

## Research Article

# The Emotional Impact of COVID-19 News Reporting: A Longitudinal Study Using Natural Language Processing

Simon L. Evans <sup>1</sup>, Rosalind Jones,<sup>1</sup> Erkan Alkan <sup>2</sup>, Jaime Simão Sichman <sup>3</sup>,  
Amanul Haque <sup>4</sup>, Francisco Bráulio Silva de Oliveira <sup>3</sup>, and Davoud Mougouei <sup>5</sup>

<sup>1</sup>Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK

<sup>2</sup>Institute of Health and Wellbeing, University of Suffolk, Ipswich, UK

<sup>3</sup>Laboratório de Técnicas Inteligentes, Escola Politécnica, Universidade de São Paulo, Av. Prof. Luciano Gualberto, Travessa 3, 158, 05508-970 São Paulo, SP, Brazil

<sup>4</sup>Social AI Lab (Engineering Building 2, 2261), Department of Computer Science, North Carolina State University, Raleigh, NC 27695-8206, USA

<sup>5</sup>School of Information Technology, Deakin University, Burwood, VIC, Australia

Correspondence should be addressed to Simon L. Evans; [simon.evans@surrey.ac.uk](mailto:simon.evans@surrey.ac.uk)

Received 16 August 2022; Revised 19 October 2022; Accepted 3 November 2022; Published 9 March 2023

Academic Editor: Zheng Yan

Copyright © 2023 Simon L. Evans et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The emotional impact of the COVID-19 pandemic and ensuing social restrictions has been profound, with widespread negative effects on mental health. We made use of the natural language processing and large-scale Twitter data to explore this in depth, identifying emotions in COVID-19 news content and user reactions to it, and how these evolved over the course of the pandemic. We focused on major UK news channels, constructing a dataset of COVID-related news tweets (tweets from news organisations) and user comments made in response to these, covering Jan 2020 to April 2021. Natural language processing was used to analyse topics and levels of anger, joy, optimism, and sadness. Overall, sadness was the most prevalent emotion in the news tweets, but this was seen to decline over the timeframe under study. In contrast, amongst user tweets, anger was the overall most prevalent emotion. Time epochs were defined according to the time course of the UK social restrictions, and some interesting effects emerged regarding these. Further, correlation analysis revealed significant positive correlations between the emotions in the news tweets and the emotions expressed amongst the user tweets made in response, across all channels studied. Results provide unique insight onto how the dominant emotions present in UK news and user tweets evolved as the pandemic unfolded. Correspondence between news and user tweet emotional content highlights the potential emotional effect of online news on users and points to strategies to combat the negative mental health impact of the pandemic.

## 1. Introduction

The mental health impact of the coronavirus disease 2019 (COVID-19) pandemic has been profound: almost all sections of society have been adversely affected. The UK's Centre for Mental Health predicts that an additional 20% of the UK's population will require mental health support due to the pandemic [1]. There is consistent evidence that the pandemic has triggered widespread stress, anxiety, and depression [2]. The social restrictions imposed have been a significant contributor to this, increasing loneliness levels

across much of the population, with consequences for mental illness symptomatology [3]. In response, many people came to rely even more on social media and online news platforms to communicate and interact and access and share information, accelerating prepandemic trends: UK adults spent 15% more time online in 2020 compared to 2019 [4]. Evidence suggests that loneliness levels in young adults have been particularly impacted by the pandemic, with loneliness amongst young adults linked to a greater increase in both their social media use and social support seeking [5]. Findings such as this, as well as from studies prepandemic,

suggest that social media and online news platforms can impact users' mental health and wellbeing [6]. Characterising the emotional content of media and user posts on social media allows deeper insights into this. Although these channels could potentially be leveraged to benefit the health and wellbeing of society as we begin to deal with the impact of the pandemic, the complex relationship between online media content and users' emotional responses and wellbeing needs to be better understood.

With the usage of social media and online news platforms at record levels, it is crucial to develop this understanding. While potential benefits include an expansion of social networks [7], better access to information, and a supportive online community, for those facing difficulties [8], significant risks and downsides have been identified. Studies conducted during the pandemic have tended to focus on social media use frequency, usually measured by self-report in cross-sectional designs; higher usage has been linked to higher stress levels [9], anxiety [10], and loneliness [11], alongside a poorer quality of life [11] and wellbeing [12]. However, it is difficult to draw inferences regarding causality from such data, and far fewer studies have investigated online news content and the responses to these on social networks. This allows more objective insights into the emotional impact on users, as well as general insights into the overall emotional status of the population at particular points in time, as the pandemic situation unfolded.

Previous work focusing on past disasters has identified important effects of exposure to online news content on mental health. In the Boston Marathon bombings of 2013, those who viewed graphic images of the attack online were subsequently at higher risk of showing posttraumatic stress disorder (PTSD) symptoms [13, 14]. Likewise, longitudinal work by Silver et al. [15] in a US sample ( $N = 2189$ ) linked higher exposure to media coverage of the 9/11 attacks with increased stress and poor health across the 3 years after the event [15]. However, the directionality is not clear-cut Thompson et al. [16] found via surveys that those showing symptoms of trauma after the Boston bombings, and after the Pulse nightclub attacks of 2016, were more likely to view media coverage of the event, which in turn exacerbated and perpetuated their trauma. Nevertheless, studies such as these point to the direct impact that online news consumption can have in times of crisis. However, these studies relied on self-report measures of both media exposure and mental health, and these isolated events involved graphic and traumatic imagery, and thus might not be very generalisable to pandemic situations. However, while not a single isolated incident like 9/11, the pandemic has had a wide and profound global impact and has produced a stream of fear-inducing news reports and content, particularly in the earlier stages of the virus' emergence and spread, and consistent with these prior findings, emerging evidence points to a connection between online news consumption and mental health during the pandemic. Based on self-reported survey responses from a probability-based US sample ( $N = 10,606$ , American Trends Panel, March 2020), Stainback et al. [17] found that those who followed the news "very closely" reported nearly 25% higher distress levels compared to those

who followed the news "fairly closely," "not too closely," or "not at all closely"; similar cross-sectional studies that focused specifically on online news consumption during the pandemic have likewise found correlations with negative affect, anxiety, and stress [18–20]. However, these studies, relying on subjective data, have limitations including social desirability and sampling bias, and larger sample sizes would be helpful to infer effects across populations.

An alternative approach is to consider the content being generated by online users and draw inferences from the emotions expressed within it. A powerful method for doing this is natural language processing (NLP) implemented using machine learning models: this approach involves analysing large datasets of user comments by automated means, to extract patterns and trends in emotions expressed. One form of NLP is sentiment analysis and emotion recognition; this measures the emotional polarity of the text, deriving measures of negative and positive emotions, and levels of specific emotions, expressed within it. Another useful element of NLP is topic modelling which can classify text into a set of preselected topics or themes; this allows inferences regarding relationships between topics and emotional responses evoked and user reactivity to certain topics. The capacity of these tools to process data on a very large scale allows for very well-powered and generalisable studies to be conducted, although the ability to consider individual-specific variables is limited since the characteristics of the users providing the content are largely unknown.

