Chapter 5
Emotional Responses Through COVID-19 in Singapore

Yingwei Yan, Wei Chien Benny Chin, Chan-Hoong Leong, Yi-Chen Wang, and Chen-Chieh Feng

5.1 Introduction

Since the first case of COVID-19 in Wuhan, China, was announced in early 2020, the virus has spread worldwide, resulting a global pandemic declared by the World Health Organization (WHO) on 11 March 2020. As a popular destination for visitors from China, Singapore was one of the first countries to be affected by the novel coronavirus with the first case of infection confirmed on 23 January 2020. While many countries adopted a wait-and-see approach to the less understood disease, Singapore was proactive in setting up measures to screen possible cases given the SARS experience in 2003. Singapore was hailed as a role model for the world in fighting against COVID-19 as it had successfully contained infection cases in the early stages while keeping everyday business as usual. Unfortunately, the success did not shield Singapore from the subsequent waves of COVID-19 infections, arising mainly from returning Singaporeans taking refuge from the pandemic outbreaks outside of China. This led Singapore into a lockdown, or what local authority called

Y. Yan
Key Lab of Guangdong for Utilization of Remote Sensing and Geographical Information System, Guangdong Open Laboratory of Geospatial Information Technology and Application, Guangzhou Institute of Geography, Guangzhou, China
e-mail: yanyingwei@u.nus.edu

Y. Yan · Y.-C. Wang · C.-C. Feng
National University of Singapore, Singapore, Singapore

W. C. B. Chin (✉)
Singapore University of Technology and Design, Singapore, Singapore
e-mail: benny_chin@sutd.edu.sg

C.-H. Leong
Singapore University of Social Sciences, Singapore, Singapore
e-mail: chanhoong@suss.edu.sg

S. Shaw and D. Sui (eds.), *Mapping COVID-19 in Space and Time*, Human Dynamics in Smart Cities, https://doi.org/10.1007/978-3-030-72808-3_5

the "circuit breaker" (CB), starting from 7 April 2020. During the CB period, non-essential services were restricted to work-from-home or complete shutdown, and all schools were moved to full online learning. While this mobility restriction evidently changed the usual activities, routines or livelihoods of the residents, the impact of CB on their emotions is less known.

Emotion (e.g., feeling) is an integral aspect of human experiences, the variation of which can be a direct consequence of crises and this in turn leads to further changes in the physical and mental dimensions of human dynamics (Seagal and Horne 2020; Shaw et al. 2016). Emotional traumas in societies as a result of a major calamity may lead to massive stress-related disorders (Ćosić et al. 2020). For COVID-19 in particular, a WHO technical guidance notes the main psychological impact to date being "elevated rates of stress or anxiety" (WHO 2020). Unlike other crisis, such as terrorist attacks and hurricanes, COVID-19 is intangible and can cause long-term uncertainties that lead to fear, loneliness, distress reactions, and mental health disorders (Aslam et al. 2020; Ćosić et al. 2020).

Recently, sentiment analysis of social media data in the context of COVID-19 pandemic has attracted attention from the research community (Barkur et al. 2020; Li et al. 2020; Lwin et al. 2020). Indeed, social media data can be harnessed to understand the responses of residents to crisis events because the data are both cost-effective to collect, and their richness in volume and spatiotemporal coverage is unrivalled against traditional data sources (Wang and Ye 2018; Yan et al. 2017; Yan et al. 2020b). With respect to human dynamics, compliance to evacuation procedures and communication behaviors of residents during disaster and crisis events have been explored using social media (Kim et al. 2017; Martín et al. 2017; Takahashi et al. 2015). These studies tend to focus on the physical (e.g., evacuation actions) and mental dimension (e.g., perspectives) of human dynamics (Seagal and Horne 2020). In terms of the emotional dimension of human dynamics, sentiment analysis of Tweets posted in crisis events has received considerable attention in recent years. These research have contributed to a better understanding of the online community dynamics in response to terrorist attack (Shaikh et al. 2017), the dynamics of social networking (Neppalli et al. 2017), the sentiments towards the refugee crisis (Öztürk and Ayvaz 2018), and the post-earthquake sentiment variations in tourism destinations (Yan et al. 2020a).

Despite the value of social media sentiment analysis in understanding emotional responses to crisis, existing studies have seldom investigated the changes of sentiments, which have the potential to shed lights on how people react when transiting from normal to crisis situations. This study performed a sentiment analysis of Tweets posted in Singapore and aimed to reveal the temporal dynamics of the online communities' emotions in the context of COVID-19 global pandemic. In addition, Singapore is a country with a high density of population (7866 km^{-2}) and multi-ethnics and multi-lingual environment (DSS 2020). Singapore resident population composed of ethnic Chinese (74.4%), Malays (13.4%), Indian (9.0%), and other minority races (3.2%). Furthermore, the citizens and permanent residents proportion composed about 70.6% of the total population, leaving 29.4% of the population as foreign employees, dependents, or students with long-term residential passes. As such, it is crucial to examine the content of all Tweets using a multicultural lens. Therefore, this research further analyzed the emotional differences of Twitter users in different ethnicities (speaking different languages) of Singapore during the pandemic.

5.2 Material and Methods

The workflow of this study is illustrated in Fig. 5.1. It started with Tweets collection, followed by six Tweets pre-processing steps, including the language detection of the retrieved Tweets, translation of those non-English Tweets into English, tokenization, stop words removal, and lemmatization of the Tweets, and removal of the Tweets from news media. Lastly, sentiment analyses were performed to reveal both the overall sentiment dynamics and the language-based sentiment dynamics. Details are introduced in Sects. 5.2.1, 5.2.2 and 5.2.3.

