiple variables to be analyzed simultaneously, and supports the modelling of complex, non-linear relations between predictors [23]. To avoid overfitting, the final tree was pruned by selecting the tree size with the lowest cross-validation error (minimized with a tree size of 8 for our dataset).

As shown in Figure 5, chronic users were more likely to be married / divorced, and to send messages either throughout the day or in the afternoon. In terms of COVID-19 responses, chronic users either had: (i) extreme fears of the COVID-19 situation (low or high), or (ii) had moderate fears paired with lower confidence in the government's response (low or moderate).

On the other hand, discursive users were more likely to be single / dating, and had either: (i) extreme levels of COVID-19 fears (either high or low), or (ii) moderate fear levels alongside Christian or Taoist affiliations. A sub-group of discursive users were – like chronic users – married / divorced, with moderate levels of COVID-19 fears. However, they were distinguished from chronic users by their high confidence in government (as compared to chronic users' lower confidence).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Finally, receiving and minimal users had similar profiles. If single / dating, both sets of users tended to have moderate levels of COVID-19 fears, had a wide range of religious backgrounds, and were distinguished by the time of day that COVID-19 related messages were received (receiving users: morning, evening, and throughout the day; minimal users: afternoon and night). If married / divorced, both sets of users tended to send messages at only one time of the day (morning, evening, or night), and were distinguished by age (receiving users: 51 years and above; minimal users: below 51 years). Table 1 describes the demographic characteristics of the four profiles.

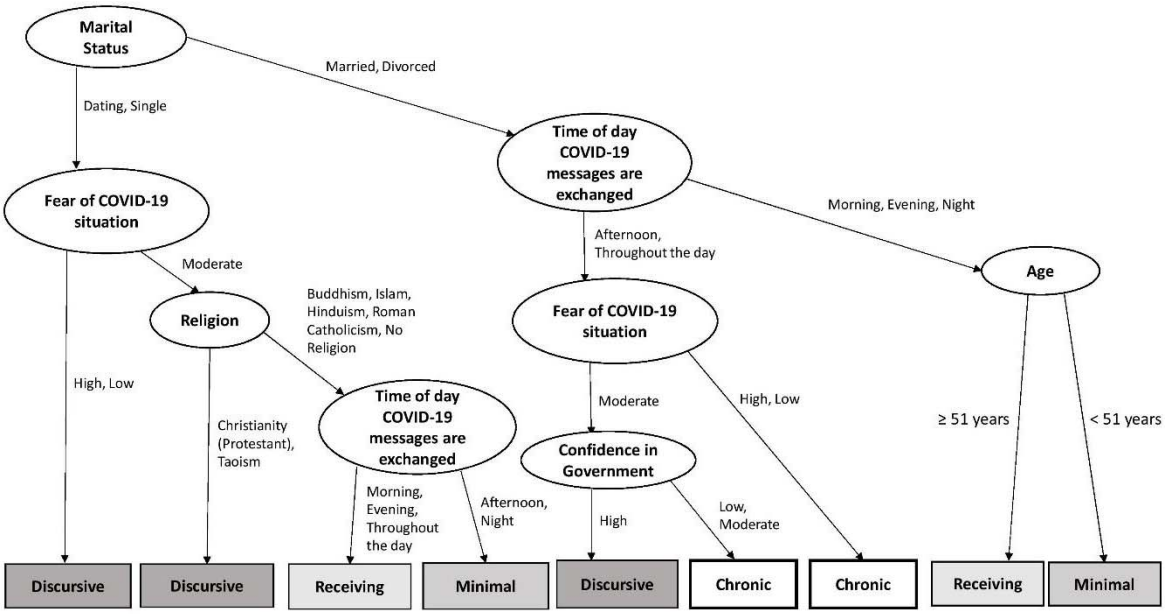

Figure 5. Classification tree analysis. Recursive partitioning was used to predict which of 4 WhatsApp usage profiles (chronic, receiving, discursive, or minimal) participants belonged to, based on baseline questionnaire measures (demographics, COVID-19 concerns, scores on the Depression, Anxiety and Stress Scale, and time of WhatsApp usage). The final tree model is presented as a flow chart, with factors chosen at each level to categorize the maximal number of participants. Marital status, time of WhatsApp usage, and age emerged as the primary predictors (model classification accuracy: 64.2%, above the chance level of 25%).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 1. Participant characteristics as a function of COVID-19 WhatsApp usage patterns.