Some studies have been published using NLP tools to investigate user responses to the pandemic on Twitter [21–33]; heightened emotions and high levels of negatively coded sentiments were present in the first half of 2020, with the economy, lockdowns, and health as prominent topics. Other studies have tracked sentiment change over time, finding an increase in negativity expressed as the pandemic situation emerged, with increased stress, isolation, and anxiety being identified [27, 34–37], with a general increase in individual user engagement with Twitter, as the pandemic situation unfolded [37]. Emotion recognition alongside topic classification can be used to identify specific emotions: Basile et al. [38] analysed emotional reactions to events (via Reddit) between January and June 2020; fear was found to predominate across the threads, and negative sentiment increased when national cases of COVID-19 increased, with sentiment also being connected to governmental responses. Likewise, Li et al. [39] found a correlation between stress levels, estimated from Twitter data in major US cities, and the number of local COVID cases. Low et al. [35] focused on Reddit users in support groups, including groups for suicide, schizophrenia, and depression, finding increased negative emotions, loneliness, and risky behaviour including suicidal ideation during the pandemic compared to before. With regard to user responses to online news during the pandemic, Aslam et al. [40] analysed nearly 150000 news headlines from the first half of 2020, across 25 global news organisations, finding that 52% of headlines were negative in polarity. However, Krawczyk et al. [41] looked at 26 million news articles from the front pages of 172 major online news sources across 11 countries between January and

October 2020; their sentiment analysis revealed widely heterogeneous emotional polarity, and only 16% of COVID-19 news articles were classified as highly negatively polarized. Nevertheless, there is evidence that COVID-related topics and themes show consistency across both online news content and Twitter posts by users. de Melo and Figueiredo [42] analysed nearly 20 thousand news media articles and over a 1.5million tweets, from Brazil. Topics and themes showed high correspondence between news media and social media, pointing to a strong influence of online news content on user postings.

Evidently, the use of NLP techniques to understand the impact of the pandemic on user wellbeing and attitudes has been fruitful, allowing large-scale inferences regarding the emotional responses of populations, as the pandemic continues to unfold. Some studies have also revealed useful insight into emotional content within online media reporting. However, studies using emotion recognition techniques have tended to focus on either user content or on news headlines, not the relationship between both. Although some small, country-specific studies have been conducted (e.g., Tune et al. [43] examined users' reactions to COVID-related media coverage in Bangladesh in the initial wave of the pandemic), there is limited literature around this, and work focusing on this relationship is needed in order to better understand the degree to which user sentiment is influenced by online news consumption. Therefore, this study analysed the emotions present in both online media content and the associated user responses to this content, to investigate how the emotions recognised in news tweets (tweets from news organisations) interact with users' emotions, in the direct comments to these posts. Studies continue to emerge covering the pandemic, and this study sets out to enhance the understanding of the pandemic's impact, focusing on temporal patterns of emotions contained in online news and the corresponding emotional responses from users, between January 2020 and April 2021. We looked specifically at UK news sources, considering a range of outlets. This allowed us to conduct analyses within particular time epochs that mapped the time course of the COVID situation in the UK, to investigate how online emotions and content linked to the UK social restrictions and virus spread. Topic modelling is helpful as it points to the key topics within online content at a specific time point. For example, Kit et al. [44] also focused on the UK and used topic modelling to identify how key themes on Twitter (such as mask wearing) tracked the public reaction to current events and government policy and found evidence of a relationship between the volume of COVID-related tweets in specific parts of the UK with the local death rates, producing what they describe as a potential "early warning system" for tracking cases.

The studies discussed above demonstrate the ability of NLP tools to provide valuable insights into population mental health to characterise the impact of the pandemic both over time and within specific user groups. Using large datasets of tweets from multiple major UK news outlets, and user responses to these, we set out to draw inferences regarding patterns of emotional content and prevailing topics in these

and the correspondence between user and news tweets. The research questions under study were

- (1) Characterise how the emotions present in UK media postings, and user content, varied across 2020/2021, considering key epochs reflecting the UK pandemic situation and the UK's response to it, with an increase in negative affect (sadness and/or anger) predicted
- (2) To investigate the degree of correlation between emotions (anger, joy, optimism, and sadness) expressed in news organisation tweets and the tweets made by users in response

Topic modelling was used to provide context regarding the emotions present, regarding prevailing themes. Characterising the emotional status of users could inform support strategies for mental health, and by studying the relationship between emotions present in users and online media content, will inform and provide better awareness of the news's impact on societal wellbeing.

## 2. Materials and Methods

*2.1. Data Collection.* We use data from Twitter, an online microblogging and social networking site publicly available and accessible to those with access to the Internet. Registered Twitter users are able to post short messages called "tweets" and interact with other users' tweets by comments, replies, and "likes." Recent years have witnessed an increase in the use of Twitter data to understand social phenomena, and it has been shown to be an effective means of tracking the public's reaction to both large- and small-scale events [45].

The data used in this project is in the format of written tweets collected using Snsrape API, an open-source web-scraping library for social networks. This API allows researchers to access public tweets via specific searches (e.g., via keywords and location). All tweets collected are shared publicly, and therefore, all users whose tweets are included in the dataset have consented to this data being accessible to the public via Twitter's privacy policy (<https://twitter.com/en/privacy>). No demographic information of the individual user of each tweet was collected, only the content and location. Ethical approval was granted via the University of Surrey. The data is organised into news tweets (posts by news organisations on their Twitter accounts) and user tweets (Twitter users' comments on these news tweets, in the form of comments to the news threads). Data used was a subset of that collected by a larger ongoing study (by Global Affects, <https://globalaffects.org/>); here, data from 19 UK-based news is used (Table 1). Tweets were sourced via the API from the preselected news channel tweets and the replies to these tweets over the period from 1 January 2020 to 30 April 2021 using a set of keywords: wuhan, ncov (abbreviation of Novel Corona Virus), coronavirus, covid, sars-cov-2, pandemic, lockdown, quarantine, social distancing, wearing masks, vaccination, vaccine, outbreak, panic buying, remote working, homeschooling.

TABLE 1: UK news media organisations.

News organisation Twitter account	Number of news tweets in the dataset	Number of user tweets in the dataset
BBC Breakfast	3536	
BBC Breaking News	324	
BBC News (UK)	9507	751700
BBC News (World)	4257	
BBC News NI	2114	
BBC Newsnight	1390	
BBC Radio 4	740	
BBC Radio 4 Today	680	
BBC Radio 5 Live	1531	
BBC Scotland News	4855	
Channel 4 News	2442	
Daily Express	14527	
Daily Mail Online	11871	119905
Metro	3550	
The Guardian	23993	272861
The Independent	29308	
The Mirror	33199	
The Sun	15713	
The Telegraph	11159	

Note: full list of UK news channels used in analysis including their corresponding frequencies within the news tweet dataset. User tweet numbers for only the three selected channels for analysis are shown.

**2.2. Data Cleaning.** To clean the data, a set of exclusion criteria was used. The following tweets were excluded: tweets from private accounts, comments on comments (leaving only direct comments to the news tweet), comments already excluded by the platform or by the user themselves, tweets that were not classified as English by Twitter, and tweets less than 3 words long. A total of 2,868,260 tweets resulted: 174,696 news tweets and 2,693,563 user tweets; this meant on average there were 15 comments per news tweet collected. For the correlation analyses, we used data relating to BBC News (UK), The Guardian, and the Daily Mail Online only (1,114,466 user tweets total). The corpus of tweets was then put through two NLP processes: (1) topic modelling and (2) emotion recognition.