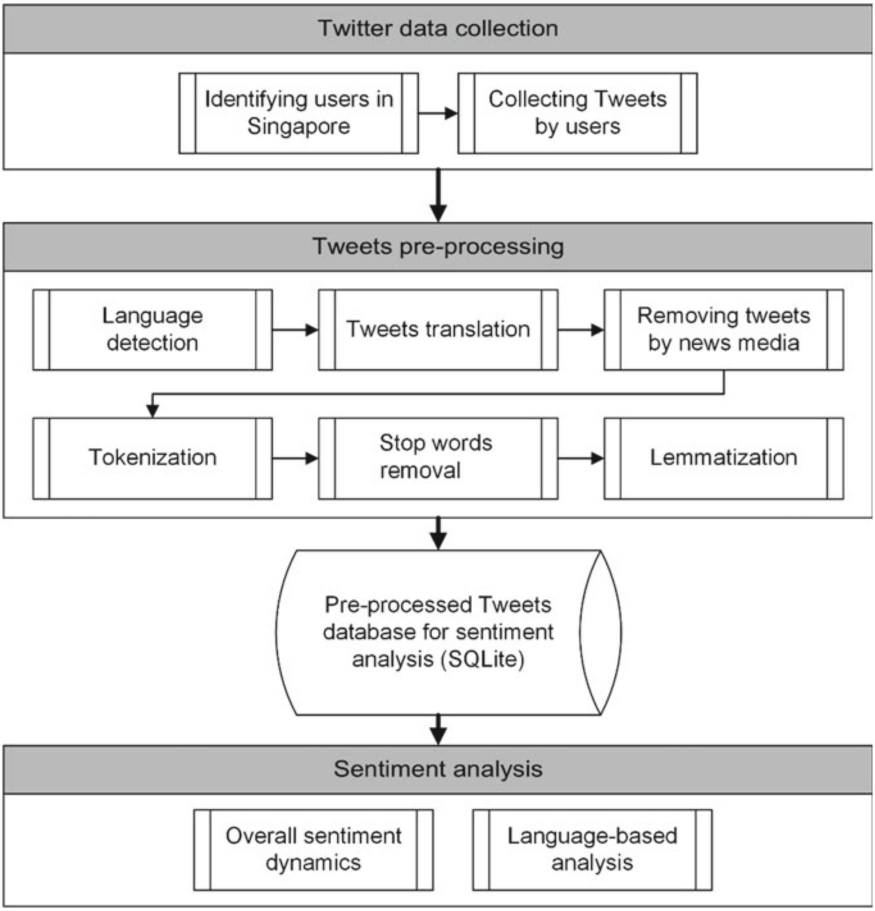

Fig. 5.1 Workflow of this study

5.2.1 Twitter Data Collection

Tweets related to COVID-19 were collected via TAGS (https://tags.hawksey.info/), a free Google Sheets template allowing us to set up and run automated collection of search results from Twitter. The geocode parameter of the Standard Search Application Programming Interface (API) of Twitter was used to retrieve Tweets that were located within Singapore. The Standard Search API takes location information preferentially from the Geotagging API, but will fall back to users' Twitter profile.

Real-time Tweets with their texts and hashtags containing the keywords of "COVID" or "coronavirus" (i.e., COVID-19-related Tweets) were collected between 16 April 2020 and 31 May 2020. After removing the non-individual (e.g. news media) accounts, a total of 12,067 users were found. Because COVID-19 pandemic started from January 2020 and Singapore has imposed semi-lockdown measures during the period of data collection, we assumed that these users resided in Singapore from January to May of 2020. After that, all the Tweets (regardless whether a Tweet was related to COVID-19 and regardless whether a Tweet was geotagged) posted by the same group of Twitter users who posted the COVID-19-related Tweets were further collected through Twitter's official API based on the user IDs. A total of 18,535,620 Tweets were collected. We retrieved the Tweets from the users posted in the first five months of 2020 for the main analysis, including (a) January and February: the nascent stage of COVID-19 outbreak when the disease was first imported to Singapore and transmitted locally (1,622,470 Tweets), (b) March: the stage where the COVID-19 outbreak escalated in Singapore (1,227,370 Tweets), (c) April and May: when the "Circuit Breaker" semi-lockdown measures were imposed (3,572,745 Tweets). We also collected the Tweets of the same months posted by the users in the preceding year (2019) to serve as the baseline for comparison (1,368,162 Tweets). All these Tweets were composed by text in different languages. Tweets composed by non-words objects (e.g., static images, GIFs, only URL links, only emoji characters or special characters) were excluded.

5.2.2 Tweets Pre-processing

After the retrieval of the Tweets, Google Sheets were used for the language detection of the Tweets and for the Tweets translation. Following Yan et al. (2020b), Tweets were further pre-processed by tokenization (cohesive strings from the Tweets were split up into single words or "tokens"), stop words removal (frequently occurring short-function words without valuable content such as "of" and "to" were removed to reduce noise), and lemmatization (the words were converted to their root form); these pre-processing steps reduced the semantic dimension of the raw Tweets for creating word vectors. All pre-processed Tweets were stored in a SQLite (3.22.0) database file for the subsequent sentiment analysis.

5.2.3 Sentiment Analysis

The sentiment analysis of this study was performed based on Plutchik (2001), which postulates eight primary bipolar emotions: joy (feeling happy) versus sadness (feeling sad); anger (feeling angry) versus fear (feeling of being afraid); trust (stronger admiration and weaker acceptance) versus disgust (feeling something is wrong or nasty); and surprise (being unprepared for something) versus anticipation (looking forward positively to something). This approach has been adopted extensively. Particularly, the National Research Council (NRC) Canada presented a large Word-Emotion Association Lexicon created through a massive online annotation project based on Plutchik's eight basic emotions (Mohammad and Turney 2013).

The NRC Word-Emotion Association Lexicon has been implemented in the Syuzhet package of R (https://www.rdocumentation.org/packages/syuzhet/versions/1.0.4). The Syuzhet package enabled us to generate emotional valence value for each Tweet. It also enabled us to identify the words in each Tweet that indicated eight basic types of human emotions (i.e., anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and the words in each Tweet that indicated positive and negative emotions. Both the overall sentiment analysis (without separating the Tweets by languages) and sentiment analysis by languages (the Tweets were separated by languages) were performed. We grouped the Tweets by days and analyzed the daily emotional valence trends in the following sections. In addition, Augmented Dickey-Fuller test (Kwiatkowski et al. 1992) were used to check if each of the emotional valence daily trends was stationary.

Hierarchical clustering was performed to investigate the clustering and hierarchical relationships among the emotional responses of the Tweets in different languages and emotion types. The tool used to perform the hierarchical clustering was a function (AgglomerativeClustering) from the Scikit-learn package of Python. To understand the similarity and differences between languages and emotions, the daily emotional dynamic for each emotion of a language was used as a data point, with each day as a vector. In data mining and statistics fields, agglomerative hierarchical clustering is a method to group similar objects (in this study, language) based on the observation patterns (i.e., the trends of the eight emotions) using a bottom-up approach. It starts with each object as a cluster, followed by calculating the differences between each pair of clusters, then groups the most similar clusters (i.e., with the smallest differences); it repeats the steps of difference calculation and cluster grouping. In this study, the trend of each emotion was used as the patterns for comparing between languages. In other words, if two languages were grouped as a cluster, it indicated that the two languages had the same trend of a specific emotion during the five-month period.