Characteristic	Chronic (n = 21)	Receiving (n = 47)	Discursive (n = 46)	Minimal (n = 37)	(%) or M (SD)
Age in years	44.1 (14.5)	41.0 (15.5)	29.7 (10.7)	34.4 (14.5)	36.35 (14.70)
Gender					
· Female	13	34	29	28	(68.8)
· Male	8	13	17	9	(31.2)
Ethnicity					
· Chinese	20	42	42	36	(92.7)
· Indian	0	2	3	0	(3.3)
· Malay	0	2	1	0	(2.0)
· Others	1	1	0	1	(2.0)
Religion					
· Christianity (Protestant)	8	17	16	13	(35.8)
· No religion	3	14	11	10	(25.2)
· Buddhism	4	9	8	11	(21.2)
· Roman Catholicism	4	4	6	2	(10.6)
· Taoism / Chinese traditional	1	0	2	1	(2.6)
beliefs	1	3	1	0	(3.3)
· Islam	0	0	2	0	(1.3)
· Hinduism					
Marital status					
· Married	13	24	8	15	(39.7)
· Single	6	15	25	12	(38.4)
· Dating	1	7	12	9	(19.2)
· Widowed / separated /	1	1	0	1	(2.0)
divorced	0	0	1	0	(0.7)
· Did not answer					
Educational level					
· 1: O Level	1	4	1	6	(7.9)
· 2: Junior College	2	5	9	9	(16.6)
· 3: ITE	1	1	1	0	(2.0)
· 4: Polytechnic/diploma	2	13	7	4	(17.2)
· 5: University (undergraduate)	11	21	21	16	(45.7)
· 6: University (postgraduate)	4	1	3	2	(6.6)
· 7: Others	0	2	3	0	(3.3)
· Did not answer	0	0	1	0	(0.7)
House type					
· 1: HDB flat: 1-2 rooms	0	0	0	1	(0.7)
· 2: HDB flat: 3 rooms	0	2	2	2	(4.0)
· 3: HDB flat: 4 rooms	2	9	10	10	(20.5)
· 4: HDB flat: 5 rooms	3	19	14	11	(31.1)
· 5: Condominium	12	12	11	10	(29.8)
· 6: Landed property	4	4	7	2	(11.3)
· Did not answer	0	1	2	1	(2.6)
Household size					
· 1	2	1	3	0	(4.0)
· 2	0	5	3	3	(7.3)
· 3	7	8	5	8	(18.5)
· 4	7	18	21	15	(40.4)
· 5+	5	15	13	11	(29.1)
·Did not answer	0	0	1	0	(0.7)

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Citizenship					
· Singapore	18	46	42	36	(94.0)
· Others	3	1	4	1	(6.0)
Country of birth					
· Singapore	17	45	38	33	(88.1)
· Other	4	2	8	4	(11.9)
Years in Singapore	39.67 (15.22)	39.60 (16.69)	26.43 (10.83)	31.65 (14.47)	33.65 (15.32)
DASS-21					
Stress	9.52 (7.12)	8.61 (7.08)	9.56 (10.13)	10.81 (8.72)	9.57 (8.47)
Anxiety	4.38 (5.28)	5.13 (5.44)	5.33 (6.59)	5.89 (7.71)	5.28 (6.36)
Depression	8.10 (6.52)	7.22 (6.86)	9.47 (9.73)	10.76 (9.54)	8.90 (8.50)
Pandemic-related concerns					
Fear of COVID-19 situation	2.29 (0.46)	2.53 (0.65)	2.22 (0.74)	2.27 (0.69)	2.34 (0.67)
Confidence in government	3.33 (0.58)	3.23 (0.63)	3.29 (0.66)	3.24 (0.72)	3.27 (0.65)
Perceived likelihood of contracting COVID-19	2.71 (0.64)	2.74 (0.53)	2.78 (0.56)	2.76 (0.60)	2.75 (0.57)

Does WhatsApp usage relate to COVID-19 concerns?

As the final thrust of the study, we ran linear mixed effects models to examine whether WhatsApp usage related to COVID-19 concerns (Table 2). As shown in Figure 6, day-to-day COVID-19 fears and thoughts fluctuated ($t(249.13) = -3.72$, $P < 0.001$; $t(297.02) = -2.36$, $P = 0.02$).