**2.2.1. Topic Modelling.** To identify topics being discussed in tweets we used latent Dirichlet allocation (LDA) [46], an unsupervised machine learning approach that finds latent topics in unlabelled text data. We used the MALLET (machine learning for language toolkit) [47] implementation of LDA. It generates a probabilistic distribution of words over latent topics, with each topic characterized by this distribution. LDA was chosen to identify topics because it is an unsupervised approach that does not require annotated data for training; also, it allows us to specify the number of latent topics to look for and produces results that are easily interpretable.

The input to an LDA model is a text corpus (tweets in this case). Since raw tweets contain a lot of uninformative

word tokens (e.g., stopwords, punctuations, special characters, and numbers), they need to be processed before feeding it to the LDA model. This was done by firstly applying a regular expression to keep words and spaces while removing numbers and underscore symbols. Then, each news tweet was converted into a list of tokens (i.e., a sequence of characters that is a useful semantic unit for processing) in the lowercase lemma format (i.e., the base or dictionary form of a word), and the team then removed all tokens with length equal to one. All tokens that start with the “@” symbol were also removed, to avoid references to other Twitter accounts. Words classified as geographical entities (i.e., countries, cities, and state names) using a named entity recognizer (NER) were removed. This ensured that the tweets were more reflective of different themes related to the COVID-19 pandemic and not biased towards grouping together tweets originating from similar geographical locations. Lastly, tokens that occurred in less than 0.005% or more than 33% of the documents were removed, so as not to allow rare and highly common tokens to bias the topic modelling.

LDA requires the number of latent topics to be specified beforehand. We used coherence, a proxy for topic quality, as a proxy to identify the optimal number of latent topics in the given tweets. Coherence is based on the distributional hypothesis that words with similar meanings tend to co-occur within similar contexts [48]. We experimented by varying the number of topics between 2 and 40 with steps of 3 and computed the coherence for each. 17 was found to be the optimal number of topics for the dataset that yields a coherence score of 0.51. These topics were manually merged into 8 main topics/themes based on the similarity of the topics with regard to COVID-19. Table 2 contains the topics identified by LDA for our dataset and the corresponding keywords.

**2.2.2. Emotion Recognition.** The RoBERTa model [49] was selected as it is retrained on ~58M tweets to capture Twitter language specifics and is fine-tuned for emotion recognition on SemEval “Affects in Tweets” dataset [50]. The model is evaluated on TweetEval [51], an evaluation benchmark for Twitter-specific classification tasks, and achieved a state-of-the-art performance of 79.8% macroaveraged F1-score on the task of emotion recognition. A macroaveraged F1 score, also referred to as macro-F1-score, is calculated based on the arithmetic mean of all the per class F1 scores. This metric treats all classes equally regardless of their support values [52].

Emotion recognition codes the data into different emotions, providing a score to each emotion for each tweet collected. This model effectively distinguishes between four core emotions: joy, optimism, sadness, and anger. The model generates a softmax score of between 0 and 1 for each emotion, and this score was used to classify each tweet. Softmax is an exponential function that normalizes the output of a model to a probability distribution over predicted classes. We classified each tweet according to the emotion with the highest softmax score, as long as this was above 0.50. Softmax scores also help identify instances that do not fall within any of the four core emotions used: a tweet containing emotions other than the ones our model can recognise usually



TABLE 2: Topic classification: topics and keywords.

Topic number	Topic theme	Keywords
1	Cases and deaths	case, state, rise, death, report, infection, confirm, numb, increase, total, high, recovery, govt, surge, record, daily, breaking, positive, number, active_case, district, hour, spike, continue, death_toll, health_ministry, rate, fatality, low, additional, jump, total_numb, today, reach, discharge, fresh, mark, department_health, county, tally, include, bring_total, register, double, past_hour, newsalert, drop, climb, live_updates, soar, recover, record_high, relate_death, count, log, yesterday, fall, patient_recover, worldwide, death_toll_rise, centre, hospitalization
2	Economic impact	plan, impact, work, crisis, business, government, hold, economy, community, deal, hit, group, fund, global, support, cancel, demand, offer, house, market, free, sign, company, launch, event, job, employee, relief, industry, provide, seek, half, service, ongoing, cost, food, postpone, cut, ensure, aid, expect, centre, push, cent, pay, budget, big, effort, pass, sector, affect, delay, boost, urge, money, sale, price, target, concern, bank, include, bill, stock, cover, power, quarter, spend, firm, poor, massive, sell, protect, season, trade, drive, benefit, growth, donate, set, financial, site, lawmaker, estimate, shock, initiative, unveil, relief_bill, revenue, social, small_business, gain, economic_recovery, apply, annual, lose_job, sabah, relief_package, fallout, programme, festival, weigh, workplace, call, performance, scheme, proposal, month, donation, file, global_economy, plant, stimulus_package, economist, employer, judge, credit, assistance, inquiry, boom, mission, legal, job_loss, world_bank
3	Educational impact	school, home, close, amid, open, student, child, return, pandemic, force, leave, reopen, shut, online, send, stay, learn, safe, parent, kid, scare, shutdown, university, miss, store, teacher, closure, education, summer, work_home, class, turn, bring, finally, virtual, option, migrant, work, space, owner, college, fee, teach, children, exam, reopen_school, safely, teen, girl, meal, mom, parents, remain_open
4	People stories	news, pandemic, die, face, people, post, man, change, family, write, woman, life, contract, challenge, resident, share, year_old, make, person, opinion, breaking, experience, reveal, big, age, good, happen, wrong, victim, future, couple, long, kill, feel, infect, problem, diagnose, great, nurse_home, survive, wife, lot, visit, hope, isolation, play, create, mother, cope, young, mental_health, jail, stress, prison, love, real, son, history, care_home, moment, complication, difficult, dozen, story, catch, film, daughter, bring, reportedly, opportunity, husband, reality, inmate, look_like, father, era, year_old_man, new_normal, mum, kind, baby, digital, leave, devastate, year_old_woman, decade, music, home, bite, funeral, art, dad, overcome, pass_away, environment, perspective, relative, song
5	Authorities and politics	response, update, live, late, american, official, watch, call, early, pm, emergency, give, meet, president, province, today, nation, release, situation, national, datum, video, lead, health_official, speak, team, accord, talk, claim, join, month, question, crisis, watch_live, minister, declare, discuss, handle, news, fight, white_house, tuesday, health_minister, leader, expect, detail, expert, relate, listen, address, election, analysis, senior, special, governor, govt, policy, week, explain, review, task_force, tomorrow, action, track, tonight, approach, friday, union, public_health, administration, follow, answer, fail, federal, vaccine_rollout, catch, begin, interview, head, outbreak, tv, model, states, win, political, guidance, conversation, defend, publish, debate, thread, commission, answer_question, trust, preparedness, failure, leadership, chief_medical_officer, episode, defeat, praise, coming, criticism, statistic, opposition, headline, voter, projection, speech, historic, government, climate_change
6	Preventative measures	Day, order, lockdown, country, travel, restriction, read, start, quarantine, measure, disease, week, police, ease, city, issue, tier, level, area, social_distance, place, announce, follow, rule, list, guideline, impose, add, end, protocol, lift, extend, mask, boris_Johnson, begin, public, suspend, curb, government, month, passenger, protest, flight, strict, wear_mask, christmas, ng, sunday, resume, anti, remain, nationwide, train, break, ncov, decide, check, curfew, roque, period, hotel, complete, relax, holiday,

TABLE 2: Continued.