5.3 Results and Discussion

5.3.1 Overall Sentiment During COVID-19

The emotional valence in the first five months of 2019 and 2020 (Fig. 5.2) shows similar percentage of Tweets exhibiting zero emotional valence. The proportion of Tweets with negative emotional valence is slightly higher in 2020 (35%) than in 2019 (29.4%). In both 2019 and 2020, more than half of the Tweets have a positive emotional valence. Aggregating the emotional valence leads to the daily average of emotional valence of 2019 and of 2020 (Fig. 5.2) both in the positive value range, suggesting a weak positive emotion throughout the study period. In addition, the emotional valence trends of the two years were stationary (Table 5.1). Nonetheless, due to the proportion of negative Tweets in 2020 is larger than in 2019, the overall trends in 2020 daily average values are lower than in 2019. To confirm that the overall emotional valence values of the two years are indeed different, an ANOVA test was conducted. The test result ($F(1{,}301) = 506.609$, $p < 0.001$) confirmed that the drop of overall emotional valence during COVID-19 period was statistically significant.

Narrowing in on the changes of emotional valence over the five-month period in 2019 and in 2020 (Fig. 5.3), the two-time series show strikingly high emotional valence on the 1 January (Day 1). By gleaning through the Tweets of 1 January, it was confirmed that the high emotional valence was reasonable as a large number of "new year greetings" were tweeted. After the new year of 2020, the overall emotional valence dipped twice, one towards the end of January (26 January, Day 26) and the other right after the start of February (8 February, Day 39). It became more positive in the subsequent month before dipping again near the end of March 2020 (29 March, Day 89). Conversely, the emotion valence did not drop but rather improved somewhat in early April at the onset of CB and remained steady with minor fluctuation around

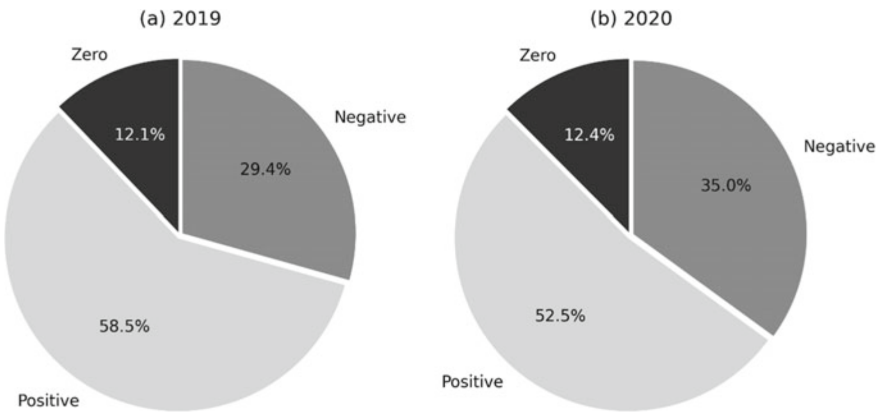

Fig. 5.2 Proportion of positive, negative, and zero emotional valence for **a** 2019 and **b** 2020

Table 5.1 Stationarity test of the emotional valence and the eight emotions for 2019 and 2020. Both trends and daily differences (differencing) were tested

Year	2019		2020	
Variable	Trend	Daily changes	Trend	Daily changes
EV	−4.72*	−7.69*	−4.99*	−11.57*
Trust	−2.26	−5.59*	−2.67	−6.54*
Anticipation	−8.38*	−8.17*	−8.66*	−12.51*
Joy	−4.21*	−5.21*	−1.68	−4.34*
Fear	−8.03*	−6.89*	−3.77*	−11.32*
Sadness	−7.48*	−6.27*	−6.81*	−7.80*
Surprise	−10.08*	−7.62*	−6.76*	−9.62*
Anger	−7.86*	−7.48*	−4.01*	−11.00*
Disgust	−3.99*	−6.05*	−7.50*	−10.61*

*$p < 0.01$; EV: emotional valence

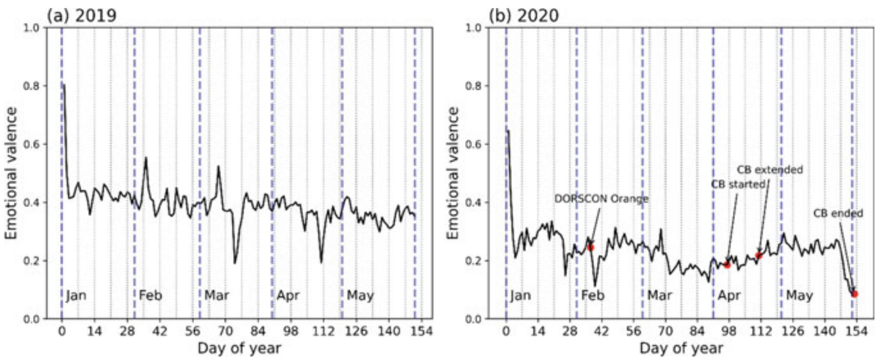

Fig. 5.3 Mean emotional valence by day-of-year of **a** 2019 and **b** 2020

0.25 until the end of CB (31 May, Day 152), when it dipped to the lowest point of the five months studied.

Breaking down the overall sentiment according to the eight basic types of emotions in both years for the five-month period exhibit similar temporal patterns (Fig. 5.4). Except for the 2019 trend for trust and the 2020 trends for trust and joy, all the other trends were stationary at 1% significance level (Table 5.1). Additionally, all the daily differences trends (i.e., differencing) were stationary (Table 5.1). The stationarity test results imply that, for the different types of emotions, although there existed certain emotional unstableness over time, there existed no significant emotional fluctuation day to day. There is a higher proportion of positive emotion types in both years, these include trust, anticipation, and joy. Overall, there is proportionately less negative emotion types, with more messages on fear and sadness than anger and disgust being observed in both years. In terms of the ordering by emotion types, there is no clear

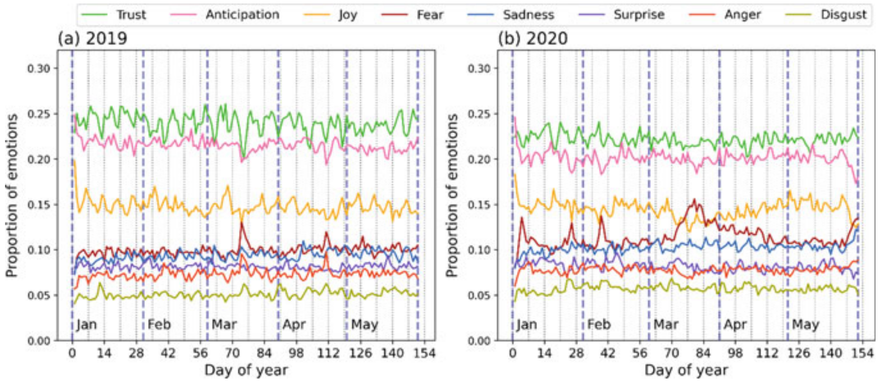

Fig. 5.4 Emotion proportion by types during COVID-19 of **a** 2019 and **b** 2020

distinction between the temporal patterns of 2019 and 2020 except that the negative types have a slightly higher proportion in 2020 as suggested by the overall sentiment analysis above. One peculiar deviation is the visible rise on the proportion of fear for about 10 days in mid-to-late-March, 2020, which may be related to the rising confirmed cases that prompted the implementation of CB in early April. In addition, it is observed that towards the end of CB, there was a visible drop of the proportion of anticipation and joy, along with a visible rise of the proportion of fear and sadness. Beside these increasing and decreasing trends, the proportion of fear had three visible one-day peaks during January (4 January and 26 January, Day 4 and Day 26) and February (8 February, Day 39) 2020, each had a corresponding drop of proportions in trust, anticipation, and joy.