For COVID-19 thoughts, there was a significant effect of WhatsApp personal chat usage at a between-subjects level; $t(164.81) = 2.36$, $P = .02$). That is, participants who handled higher levels of COVID-19 content in their personal chats reported having more COVID-19 thoughts (relative to participants who handled lower levels of COVID-19 content). However, the corresponding effect for group chats was not significant; $t(141.17) = 0.89$, $P = .37$. At the within-subjects level, neither day-to-day fluctuations in personal nor group chat activities significantly predicted COVID-19 thoughts (smallest $P = .68$).

For COVID-19 fears, we found no significant effect of any WhatsApp usage variable (smallest $P = .17$). (For sensitivity analyses, we repeated both models with

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

group membership as a fixed factor in place of personal and group chat usage, and our primary conclusions did not change; Multimedia Appendix 3.)

Table 2. Parameter estimates for the multi-level model of thoughts about COVID-19 (Model 1) and fear of COVID-19 (Model 2) as a function of participants' daily WhatsApp use (personal-chat and group-chat).

[Model 1] Outcome: Thoughts about COVID-19						
Fixed effects	Estimate	SE	*t*	*P*	CI_{95} lower	CI_{95} upper
Intercept	2.18	0.07	32.81	<0.001	2.05	2.31
Time (centered)	-0.03	0.01	-2.36	0.02	-0.05	-0.00
Daily personal chat usage (between)	0.04	0.02	2.36	0.02	0.00	0.07
Daily personal chat usage (within)	0.00	0.01	0.42	0.68	-0.01	0.02
Daily group chat usage (between)	0.05	0.06	0.89	0.37	-0.06	0.17
Daily group chat usage (within)	-0.00	0.03	-0.08	0.93	-0.06	0.05
Random effects	Estimate	SE	Z	*P*	CI_{95} lower	CI_{95} upper
Intercept (between)	0.56	0.08	6.89	<0.001	0.42	0.75
Residual (within)	0.37	0.02	14.90	<0.001	0.33	0.43
Autocorrelation (within)	0.24	0.05	4.97	<0.001	0.14	0.33
[Model 2] Outcome: Fear of COVID-19						
Fixed effects	Estimate	SE	*t*	*P*	CI_{95} lower	CI_{95} upper
Intercept	2.10	0.06	36.37	<0.001	1.98	2.21
Time (centered)	-0.03	0.01	-3.72	<0.001	-0.05	-0.02
Daily personal chat usage (between)	0.01	0.01	0.85	0.39	-0.02	0.04
Daily personal chat usage (within)	0.01	0.01	1.22	0.24	-0.01	0.02
Daily group chat usage (between)	0.02	0.05	0.49	0.62	-0.07	0.12
Daily group chat usage (within)	-0.03	0.02	-1.42	0.17	-0.06	0.01
Random effects	Estimate	SE	Z	*P*	CI_{95} lower	CI_{95} upper
Intercept (between)	0.44	0.06	7.47	<0.001	0.34	0.58
Residual (within)	0.21	0.01	13.83	<0.001	0.18	0.23
Autocorrelation (within)	0.26	0.05	5.16	<0.001	0.16	0.35