Topic number	Topic theme	Keywords
7	Virus spreading	ahead, national, august, arrest, stage, set, require, tough, arrive, urge, enforce, date, breach, street, monitor, extension, airport, limit, arrival, fine, plan, region, need_know, quarantine_facility, advice, crowd, enhance_community_quarantine, village, ban, johnson, affect, relaxation, airline, personnel, average, activity, decision, violate, local government, phase, authority, pub, photos, tourism, mandatory, exercise Patient, test, outbreak, hospital, people, spread, health, find, warn, doctor, study, virus, battle, variant, fear, treatment, result, wave, facility, symptom, treat, risk, care, infect, chinese, medical, positive, europe, system, show, control, nurse, strain, staff, conduct, deadly, chief, detect, stop, fight, identify, inside, research, grow, line, screen, cluster, build, lab, expert, cure, researcher, scientist, admit, carry, prevent, body, expose, concern, healthcare_worker, negative, worry, run, transmission, condition, contact, evidence, world_health_organization, bed, wait, epidemic, shift, reduce, link, health_worker, population, thecable, suspect, country, work, human, emerge, fire, test_result, lock, officer, undergo, recover, test_kit, global, critical, ventilator, sample, panic, nhs, asymptomatic, ncdc, doctors, develop, quickly, sick, antibody, wrong, rapid, winter, ward, scientists, spark, describe, high_risk, lead, frontline, exposure, resurgence, private_hospital, trace, expert_warn, duty, close_contact, epicentre, animal, army, long, origin, overwhelm, caution, oxygen, claim, save_live, mutation, danger, ppe, researchers, epicenter, intensive_care, hospitals, ability, palliative, infectious, gravitas, dedicate, rapid_test, scam, equipment, transfer, fears
8	Vaccines and vaccination	Vaccine, dose, receive, astrazeneca, trial, supply, ready, jab, shoot, approve, pfizer, roll, develop, effective, eu, development, vaccinate, shot, deliver, distribution, race, delivery, covidvaccine, administer, volunteer, produce, distribute, approval, johnson_amp_johnson, clinical_trial, rollout, vaccination, fda, moderna, manufacture, russian, potential, sinovac, pfizer_biontech, vaccine_candidate, give, manufacturer, experimental, shipment, covishield, procure

Note: example of keywords used by the NLP model to classify tweets into the 8 topics. The table does not include every keyword used by the model.

gets a more homogeneous distribution of confidence scores across the four emotions it can detect; when none of the emotion softmax scores exceeds 0.50, we label such tweets as being “undefined.”

As an example, the sentence “I hate the lockdown, it feels so isolating.” gets scored: anger .230, sadness .746, joy .010, and optimism .013. Since sadness has the highest softmax score and it is above 0.50, we label this tweet as “sad.”

**2.3. Analysis.** To provide an overall picture of the topic prevalence and emotion levels over the timeframe under study, the entire UK news tweet dataset was considered first (all 19 channels). This allowed us to depict the wider trends in emotion and topic. Next, we focused on three specific channels: BBC News (UK), The Guardian, and the Daily Mail, to formally investigate correlations between news tweets and user tweets. The three channels were selected as they represented three major mainstream UK media channels with large followings but with different audiences and political leanings: according to a YouGov poll in 2017, The Guardian was regarded as the most left-wing UK newspaper included in the survey, and the Daily Mail was the most right-wing [53]. The BBC is officially a politically neutral organisation. Conducting the analyses on each channel separately not only allowed us to infer channel-specific effects but also to check for replicability across channels.

*Effect of time.* Similar to the methodology of Su et al. [54], to create the time epochs for the analyses, time periods were chosen according to the UK Government data on the COVID “waves” of infections [55] and the corresponding

imposition/easing of social distancing measures. The four time epochs were

- (1) Early phase, before UK social distancing imposed: Jan 2020–22nd March 2020
- (2) 23 March 2020–30 May 2020: “Wave One,” social distancing measures imposed
- (3) 31 May 2020–6 September 2020: lower numbers of cases, restrictions eased
- (4) 7 September 2020–April 2021: “Wave Two,” increased number of cases

Only the BBC News (UK) news tweet dataset and corresponding user tweet datasets were used for this analysis, as (a) it was the largest of all the media channel user datasets and (b) the BBC is politically neutral.

### 3. Results and Discussion

Each tweet represented one unit of data in the study; each tweet had multiple variables derived from it including date, tweet content, topic, prevalent emotion, and emotion scores for anger, joy, optimism, and sadness (ranging between 0 and 1, for each). News tweets were labelled with which organisation they were from, and user tweets were labelled according to which news tweet they were commenting on.

**3.1. Emotions and Topics in the News Tweets (Full Dataset).** The full news tweet dataset consisted of 174,696 individual

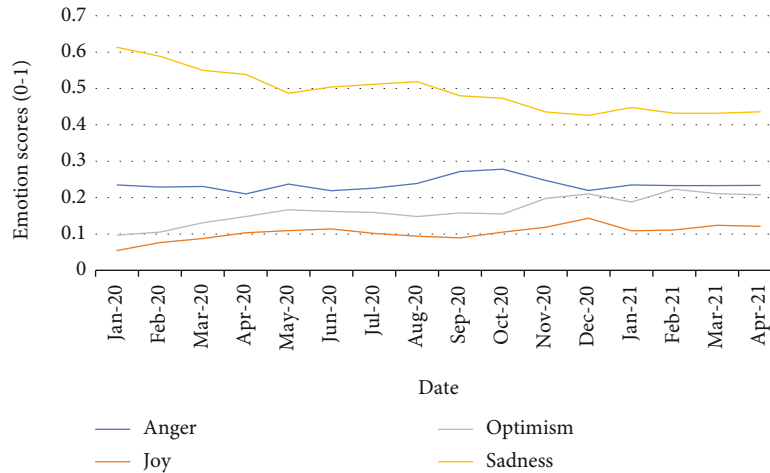


FIGURE 1: Monthly average emotion scores in the news tweets.

TABLE 3: Frequency of topics in news tweets.

Topic	Counts	Total %
Cases and Deaths	18890	10.8%
Economic Impact	30419	17.4%
Educational Impact	19308	11.1%
People Stories	11379	6.5%
Authorities and Politics	16643	9.5%
Preventative Measures	34408	19.7%
Virus Spreading	35900	20.5%
Vaccines and Vaccination	7749	4.4%

Note: table shows the distribution of each classified topic across the entire news tweet dataset.

tweets from 19 media channels across the UK, see Table 1 for the number of tweets per channel. The average news tweet emotion scores were calculated across the period January 2020 to April 2021, see Figure 1.

Overall, 50.5% of news tweets were classified as sad, 14.4% as angry, 9.3% were optimistic, 5.9% were joyful, and 19.9% of news tweets were labelled undefined. The average anger score across the news tweet data set was  $M = .235$  ( $SD = .241$ ), for joy  $M = .105$  ( $SD = .182$ ), optimism  $M = .167$  ( $SD = .197$ ), and sadness  $M = .492$  ( $SD = .303$ ). Supplementary Table 1 presents average emotion scores for each of the news channels.