5.3.2 Emotional Responses to COVID-19 in Singapore

In general, the results show that Singapore had a high level of trust and positive attitude during the 2020 pandemic (slightly less positive compared with the figures for 2019) (Figs. 5.2, 5.3 and 5.4). Singapore government's response to the crisis was swift. Within days of the Wuhan lockdown on 23 January (Day 23, the fear emotion increased after the Wuhan lockdown, see Fig. 5.4), free face masks were provided, social distancing was required, temperature checks were implemented in schools and work places, and hand sanitizers were provided where necessary. In the meanwhile, a network of about 900 designated clinics called Public Health Preparedness Clinics (PHPCs) was activated to help the authorities better detect and manage COVID-19 outbreak as the first line of treatment (ST 2020). Citizens and permanent residents would pay an affordable flat rate fee for consultation and treatment at PHPCs. This encouraged people to consult doctors and allowed the government to have enhanced surveillance systems and extensive testing operations to identify cases at their early diagnoses. Anyone infected was sent to hospitals or community isolation facilities. People entering Singapore must serve a quarantine during which their health statuses were monitored; breaking quarantine will be a criminal offense.

Additionally, the government was actively identifying clusters and linkages of COVID-19 infections, tracking where every patient had been and identify everyone they had interacted with since becoming infected in order to break the chain of transmission. Furthermore, residents were encouraged to download a tracing mobile application for the authorities to check the places they had visited and who they had come in contact. With these measures in place, people in Singapore would know where the exposed risks were, so that they could adapt their behaviors and feel confident that the virus was under control. As a result of these measures, most of the workplaces, schools, resultants, cafes, and bars stayed open until the start of the CB.

Following several cases of infections without any links to previous cases or travel history to mainland China, Singapore raised its Disease Outbreak Response System Condition (DORSCON) level to Orange from Yellow on 7 February, 2020 (Day 38) (CNA 2020). This resulted in the fear emotion of the general public (Fig. 5.4) and panic buying and stock piling were reported. However, the government soon urged calm and assured that there was "no risk of us running a shortage of essential food or household items" (BT 2020).

Facing the wave of outbreak before the CB in early April, the proportion of fear emotion had increased and the proportion of joy emotion decreased (Fig. 5.4). The panic lasted only a short time, as the emotional valence curve of fear descended and the emotional valence curve of joy ascended very soon after. The level of community transmission (COVID-19 spreads in such a way that the source of origin of the infection is unknown) was still low during the CB. The dramatic increase of the infections mainly occurred within the foreign workers dormitories, but they had little interactions with the general public (Today 2020). The mortality rate in Singapore remained low (Worldometer 2020).

5.3.3 Sentiment Analysis by the Top 10 Languages

The multi-lingual nature of Singapore was evidently captured in Twitter messages during the study period, with a total of 105 languages found in the Tweets being analyzed in this study. The major ethnic groups in Singapore residents are Chinese, Malay, and Indian (DSS 2020) and their corresponding languages are, respectively, Chinese language, Malay language, and Tamil or Hindi languages. Using 18,000 Tweets as the threshold, the 10 most frequently used languages were identified (Table 5.2). The four languages of the major ethnic groups were contained in the top 10 languages. English was most frequently used, followed by Malay. The frequencies corresponded well to the fact that English and Malay are the main language and the national language of Singapore, respectively. The next language used is not Chinese even though ethnic Chinese accounts for 74% of Singaporeans as of 2019 (DSS 2020). Rather, the number of Tweets in Chinese is in the seventh place, suggesting the

Table 5.2 The top 10 languages used in COVID-19 related Tweets in Singapore

Rank	Language	Number of Tweets	Percentage (%)
1	English	5,433,782	84.51
2	Malay	259,485	4.04
3	Japanese	170,802	2.66
4	Indonesian	151,604	2.36
5	Korean	84,960	1.32
6	Filipino	61,019	0.95
7	Chinese	48,855	0.76
8	Tamil	40,946	0.64
9	Thai	30,894	0.48
10	Hindi	18,578	0.29
	Other languages	128,524	1.99
	Total	6,429,449	100.00

possibility that majority of ethnic Chinese in Singapore tweeted in English. Coupled with the fact that the use of Chinese trails Japanese, Indonesian, Korean, and Filipino, it is plausible that non-English Tweets corresponded to non-local communities. The Tweets distribution of the top 10 languages may reflect the penetration rates of Twitter in different ethnicities.

Focusing on the top 10 languages used in the Tweets during the COVID-19 period, Fig. 5.5 shows the daily average emotional valence over the study period. All except Japanese and Tamil languages were stationary at 1% significance level (Table 5.3); the daily differences trends were all stationary (Supplementary Table 5.1). The stationarity test results suggested that, for the Twitter users in different language groups, despite certain emotional unstableness over time, there existed no significant emotional fluctuation day to day. Because a large percentage of the Tweets were in English, the trend of emotional valence values for English (Fig. 5.5b) was similar to the overall trend (Fig. 5.5a), both of which had relatively low uncertainty. The fluctuation patterns of the other nine languages (Fig. 5.5c–k) were different from each other, suggesting that the people of different ethnicities were experiencing different sentiments during the first five months of COVID-19. Alternatively, sentiments reflected in English, Japanese, and Chinese Tweets, were relatively stable compared to the sentiments reflected in the other seven languages, which not only fluctuated over time but also exhibited different patterns in the five months studied. Both Japanese and Chinese had a moderate fluctuation in January, and the fluctuations stabilized from February to the end of May. The sentiments in Malay and Indonesian, the two languages that were highly cognate (Feng and Mark 2017), had moderate fluctuations but the former had a larger fluctuation in the first two months and a smaller fluctuation starting from March, whereas the latter had a larger fluctuation in the first two months, following by a small fluctuation in March, and a moderate to large fluctuation in April and May. This further reinforced the idea that the users of the two languages were likely from separate, local and non-local communities.