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

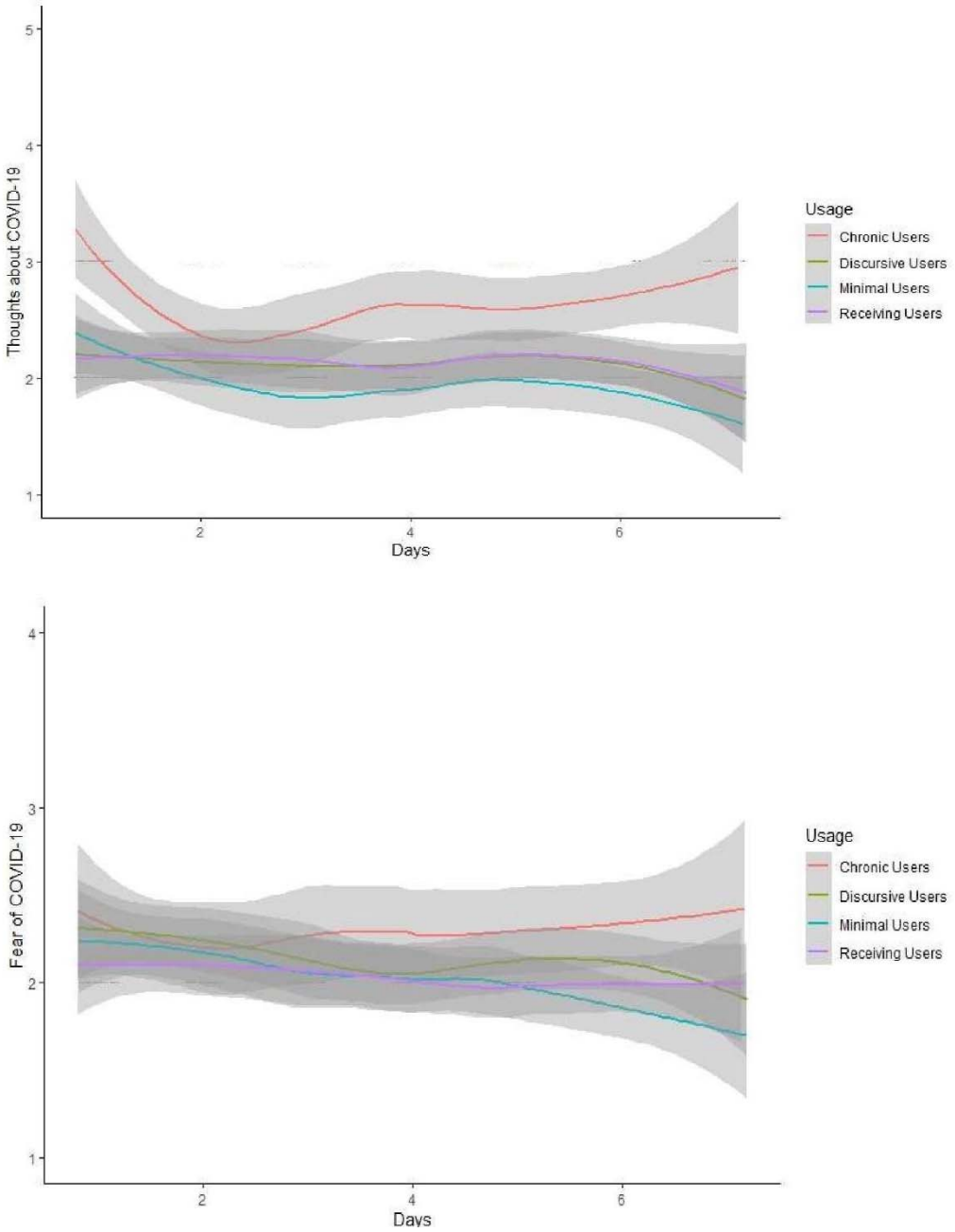

Figure 6. COVID-19 related thoughts and fears over a week. Day-to-day variations in COVID-19 thought (top) and fear levels (bottom), as a function of WhatsApp user profiles. The shaded grey areas represent 95% confidence intervals.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Discussion

The ongoing pandemic has drawn attention to the role of social media in public health. Against this backdrop, we present the first infodemiological study documenting the spread of COVID-19 content through WhatsApp. By tracking daily WhatsApp usage for one week, we found that: (i) nearly every participant engaged in "COVID-19 chatter", and that (ii) participants were more likely to share or receive forwarded messages than to engage in conversation about COVID-19.

Our findings have several implications for health communication. First, the volume of forwarded messages we observed raises concern. On other social media platforms, forwarding behaviors have been linked to the spread of misinformation. For example, a study of 4.5 million Twitter posts found that misinformation was 70% more likely to be shared than posts that were true; correspondingly, any single re-tweet had a higher probability of containing false than truthful news [15]. Although analogous research has not been conducted on WhatsApp, the app's developers have likewise targeted forwarded messages as being at high-risk for misinformation [4,5].

To the extent that forwarded messages carry misinformation [1, 2], our latent profile analyses revealed that 1 in 10 participants were 'chronic users' who received and shared a large volume of these messages. Notably, chronic users disseminated an average of 14 forwarded messages during the week – approximately 5 times the number of messages sent by all participants in the study. This is reminiscent of research from other social media platforms (e.g., Twitter), where a small group of "supersharers" and "superconsumers" are responsible for the bulk of misinformation shared [26]. Given the potential influence of this group, further research is needed to understand: (i) the profile of chronic users, (ii) the reasons why they forward messages, and (iii) how their forwarding activities may influence outcomes during health crises.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Aside from chronic users, our study also found that 1 in 3 participants were 'receiving users' who had high exposure to forwarded COVID-19 content. Receiving users tended to be older (in line with misinformation studies on Facebook [27]) but were otherwise moderate in profile - whether in terms of COVID-19 fears or in religion (coming from a diverse religious background). Although this group did not spread forwarded messages themselves, their high exposure may nonetheless leave them susceptible to false beliefs. Again, we urge further research to understand how WhatsApp use within this group influences their health behaviors during crises.