Table 3 shows the overall frequencies of each topic, and Figure 2 depicts how the distribution of topics in the news tweets changed over time. The most common topic was “Virus Spreading” at 20.5%, and the least was “Vaccines And Vaccination” at 4.4%, although the prevalence of “Vaccines And Vaccination” only became apparent from midway through 2020.

The average emotion scores for each topic were then examined, see Table 4 for the mean emotion scores, by topic. The topic with the highest anger score was “Preventative Measures” ( $M = .339$ ,  $SD = .274$ ). The topic with the highest joy and optimism score was “Vaccines And Vaccination” ( $M = .172$ ,  $SD = .210$  and  $M = .312$ ,  $SD = .245$ , respectively).

The topic with the highest sadness score was “Cases and Deaths,” ( $M = .633$ ,  $SD = .264$ ).

**3.2. Correlation between News and User Emotions.** To investigate whether the news tweet emotion levels were correlated to the user tweet emotion levels, correlation analyses were carried out. These analyses were conducted using news and user tweets from three selected news channels: BBC News (UK), the Daily Mail Online, and The Guardian. User tweets were direct comments and replies to the news tweets. In total, the user tweets consisted of 1,114,466 user tweets across the time-frame under study (1 January 2020 to 30 April 2021). The average emotion scores for every month for each chosen news tweet were calculated. The same process was repeated for their user comments, and then, a correlation matrix was created for each of the three channels (see Tables 5–7). Spearman’s Rho coefficients were used due to the data being nonparametric. Figures 3–8 show emotion levels over time, by channel, for news tweets (Figures 3, 5, and 7), and corresponding user tweets (Figures 4, 6, and 8).

These highlight significant correlations between user emotions and news emotions: significant positive correlations emerged for all emotions and across all three channels. Based on the test statistics (Tables 5–7), anger consistently emerged as having the strongest correlation between user and news tweets (compared to the other emotions studied) across all three datasets (for anger: BBC News:  $rs = .14$ ,  $p < .001$ ; The Guardian:  $rs = .19$ ,  $p < .001$ ; and the Daily Mail Online:  $rs = .12$ ,  $p < .001$ ).

**3.3. Changes in News and User Emotions across Specific Time Epochs (BBC News (UK)).** We then conducted statistical tests to investigate changes in emotion level by epoch. Four key epochs were distinguished, representing (1) Early phase (before UK social distancing imposed), (2) Wave One (spring 2020), (3) low cases/minimal restrictions (summer 2020), and (4) Wave Two (autumn 2020–spring 2021). One-way ANOVA was conducted to identify differences in news and user emotions between these four epochs, using the BBC News (UK) tweets (news tweets and corresponding

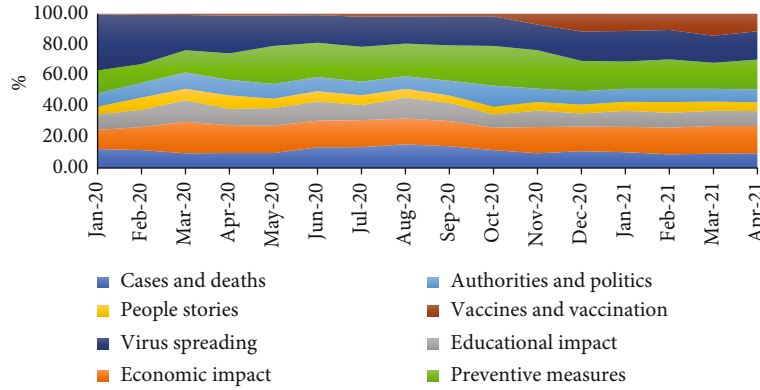


FIGURE 2: Figure illustrating the average distribution of the 8 topics across all channels, over time. Distribution was calculated as a percentage of all tweets per month.

TABLE 4: Average news tweet emotion scores per topic.

Topic	Anger	Joy	Optimism	Sadness
Cases and Deaths	.174	.080	.112	.633
Economic Impact	.230	.104	.182	.483
Educational Impact	.217	.101	.142	.540
People Stories	.139	.104	.135	.622
Authorities and Politics	.266	.135	.204	.395
Preventative Measures	.339	.103	.162	.395
Virus Spreading	.204	.093	.164	.539
Vaccines and Vaccination	.210	.172	.312	.299

Note: table showing mean emotion scores in news tweet dataset, across entire period, for all news channels. Scores are emotion levels coded by the NLP model, ranging 0 to 1.

user tweets) only (see Table 8 for the descriptive statistics per emotion, by epoch).

For the news tweet dataset, the assumption of homogeneity of variance was violated, due to Levene’s test being significant for all emotion categories. Therefore, Welch’s  $F$ -statistics are reported. There was a main effect of epoch for all emotion categories: anger:  $F(3, 2797.24) = 14.91$ ,  $p < .001$ ; joy:  $F(3, 2744.49) = 15.55$ ,  $p < .001$ ; optimism:  $F(3, 2702.71) = 71.34$ ,  $p < .001$ ; and sadness:  $F(3, 2665.72) = 116.29$ ,  $p < .001$ . Post hoc tests (Games-Howell due to unequal variances) for each emotion showed changes between each epoch with significant increases in anger scores between epochs one and four ( $p < .001$ ), epochs two and four ( $p < .001$ ), and epochs three and four ( $p < .001$ ). Joy levels increased between each epoch, with significant increases between time epochs one and three ( $p < .01$ ), one and four ( $p < .001$ ), two and four ( $p < .001$ ), and epochs three and four ( $p < .001$ ). Optimism levels ended significantly higher with significant increases in optimism between epochs one and four ( $p < .001$ ), two and four ( $p < .001$ ), and three and four ( $p < .001$ ). Lastly, sadness levels decreased across all epochs. Sadness significantly decreased between epochs one and two ( $p < .01$ ), one and three ( $p < .001$ ), and one and four ( $p < .001$ ) and also between epochs two and four ( $p < .001$ ) and three and four ( $p < .001$ ).

For the corresponding user tweet dataset, the assumption of homogeneity of variance was also violated due to Levene’s test being significant. Therefore, Welch’s  $F$ -statistics are reported. As in the news tweets, there was a main effect of epoch for all four mean emotion scores: anger:  $F(3, 141212.21) = 183.46$ ,  $p < .001$ ; joy:  $F(3, 141110.97) = 66.50$ ,  $p < .001$ ; optimism:  $F(3, 140558.84) = 12$ ,  $p < .001$ ; and sadness:  $F(3, 139859.6) = 216.47$ ,  $p < .001$ . Post hoc tests (Games-Howell) were then conducted. Similar to the news tweets, anger levels increased throughout, with user anger scores significantly increasing across the four epochs ( $p < .001$  for all pairwise comparisons). Joy levels were seen to increase across the first 3 epochs and then decline: there was a significant increase from epoch one versus two and three ( $p < .001$ ), followed by a significant decrease in joy levels in epoch four, compared to both epochs two and three ( $p < .001$ ). On the other hand, optimism levels significantly decreased in epochs two, three, and four compared to epoch 1 ( $p < .001$  for all), but there was a significant increase in optimism between epochs three and four ( $p < .01$ ). However, as in the news tweets, sadness levels decreased across the epochs, with a significant decrease in sadness across the first 3 epochs (all  $p$  values  $< .001$ ) and between epochs three and four ( $p < .05$ ).