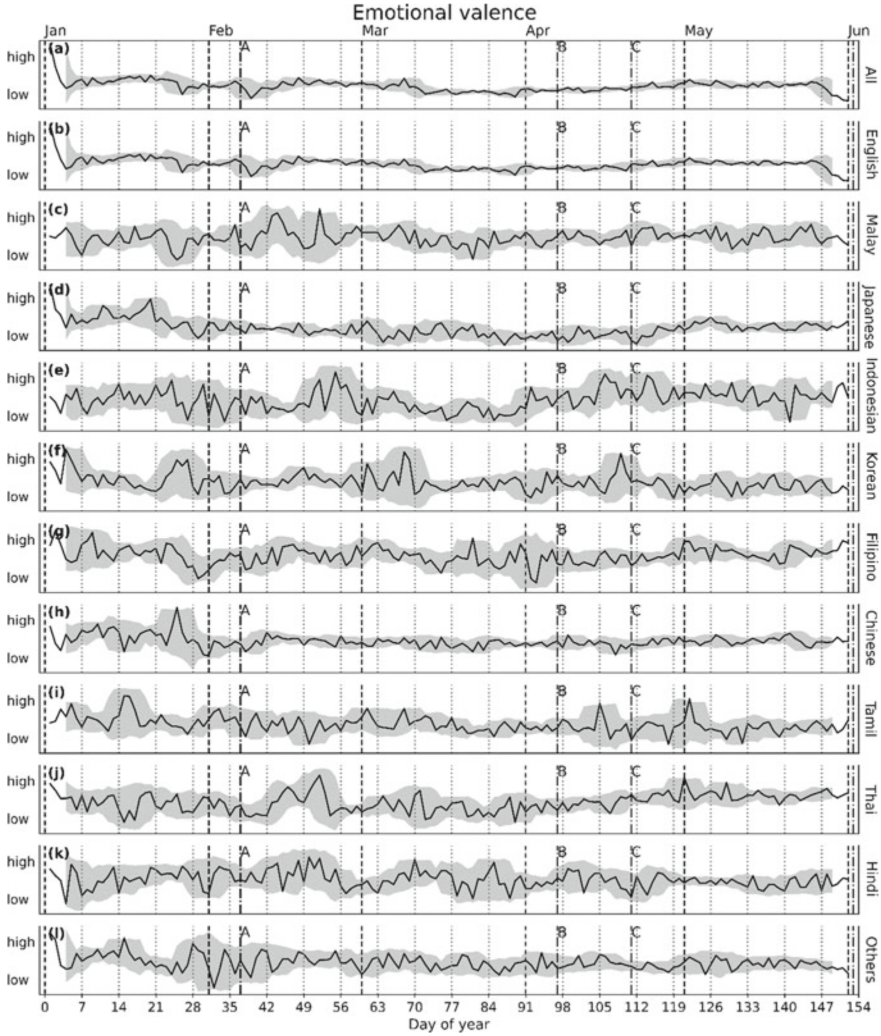

Fig. 5.5 Daily average of the emotional valence values for 2020 by the top 10 languages used. The black lines indicate the daily average value, and the gray area shows the trends of ±3 days moving average with ±2 standard deviation. The vertical dash-dot lines indicate the three key dates: **a** DORSCON level to Orange, **b** CB measures started, **c** the announcement of CB being extended to 1 June

The remaining five languages showed persistent fluctuations throughout the study period. In general, most languages exhibited a significant drop of emotional valence values in the last two weeks of January, some earlier, including Japanese (decrease starting from Day 20 to Day 29) and Thai (dropped to a valley on Day 15 to Day 17), and some later, including Indonesian, Korean, Filipino, Chinese (these four dropped

Table 5.3 The stationarity test for the trends of emotional valence and the eight emotions. The languages were labelled with the corresponding language code (ISO 639–3)

Lang	EV	Trust	Anticip	Joy	Fear	Sad	Surpr	Anger	Disgust
ENG	−5.12*	−2.76	−8.50*	−4.86*	−3.88*	−7.06*	−7.38*	−3.88*	−7.44*
MSA	−8.23*	−8.63*	−4.72*	−9.12*	−8.20*	−8.59*	−8.01*	−8.95*	−7.47*
JPN	−2.13	−1.73	−2.25	−1.83	−2.26	−2.05	−6.45*	−10.75*	−4.64*
IND	−3.48*	−5.47*	−7.38*	−2.45	−7.56*	−9.77*	−4.79*	−8.56*	−9.26*
KOR	−8.20*	−10.06*	−11.63*	−9.05*	−7.02*	−9.11*	−10.16*	−6.06*	−9.47*
FIL	−5.08*	−11.83*	−8.15*	−11.23*	−7.79*	−11.48*	−11.60*	−10.00*	−7.03*
ZHO	−8.18*	−9.12*	−10.87*	−2.29	−2.57	−9.32*	−9.31*	−5.77*	−7.96*
TAM	−2.52	−10.50*	−11.02*	−10.43*	−4.12*	−4.30*	−10.54*	−12.14*	−6.13*
THA	−7.10*	−12.24*	−7.95*	−6.07*	−1.13	−7.43*	−3.11	−10.88*	−3.11
HIN	−7.01*	−12.82*	−5.83*	−11.26*	−4.94*	−11.35*	−12.20*	−11.70*	−10.48*

*$p < 0.01$; EV: emotional valence; Anticip.: Anticipation; Sad.: Sadness; Surpr.: Surprise

Supplementary Table 5.1 Supplementary stationarity test for the daily changes of the language-based emotional valences

Lang	EV	Trust	Anticip	Joy	Fear	Sad	Surpr	Anger	Disgust
ENG	−11.19*	−6.44*	−10.89*	−7.73*	−11.60*	−7.48*	−9.96*	−11.27*	−7.81*
MSA	−7.85*	−6.42*	−11.27*	−5.36*	−7.46*	−6.13*	−7.03*	−7.67*	−7.17*
JPN	−7.59*	−8.80*	−8.57*	−8.60*	−7.71*	−8.80*	−6.28*	−7.85*	−8.95*
IND	−5.34*	−7.40*	−7.29*	−4.87*	−9.17*	−10.38*	−6.49*	−7.28*	−7.19*
KOR	−8.61*	−7.33*	−6.53*	−7.16*	−6.99*	−6.97*	−7.84*	−9.72*	−7.12*
FIL	−7.31*	−8.04*	−7.04*	−7.63*	−7.56*	−9.18*	−7.83*	−9.21*	−7.95*
ZHO	−6.96*	−6.97*	−4.62*	−5.24*	−9.25*	−6.02*	−7.93*	−5.37*	−8.27*
TAM	−6.70*	−9.29*	−9.57*	−7.75*	−8.66*	−12.76*	−9.27*	−6.15*	−7.43*
THA	−9.24*	−10.22*	−8.27*	−6.39*	−6.06*	−10.93*	−9.95*	−6.94*	−9.96*
HIN	−9.77*	−8.47*	−10.81*	−6.83*	−6.36*	−6.40*	−6.49*	−8.70*	−6.73*

*$p < 0.01$; Anticip.: Anticipation; Sad.: Sadness; Surpr.: Surprise

after the peak of Chinese New Year between Day 21 and Day 28), Malay (which had a valley on Day 25), and Hindi (dropped to a valley on Day 30 and Day 31).