Taken together, our taxonomy of WhatsApp user profiles provides a basis for targeted risk communication. Our findings suggest that public health agencies may need to reach out proactively to chronic and receiving users who handle the bulk of forwarded COVID-19 content on WhatsApp. One possible intervention may be to encourage these users to subscribe to official WhatsApp channels for updates (e.g., from the World Health Organization) [2], capitalizing on their pre-existing readiness to use the platform.

Finally, aside from forwarded content, we also found that WhatsApp users who discussed COVID-19 in their personal chats were more likely to think about COVID-19 through the day. As similar forms of rumination (involving frequent and persistent thoughts) have been linked to clinical depression [24], this finding may implicate COVID-19 chatter as a risk factor for poorer well-being [25]. Future studies should explore this possibility and the potential mechanisms involved.

Limitations

In reporting these findings, we note several limitations. First, we opted to study WhatsApp, the most widely used messenger application. At this juncture, it is unclear whether our results will generalize to other messenger applications (e.g., Facebook Messenger, Telegram).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Second, our recruitment strategy was limited by the nature of the pandemic. Owing to infectious disease protocols and the short time period when crisis-related communication was high [28], we relied on online recruitment. Further research is needed to examine whether our findings generalize to the broader population.

Finally, although the experience sampling method captured WhatsApp usage in participants' naturalistic settings, the method nonetheless required self-reports. Extending our findings, future studies will profit from having objective metrics of WhatsApp usage.

Conclusions

In conclusion, we applied the experience sampling method to capture for the first time "COVID-19 chatter" on WhatsApp. In total, we tracked 924 days of chatter *in situ,* revealing: (i) the sheer prevalence of WhatsApp usage; (ii) a typology of WhatsApp users; and (iii) a link between usage patterns and constant thoughts about the pandemic. These findings bring the field one step closer to digital phenotyping through WhatsApp, spurring future research on the role of messenger applications in health communication.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Acknowledgements

RREW, KJQH, JWEC contributed to the study design and collected the data. EYQT and SYE analyzed the data and wrote the first draft of the manuscript. JCJL and EMWT conceived of the study, had oversight of the overall project, and supervised both data analyses and manuscript preparation. All authors provided feedback on the final version of the manuscript. This research was funded by a grant awarded to JCJL from the NUS Centre for Trusted Internet and Community (Grant number: CTIC-RP-20-09). The first author's (EYQT) involvement was funded by a centre grant awarded to the Centre for Sleep and Cognition. This study was approved by the Yale-NUS College Ethics Review Committee.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

COVID-19: Coronavirus disease

GMM: Gaussian Mixed Models

BIC: Bayesian Information Criterion

ICL: Integrated Completed Likelihood Criterion

BLRT: Bootstrap Likelihood Test

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. Singh M. WhatsApp hits 2 billion users, up from 1.5 billion 2 years ago. TechCrunch. 12 February, 2020. http://techcrunch.com/2020/02/12/whatsapp-hits-2-billion-users-up-from-1-5-billion-2-years-ago/ [accessed Dec 1, 2020]

2. Liu JCJ, Tong EMW. The relation between official WhatsApp-distributed COVID-19 news exposure and psychological symptoms: Cross-sectional survey study. J Med Internet Res. 2020; 22. doi: 10.2196/22142

3. Kantar. COVID-19 barometer: Consumer attitudes, media habits and expectations. Kantar. 3 April 2020. https://www.kantar.com/Inspiration/Coronavirus/COVID-19-Barometer-Consumer-attitudes-media-habits-and-expectations [accessed Dec 1, 2020]

4. Newton C. WhatsApp puts new limits on the forwarding of viral messages. The Verge. 7 April 2020. https://www.theverge.com/2020/4/7/21211371/whatsapp-message-forwarding-limits-misinformation-coronavirus-india [accessed Dec 1, 2020]

5. Singh M. WhatsApp's new limit cuts virality of "highly forwarded" messages by 70%. TechCrunch. 27 April 2020. http://techcrunch.com/2020/04/27/whatsapps-new-limit-cuts-virality-of-highly-forwarded-messages-by-70/ [accessed Dec 1, 2020]

6. Bapaye JA, Bapaye HA. Demographic factors influencing the impact of coronavirus-related misinformation on WhatsApp: Cross-sectional questionnaire study. JMIR Public Health Surveill. 2020;7. PMID:33444152 doi: 10.2196/19858