## 4. Discussion

By leveraging NLP techniques, this study has provided a large-scale picture of online emotional responses by Twitter users, to COVID-related news. The aim was to characterise how prevailing emotions evolved between the virus first being publicised at the start of 2020 and April 2021. Topic modelling was used to provide context to the findings. We classified emotions within a large corpus of news tweets from leading UK media channels, the user emotions in response, and the correlations between the two. Finally, focusing just on the largest dataset (BBC News UK), we conducted analyses of effects between four specific time epochs, chosen based on the timeline of social distancing and “waves” of infections, between January 2020 and April 2021.

Consistent emotion patterns in the news and user tweets across the different news channels were observed. Sadness



TABLE 5: BBC News (UK) correlation matrix showing Spearman’s Rho coefficients between news and user emotion scores.

	1	2	3	4	5	6	7	8
1 News anger	—							
2 News joy	-.20**	—						
3 News optimism	-.12**	.66**	—					
4 News sadness	-.35**	-.47**	-.57**	—				
5 User anger	.14**	-.05**	-.01**	-.06**	—			
6 User Joy	-.08**	.09**	.04**	-.02**	-.71**	—		
7 User optimism	-.03**	.04**	.08**	-.03	-.57**	.6**	—	
8 User sadness	-.09**	-.02**	-.02**	.1**	-.47**	.12**	.25**	—

Note: scores are emotion levels coded by the NLP model from 0 to 1. Columns 1-8 represent the 8 categories labelled in the first column. \*\* $p < .001$ .

TABLE 6: The Guardian correlation matrix showing Spearman’s Rho coefficients between news and user emotion scores.

	1	2	3	4	5	6	7	8
1 News anger	—							
2 News Joy	-.32**	—						
3 News optimism	-.27**	.67**	—					
4 News sadness	-.37**	-.33**	-.43**	—				
5 User anger	.19**	-.06**	-.03**	-.09**	—			
6 User Joy	-.12**	.11**	.06**	.00	-.71**	—		
7 User optimism	-.05**	.05**	.09**	-.01**	-.56**	.58**	—	
8 User sadness	-.11**	-.01**	-.01*	.12**	-.42**	.07**	.22**	—

Note: scores are emotion levels coded by the NLP model from 0 to 1. Columns 1-8 represent the 8 categories labelled in the first column. \* $p < .01$  and \*\* $p < .001$ .

TABLE 7: The Daily Mail Online correlation matrix showing Spearman’s Rho coefficients between news and user emotion scores.

	1	2	3	4	5	6	7	8
1 News anger	—							
2 News Joy	-.28**	—						
3 News optimism	-.22**	.75**	—					
4 News sadness	-.47**	-.36**	-.36**	—				
5 User anger	.12**	-.02**	.004	-.07**	—			
6 User Joy	-.04**	.04**	.04**	-.03**	-.71**	—		
7 User optimism	-.03**	.09**	.09**	-.006**	-.54**	.58**	—	
8 User sadness	-.11**	-.02**	-.02**	.15**	-.44**	.09**	.24**	—

Note: scores are emotion levels coded by the NLP model from 0 to 1. Columns 1-8 represent the 8 categories labelled in the first column. \*\* $p < .001$ .

was the most prevalent emotion across channels and throughout the time period under study. Analysing correlations between news and user tweet emotions for three major UK news sources (BBC News (UK), the Daily Mail Online, and The Guardian) revealed significant correlations, across all emotions and for all 3 sources. Regarding the effect of time epoch, ANOVA results analysing the BBC News data revealed significant effects of time epoch on emotion levels for both news and user tweets. In the news tweets, anger levels increased throughout the time epochs, and sadness declined, although sadness remained the most prevalent emotion in the news tweets throughout, with just over 50%

being classified as “sad”; this remained relatively consistent despite the significant change in scores between time epochs. Both joy and optimism levels increased across epochs. For user tweets, anger followed the same pattern, significantly increasing across time epochs. Likewise, sadness decreased throughout. Unlike the news tweets, joy levels were seen to increase across the first 3 epochs and then decline in epoch 4 (which represented Wave Two of infections). For optimism, levels were significantly lower in epochs two, three, and four compared to epoch 1 ( $p < .001$  for all), but there was a significant increase in optimism between epochs three and four. This could potentially indicate that despite

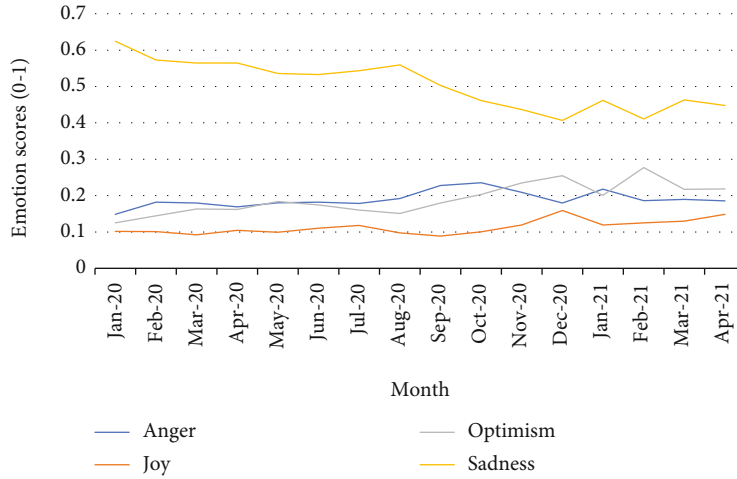


FIGURE 3: Chart showing monthly average emotion scores across the BBC News (UK) news tweets over time.

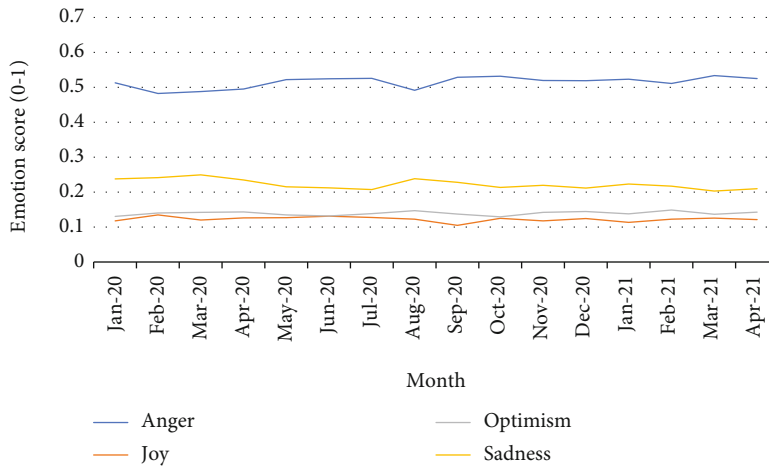


FIGURE 4: Chart showing monthly average emotion scores across the BBC News (UK) user tweets over time. User tweets are reactions to the BBC News (UK) news tweets shown in Figure 3.