The overall emotion pattern (Fig. 5.6b) reveals that more positive emotion types are more dominant given that they account for higher proportions of the Tweets, and the Tweets related to trust in every language is generally above 20%. Nevertheless, for the proportion of emotion by types, different languages have exhibited different patterns (Fig. 5.6). Japanese is unique in that the anticipation is always slightly higher than trust, while in other languages either trust is mostly higher than anticipation or their proportions are indistinguishable (Fig. 5.6). Tweets in five languages show clear "gaps" (generally 5–10%) between one or more positive emotion types and other emotion types. For the Tweets in English and Japanese, trust and anticipation are the

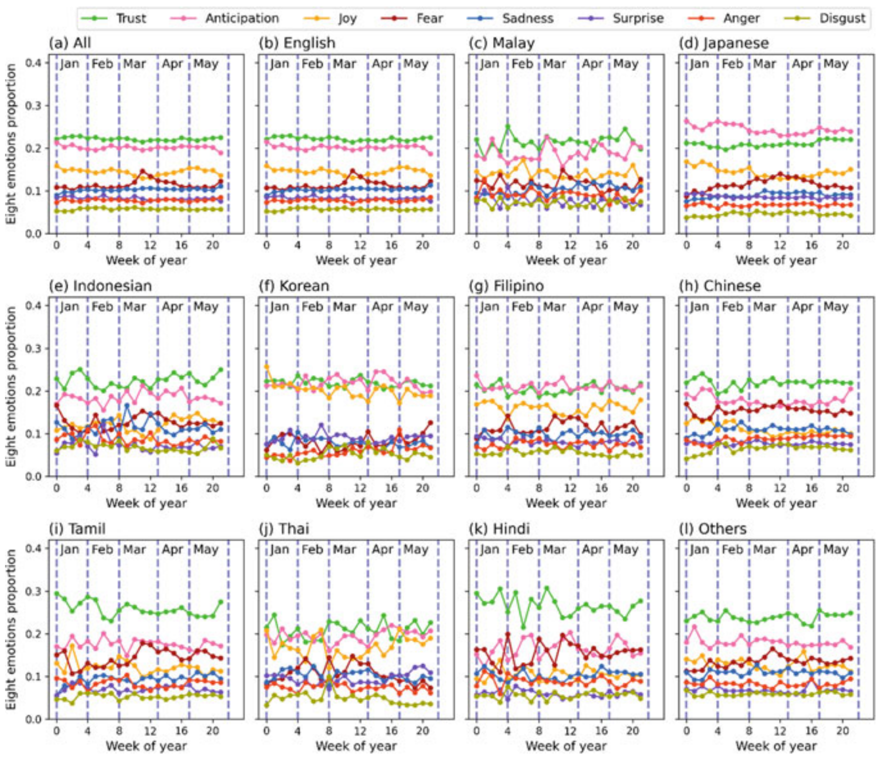

Fig. 5.6 Patterns of the weekly average proportion of the eight emotions by **a** all languages; **b**–**k** the top 10 languages analyzed; and **l** all other languages

two emotion types separated from the rest emotion types. For the Tweets in Korean, trust, anticipation, and joy are clearly more dominant than the rest emotion types. For Tamil and Hindi, trust stands out from the rest emotion types. The remaining languages exhibit mixtures of emotions although they may at some point exhibit separations between some positive emotion types with other types (Fig. 5.6). The most intriguing temporal pattern of individual emotion types across all languages is the rise of the proportions of Tweets related to fear around mid-March, except for Korean. The proportion of Tweets in Korean did exhibit a rise in March, but it is near the end of March. Most of the trends in Fig. 5.6 were stationary at 1% significance level (Table 5.3). All of the corresponding daily differences were also tested, and were stationary at 1% significance level (Supplementary Table 5.1). The stationarity test results suggest that, for the Twitter users in different language groups and for the different types of emotions, although there existed certain emotional unstableness over time, there existed no significant emotional fluctuation day to day.

To explore the similarities and differences of emotional responses among languages, each of the eight daily average emotion proportions was analyzed using hierarchical clustering. The results are shown in dendrograms (Fig. 5.7). Some

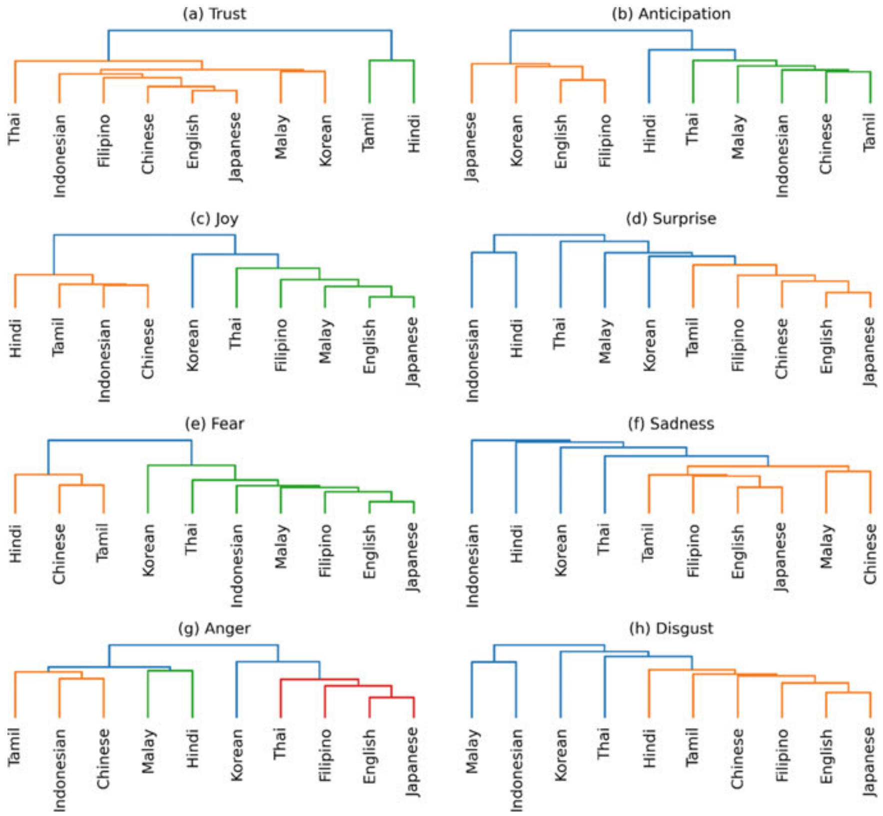

Fig. 5.7 The hierarchical clustering dendrograms for the daily average of the eight emotion proportions in Tweets