7. Luk TT, Zhao S, Weng X, Wong JY-H, Wu YS, Ho SY, Lam TH, Wang MP. Exposure to health misinformation about COVID-19 and increased tobacco and alcohol use: a population-based survey in Hong Kong. Tob Control. 2020. PMID:32855353 doi: 10.1136/tobaccocontrol-2020-055960

8. Lwin MO, Lu J, Sheldenkar A, Schulz PJ, Shin W, Gupta R, Yang Y. Global Sentiments Surrounding the COVID-19 pandemic on twitter: Analysis of twitter trends. JMIR Public Health Surveill. 2020;6. PMID:32412418 doi: 10.2196/19447

9. Chen E, Lerman K, Ferrara E. Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. JMIR Public Health Surveill. 2020;6. PMID:32427106

10. Kumar VG, Gireesh Kumar V, Koovakkai D. Social media misinformation in disaster situations: A case of 2018 Kerala flood. Libr Prog, Int. 2019;39:289-296. doi: 10.5958/2320-317x.2019.00032.1

11. Simon T, Goldberg A, Leykin D, Adini B. Kidnapping WhatsApp – Rumors during the search and rescue operation of three kidnapped youth. Comput Hum Behav. 2016; 64:183–190. doi: 10.1016/j.chb.2016.06.058

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

12. Lottridge D, Bentley FR. Let's hate together: How people share news in messaging, social, and public networks. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; New York. New York, USA: Association for Computing Machinery; 2018. https://dl.acm.org/doi/10.1145/3173574.3173634 doi: 10.1145/3173574.3173634

13. Csikszentmihalyi M, Larson R. Validity and reliability of the experience-sampling method. In: Csikszentmihalyi M, editor. Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi. Dordrecht: Springer; 2014:35–54. doi: 10.1007/978-94-017-9088-8_3

14. Hektner JM, Schmidt JA, Csikszentmihalyi M. Experience sampling method: Measuring the Quality of Everyday Life. Thousand Oaks, California: SAGE; 2007. ISBN: 9781412925570

15. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science. 2018 Mar 9; 359:1146–1151. PMID:29590045

16. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther.1995 Mar; 33:335–343. PMID:7726811

17. Mesch GS, Schwirian KP. Vaccination hesitancy: Fear, trust, and exposure expectancy of an Ebola outbreak. Heliyon. 2019 Jul; 5. PMID:31367685

18. Scrucca L, Fop M, Murphy TB, Raftery AE. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. R J. 2016 Aug;8:289–317. PMID:27818791

19. Wade TD, Crosby RD, Martin NG. Use of latent profile analysis to identify eating disorder phenotypes in an adult Australian twin cohort. Arch Gen Psychiatry. 2006;63:1377-1384. doi: 10.1001/archpsyc.63.12.1377

20. Bolger N, Laurenceau J-P. Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research. New York, NY: Guilford Press; 2013. ISBN:97814625061

21. R Core Team. R: A language and environment for statistical computing, Version 3.5.0. R Foundation for Statistical Computing, Vienna, Austria; 2018.

22. IBM Corp. SPSS Statistics for Windows, Version 23.0. IBM Corp. Armonk, NY; 2015.

23. Therneau TM, Atkinson EJ, Others. An introduction to recursive partitioning using the RPART routines. Technical report Mayo Foundation; 1997. https://www.mayo.edu/research/documents/biostat-61pdf/doc-10026699 [accessed 21 Jun, 2021]

24. Moulds ML, Kandris E, Starr S, Wong ACM. The relationship between rumination, avoidance and depression in a non-clinical sample. Behav Res Ther. 2007; 45: 251–261. doi: 10.1016/j.brat.2006.03.003

medRxiv preprint doi: https://doi.org/10.1101/2020.09.29.20203646; this version posted October 13, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

25. Blabst N, Diefenbach S. WhatsApp and wellbeing: A study on WhatsApp usage, communication quality and stress. Proceedings of the 31st International BCS Human Computer Interaction Conference; 2017 3 - 6 July. Boston: Science Open; 2017. doi: 10.14236/ewic/hci2017.85

26. Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D. Fake news on Twitter during the 2016 U.S. presidential election. Science. 2019 Jan 25;363(6425):374–378. PMID:30679368

27. Guess A, Nagler J, Tucker J. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Sci Adv. 2019 Jan; 363:374-378. PMID:30662946

28. Nsoesie EO, Cesare N, Müller M, Ozonoff A. COVID-19 misinformation spread in eight countries: Exponential growth modeling study. J Med Internet Res. 2020 Dec 15; 22. PMID:33264102