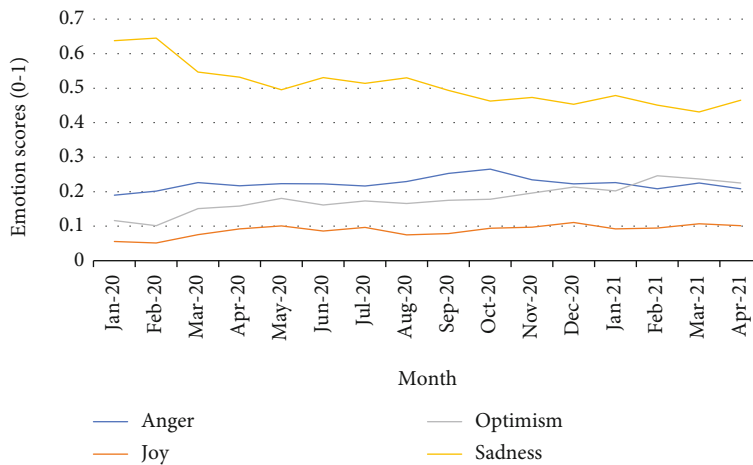


FIGURE 5: Chart showing monthly average emotion scores across The Guardian news tweets over time.

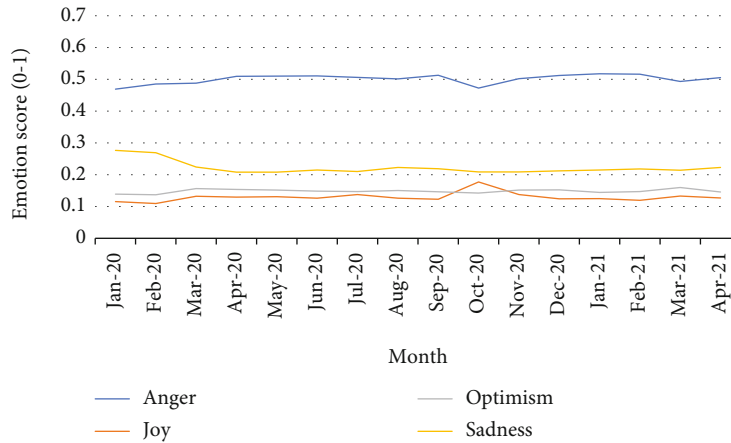


FIGURE 6: Chart showing monthly average emotion scores across The Guardian’s user tweets over time. User tweets are reactions to The Guardian’s news tweets shown in Figure 5.

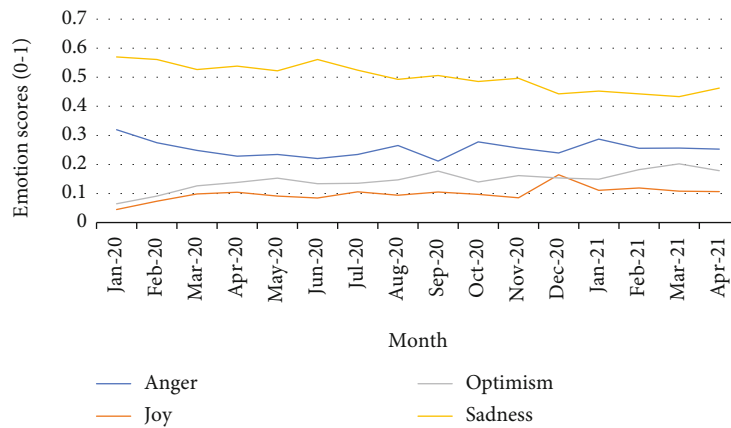


FIGURE 7: Chart showing monthly average emotion scores across the Daily Mail Online news tweets over time.

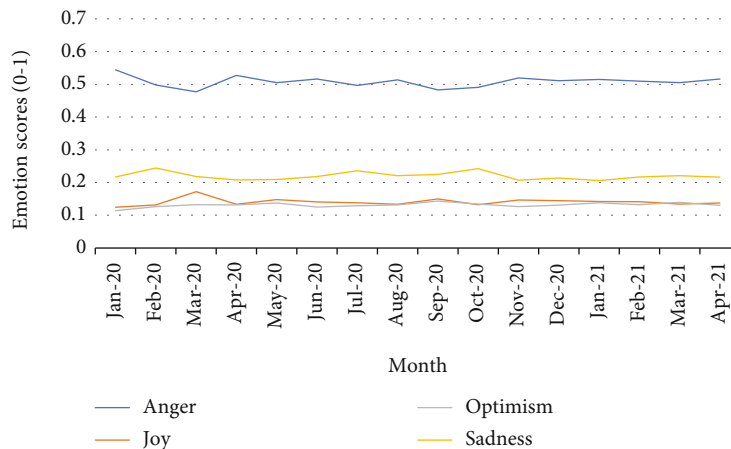


FIGURE 8: Chart showing monthly average emotion scores across the Daily Mail Online user tweets over time. User tweets are reactions to the Daily Mail Online news tweets shown in Figure 7.

infection rates increasing and lockdown measures being reimposed (impacting levels of joy amongst the user tweets), there was nevertheless growing optimism that vaccine roll-out could combat the virus spread, going forward.

Previous research has identified correlations between stress levels and local COVID infection rates [38, 39], with those authors directly attributing this to news reporting. Here, epochs represented key periods in terms of infections

TABLE 8: BBC news (UK) tweet and user tweet dataset descriptives, by epoch.

Emotion	Epoch	BBC news tweets		Corresponding user tweets	
		<i>N</i>	Mean (SD)	<i>N</i>	Mean (SD)
Anger	1	670	.176 (.005)	34054	.487 (.002)
	2	2463	.176 (.004)	184954	.507 (.0008)
	3	1841	.184 (.004)	133330	.516 (.0009)
	4	4533	.204 (.003)	399359	.524 (.0005)
Joy	1	670	.089 (.005)	34054	.117 (.001)
	2	2463	.102 (.003)	184954	.127 (.0006)
	3	1841	.109 (.004)	133330	.127 (.0007)
	4	4533	.124 (.003)	399359	.120 (.0004)
Optimism	1	670	.154 (.007)	34054	.145 (.001)
	2	2463	.171 (.004)	184954	.139 (.0005)
	3	1841	.163 (.004)	133330	.138 (.0005)
	4	4533	.226 (.003)	399359	.140 (.0003)
Sadness	1	670	.251(.010)	34054	.251 (.001)
	2	2463	.285 (.006)	184954	.226 (.0006)
	3	1841	.277 (.006)	133330	.219 (.0007)
	4	4533	.291 (.004)	399359	.216 (.0004)

and lockdown restrictions. The increasing anger levels over time, observed in both the news and user tweets, is likely a reflection of wider sociopolitical factors including frustration with the lockdown measures and dissatisfaction with governmental responses. The topic with the highest anger score was “Preventative Measures,” supporting this. The strong correlations between news and user tweets in terms of emotional classification provide good evidence that news content has been influencing user emotions. However, it should be noted that the direction of causality cannot be inferred based on these data alone. Overall, our prediction that there would be an increase in negative affect was supported, with user anger scores significantly increasing across the four epochs (Table 8); this finding indicates a trend of worsening mental health in terms of anger and stress across 2020/2021, in this sample. However, only anger increased rather than both anger and sadness: encouragingly, sadness was seen to decrease across the epochs. Also, as noted above, there was some indication that despite levels of joy decreasing as the Wave Two of infections took hold, optimism increased. This suggests reliance and a sense of hope that virus spread, and associated mortality, could be curtailed. Topic modelling supports this: the topic with the highest joy and optimism score was “Vaccines and Vaccinations,” which was not prominent in the news until the end of 2020, since this is when the vaccine rollout began in the UK.