emotions exhibited similar clustering patterns such as (1) joy and fear as a group, and (2) surprise, sadness and disgust as another group, demonstrating the intra-group language similarity. The main difference between joy and fear is on Indonesian; it belongs to the left cluster (orange) in joy, but to the right cluster (green) in fear. Additionally, surprise, sadness, and disgust have a similar clustering pattern. For these three emotions, Tamil, Filipino, English, and Japanese have similar trends; they are among the lowest levels in all three emotions. Indonesian joins the tree in the last level for most of the emotions, indicating that it is more different from the other nine languages. Hindi also differs from the other eight languages for surprise and sadness emotions. English and Japanese usually appear at the lowest level, indicating their similarity in the trends in most emotions (except for anticipation, in which they appear in third lowest level).

5.3.4 COVID-19 Outbreaks in a Multi-ethnics and Multi-lingual Environment

Singapore has a high population density and many diverse ethnic groups. It is a norm among Singapore residents to live next to others who speak different languages. This cultural plurality is also evident in the virtual community of social media users (e.g., Twitter). From our data—the Tweets from people who live in Singapore—the Tweets in English accounted for 84.5%, the Tweets in nine other major languages represent 14%, and the Tweets in all other languages 2% (Table 5.2). Despite the high population density, the people in different language groups reported different emotional dynamics over the first five months in 2020 (Figs. 5.5 and 5.6), meaning that the people of different ethnics experienced the impacts of COVID-19 differently, even though they stayed next to each other. There are several possible reasons for the differences. First, the people of different ethnics received information from different types of media. Different sources of news may have different time-lag effects, causing the temporal differences in the emotional variations. Second, people from different ethnic community may embrace different attitudes in coping with the pandemic, hence experienced the outbreak differently and expressed their feelings using varying emotions. Different languages may express the same event but emphasises different sentiments. Some may describe more terms in "fear", while others may use "angry". Third, the non-resident population may be concerned about their families or relatives residing in their home and other countries that encountered different COVID-19 challenges compared to the situation in Singapore. Fourth, despite the highly diverse residential environment—the result of the social mixing programs in Singapore—people in the same ethnic group tend to gather around certain locations, and the clusters might appear in different locations in different time, leading to the differences in emotional dynamics.

The hierarchical clustering results (Fig. 5.7) showed the similarities and differences of languages in different emotion types. Some languages demonstrated similarities across multiple emotions, such as English and Japanese. Some emotions have similar language clustering patterns, such as joy and fear as a group, surprise, sadness, and disgust as another group, implying that the intra-group dynamics of the emotions were related. The inter-group differences indicated the underlying factors that stimulated the emotion dynamics were different. Those factors could be related to the four causes mentioned above or to the nature of the languages, which require further in-depth investigations.

While some similar clustering patterns exist among languages and emotions, the results showed more differences than similarities among the languages' dynamic patterns in terms of the emotions. As such, the multi-lingual and multi-ethnic diversity perspective should be considered when investigating the impacts of the disease outbreak on the people in Singapore. Granted, existing efforts in fighting against

COVID-19 have considered this perspective by, for example, delivering useful information in multiple languages and dialects. The sentiment analysis results, nonetheless, showed that considering strategies amenable to the needs of different communities to minimize negative sentiments is still needed. One strategy may be to ramp up counselling, social, and other related services in a way that is sensitive to the difficulties facing specific communities so as to reduce the chances of further worsening of their mental wellbeing. The results and the methodological framework presented here demonstrates a means that enables the identification of the sentiment changes by communities or groups in a crisis event, such as the COVID-19 pandemic, and forms the basis from which resource provisions can be more precisely rendered and more adaptive to the needs on the ground.

Beyond informing policies, the sentiment analysis presented here suggested the importance the need to consider language as a factor in the analytical framework of crisis management, especially in regions with diverse ethnicities and possible diverse spoken languages, such as Singapore. As the diversity of ethnics in population is usually neglected in previous studies because of the difficulties in data collections and data processing, it is challenging if not possible to compare the expression of same emotion from two people from different ethnicities. Using Singapore as a case study and the languages of Tweets as a surrogate to represent different ethnic groups, it is demonstrated that considering ethnicity is indispensable in such analyses. Practically, this study also showcased one possible way to incorporate the notion of ethnicity in the understanding of the dynamic responses to a pandemic event.

5.4 Conclusions and Future Works

Social media platforms offer residents venues to express freely their feelings and opinions. In a crisis situation, mining the reactions of people on social media could be an efficient way to evaluate the effectiveness of the intervention measures and the decisions made by the related authorities. In this study, sentiment analysis was performed with Tweets posted in Singapore in 2019 and 2020, revealing the variations in the emotion dimension of human dynamics in the context of COVID-19. The study found that people in Singapore generally had a high level of trust and positive attitude facing the crisis, which could be a result of the government's smart crisis response. The study also revealed the differences in the emotions of the people in different language groups, highlighting the uniqueness in the emotional reactions of Singapore as a multi-ethnic and multi-lingual nation.

The significance of this study is at least threefold. First, understanding the social emotional impacts of the first and largest global pandemic event in the twenty-first century allows us to better understand how people cope with such events and how it affects people's well-being and social resilience. Second, the insights generated in this study may also facilitate the authorities' evaluation of the impact of their anti-epidemic measures on people's mental states. Third, this study contributes to the literature on the policies that assist people to address possible emotional reaction to a pandemic.

For future works, first, this study focused on only the Tweets' text contents and the Tweets' visual contents (i.e., pictures and memes) were not included in the analysis. It would be interesting to take the visual contents into consideration in future works. Second, topic modeling such as Latent Dirichlet Allocation can be coupled with the sentiment analysis to discover more hidden details and provide more contexts for better explaining the sentimental patterns. Third, emotion is a complex and mixed product of human feelings, but this study models the sentiment of Tweets using only the eight basic emotions. Future studies are suggested to explore more on the emotions, e.g., the stronger or weaker emotions and the integrated feelings. In addition, as a data source, the sampling issue and bias of social media related to demographics, culture, user behavior, and even its own API (at most one percent sample of all Tweets are retrievable) should be addressed in future research.

Acknowledgements We thank the support from the GDAS' Project of Science and Technology Development (grant number: 2020GDASYL-20200103005), the National Natural Science Foundation of China (grant number: 41901330), and the Singapore University of Technology and Design (grant number: Cities Sector: PIE-SGP-CTRS-1803). The funding sources have no involvement in study design, collection, analysis and interpretation of data, writing of the report, and the outlet of the manuscript.

References

Aslam, F., Awan, T. M., Syed, J. H., Kashif, A., & Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanities and Social Sciences Communications, 7,* 23. https://doi.org/10.1057/s41599-020-0523-3.

Barkur, G., & Vibha, Kamath, G. B. (2020). Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: Evidence from India. *Asian Journal of Psychiatry, 51*, 102089. https://doi.org/10.1016/j.ajp.2020.102089.

BT. (2020). Panic buying hits Singapore after virus alert raised. Retrieved August 13, 2020, from https://www.businesstimes.com.sg/government-economy/panic-buying-hits-singapore-after-virus-alert-raised.

CAN. (2020). Coronavirus outbreak: Singapore raises DORSCON level to Orange; schools to suspend inter-school, external activities. Retrieved August 7, 2020, from https://www.channelnewsasia.com/news/singapore/wuhan-coronavirus-dorscon-orange-singapore-risk-assessment-12405180.

Ćosić, K., Popović, S., Šarlija, M., & Kesedžić, I. (2020). Impact of human disasters and Covid-19 pandemic on mental health: Potential of digital psychiatry. *Psychiatria Danubina, 32,* 25–31. https://doi.org/10.24869/psyd.2020.25.

DSS. (2020). Population and population structure: Singapore residents by age group, ethnic group and sex, end June, annual. Retrieved August 20, 2020, from https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/latest-data.

Feng, C.-C., & Mark, D. M. (2017). Cross-linguistic research on landscape categories using GEOnet names server data: A case study for Indonesia and Malaysia. *The Professional Geographer, 69,* 567–578. https://doi.org/10.1080/00330124.2017.1288575.

Kim, I.-H., Feng, C.-C., Wang, Y.-C., Spitzberg, B. H., & Tsou, M.-H. (2017). Exploratory spatiotemporal analysis in risk communication during the MERS outbreak in South Korea. *The Professional Geographer, 69,* 629–643. https://doi.org/10.1080/00330124.2017.1288577.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics, 54,* 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y.

Li, S., Wang, Y., Xue, J., Zhao, N., & Zhu, T. (2020). The impact of COVID-19 epidemic declaration on psychological consequences: A study on active Weibo users. *International Journal of Environmental Research and Public Health, 17,* 2032. https://doi.org/10.3390/ijerph17062032.

Lwin, M. O., Lu, J., Sheldenkar, A., Schulz, P. J., Shin, W., Gupta, R., & Yang, Y. (2020). Global sentiments surrounding the COVID-19 pandemic on Twitter: Analysis of Twitter trends. *JMIR Public Health and Surveillance, 6,* e19447. https://doi.org/10.2196/19447.

Martín, Y., Li, Z., & Cutter, S. L. (2017). Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew. *PLoS ONE, 12,* e0181701. https://doi.org/10.1371/journal.pone.0181701.

Mohammad, S. M., & Turney, P. D. (2013). NRC emotion lexicon. National Research Council Canada, Canada.

Neppalli, V. K., Caragea, C., Squicciarini, A., Tapia, A., & Stehle, S. (2017). Sentiment analysis during Hurricane Sandy in emergency response. *International Journal of Disaster Risk Reduction, 21,* 213–222. https://doi.org/10.1016/j.ijdrr.2016.12.011.

Öztürk, N., & Ayvaz, S. (2018). Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis. *Telematics and Informatics, 35,* 136–147. https://doi.org/10.1016/j.tele.2017.10.006.

Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist, 89,* 344–350.

Seagal, S., & Horne, D. (2020). Human dynamics for the 21st century. *Systems Thinker*. Retrieved July 17, 2020, from https://thesystemsthinker.com/human-dynamics-for-the-21st-century/.

Shaikh, S., Feldman, L. B., Barach, E., & Marzouki, Y. (2017). Tweet sentiment analysis with pronoun choice reveals online community dynamics in response to crisis events. *Advances in cross-cultural decision making* (pp. 345–356). Cham: Springer International Publishing.

Shaw, S.-L., Tsou, M.-H., & Ye, X. (2016). Editorial: Human dynamics in the mobile and big data era. *International Journal of Geographical Information Science, 30,* 1687–1693. https://doi.org/10.1080/13658816.2016.1164317.

ST. (2020). Coronavirus: Designated clinics being activated from today. Retrieved August 7, 2020, from https://www.straitstimes.com/singapore/health/designated-clinics-being-activated-from-today.

Takahashi, B., Tandoc, E. C., & Carmichael, C. (2015). Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines. *Computers in Human Behavior, 50,* 392–398. https://doi.org/10.1016/j.chb.2015.04.020.

Today. (2020). 89% of migrant workers at dorms cleared of Covid-19: MOM. Retrieved August 7, 2020, from https://www.todayonline.com/singapore/89-per-cent-migrant-workers-at-dormitories-cleared-of-covid-19-mom.

Wang, Z., & Ye, X. (2018). Social media analytics for natural disaster management. *International Journal of Geographical Information Science, 32,* 49–72. https://doi.org/10.1080/13658816.2017.1367003.

WHO. (2020). World Health Organization: Mental health and COVID-19. Retrieved July 18, 2020, from https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/technical-guidance/mental-health-and-covid-19.

Worldometer. (2020). Coronavirus cases and deaths. Retrieved August 28, 2020, from https://www.worldometers.info/coronavirus/country/singapore/.

Yan, Y., Chen, J., & Wang, Z. (2020a). Mining public sentiments and perspectives from geotagged social media data for appraising the post-earthquake recovery of tourism destinations. *Applied Geography, 123,* 102306. https://doi.org/10.1016/j.apgeog.2020.102306.

Yan, Y., Eckle, M., Kuo, C.-L., Herfort, B., Fan, H., & Zipf, A. (2017). Monitoring and assessing post-disaster tourism recovery using geotagged social media data. *ISPRS International Journal of Geo-Information, 6,* 144. https://doi.org/10.3390/ijgi6050144.

Yan, Y., Feng, C.-C., Huang, W., Fan, H., Wang, Y.-C., & Zipf, A. (2020b). Volunteered geographic information research in the first decade: A narrative review of selected journal articles in GIScience. *International Journal of Geographical Information Science, 34,* 1765–1791. https://doi.org/10.1080/13658816.2020.1730848.