Emotions in the news channels selected were significantly positively correlated to the equivalent emotions in the user reactions. This supports the notion that emotional content in the news triggers and/or exacerbates corresponding emotional responses in users, suggesting that emotions projected by news channels online are having an important impact on users.

This study provides unique insight by virtue of its methodology, whereby the user’s comments are analysed directly

in relation to the original news tweet it is a response to. Findings thereby add detail to previous literature showing that high levels of news consumption during the pandemic are linked to increased levels of stress, anxiety, and depression symptoms [56]. Emotions in the news, such as sadness and anger, are positively correlated with the user’s emotional response to it and this could contribute to the higher stress and depression observed in heavy media consumers; this is particularly relevant during the time period under study, where we found that sadness was the most prominent emotion in the news tweets.

Work reviewing research into the phenomenon of “Digital Emotion Contagion” by Goldenberg and Gross [57] highlighted how users’ emotions are influenced by the emotions of content they are interacting with online, and that these emotions spread across users as a result. Also, there is evidence that this effect is actively encouraged by media organisations in that they aim to generate content that generates intense user emotions [58]. Also, in terms of tweet counts and sharing online during the pandemic, Nanath and Joy [59] found that two of the topics most likely to spread amongst users were mental health topics and content with negative emotional impact. They also found that these trends had changed as a result of the pandemic, highlighting the dynamic nature of the online world and the unique impact the pandemic has had on it. The current study adds further insights, showing how topics and emotional patterns have evolved over the time period under study, with correspondence between news and user tweets, highlighting the role of emotional contagion, and the impact of media news reporting, on users’ emotions.

Figures 3–8 highlight the relationships between the emotion levels in the news versus the emotion levels in the corresponding user tweets. Despite the ANOVA tests revealing significant changes in emotion scores and both the news



and users following a pattern of increased anger and decreased sadness, what remains consistent across all channels and throughout the whole period is that the news showed high sadness levels (above all other emotions), and in the user tweets, the prevailing emotion was anger. Thus, it seems that news outlets might exaggerate negativity in their headlines: historically, news organisations have reported more negative news in print and on television in order to attract readers [60]; evidence suggests that this policy also applies to online news reporting [58]. However, it seems that from the data here, this has taken the form of a bias towards sadness rather than anger in the context of the pandemic, while user tweets have instead trended towards anger in response. Theories of emotion highlight the differences between anger and sadness regarding their cognitive, physiological, and behavioural responses. Both emotions would be classed as negative in affect, but evidence points to them having different outcomes, with anger being linked to increased physiological activity and a drive to overcome obstacles and sadness being linked to reduced activity and a focus on goal failure [61]. Anger has been linked to dominance and control, which also impacts how a person confronts risk, leading to more optimistic judgments [62, 63]. In a world of uncertainty, fear, risk, and a lack of control, it is understandable why people are perhaps more inclined to express and share content involving anger rather than sadness, as it is a more proactive response that counteracts feelings of helplessness. However, compared to sadness, anger causes a heightened physiological stress response, activating the autonomic nervous system and triggering increases in blood pressure and heart rate [64]. Chronic stress and anger have a range of negative consequences for health, Yaribeygi et al. [65] highlight effects across the cardiovascular, gastrointestinal, and endocrine systems, brain, and immune system functioning. This shows the negative effects of prevailing anger could be having on users. To manage this risk, strategies could include targeted messaging on media platforms to protect the wellbeing of users. For example, reminding users to take time away from the screen, or the importance of self-care and techniques to manage stress could be helpful.

We identified consistently high anger levels amongst UK Twitter users over the time period studied, and this is in line with some previous work. Li et al. [39] analysed the affective trajectories of users based on Twitter and Weibo data from Jan to May 2020. Their results showed that in China, anger intensity first went up steeply at the initial stage of the outbreak and then had another sharp increase around February 8th: this peak was attributed to the death of Wenliang Li, the ophthalmologist who issued the warnings about the virus. The intensity of anger then gradually decreased, with no further spikes. In the USA, their data showed that anger levels were low initially but showed a rapid rise from the beginning of March 2020; analyses suggested that the topics of “Trump,” “lockdown,” and “government” were the key triggers for this. Basile et al. [38] studied emotional content in Reddit posts amongst users from 5 European countries, and New York City, for the period Jan to June 2020. They found that the UK showed high levels of anger compared

to the other European countries, supporting our findings, noting that expression of anger in both the UK and New York City was the highest amongst the countries included, with values increasing across the time period under study. They found a similar pattern for fear. Moreover, there was a decrease in trust over the time period, and levels of trust were the lowest amongst all the considered countries. Thus, they argue that because both UK and NYC adopted quite a late and initially soft approach to managing the pandemic, the patterns of emotional responses from users in these locations reflect anger directed towards the late response of their governments, to the crisis. As discussed above, emotion contagion amongst users [57], and the ways that online news perpetuates and exacerbates distress and worry [16], also likely contributes to the pattern of findings identified here and other comparable studies.

This study has identified a lack of balance in the news tweets regarding positive and negative affect. The news channel content contained consistently high levels of sadness, followed by anger, across the entire 16-month period (Figure 1). Levels of joy and optimism remained very low. Of course, during a global health crisis, perhaps these emotion levels reflect the stark reality, and it would be unlikely that the news could have reported the pandemic-related events without sadness and anger prevailing. However, in the UK, the vaccine rollout began at the very end of 2020, which perhaps underlies the significant increases in user optimism seen in the last time epoch. In the news, joy and optimism levels did increase across the epochs; nevertheless, sadness remained far more prominent. Based on the correlation analyses, we provide further evidence around how negative emotional news channel content links to emotions expressed by users in response. With increasing numbers of people using platforms like Twitter to join the online community and to access the news (a trend exacerbated by pandemic-related social restrictions), awareness of how media organisations impact user emotions and mental health is important; follow-up longitudinal work is necessary to ascertain causality behind the correlations identified here and study longer-term effects. Further work is also needed to generalise findings to other countries. Here, user tweets were assumed to reflect UK sentiment due to being comments on UK news; however, we did not collect personal demographic information on the users, so it is likely that a proportion of the user comments were from individuals outside of the UK. Also, while our model is effective at classifying core emotions, follow-up work could take this further to focus on mental health problems associated with sadness and anger. This was seen in the work by Li et al. [39] where the lexicon from the PHQ-9 was used to identify emotional content indicating possible clinical depression. Earlier work on the news and trauma, as outlined in the introduction, found that victims of trauma are more likely to consume the news, and this then perpetuates their trauma [16]. Thus, it should be noted that the current findings around the user comments might have been biased towards individuals more negatively impacted by the pandemic, which then resulted in more engagement with online news and stronger negative sentiment expressed. On the other hand, a strength of this

study is that we utilised a large corpus of user comments and considered a range of major UK news outlets, enhancing generalisability. In the UK, 54% of the online population uses Twitter [4]: there are approximately 17 million Twitter users in the UK, representing around a quarter of the UK population. However, Mellon and Prosser [66] found that Twitter users are not representative of the general UK population, in terms of age, political opinions, and education levels; also, it should be noted that those who comment on the posts of news organisations represent a specific subset of Twitter users. Therefore, it is important to recognise that despite analysing over one million user comments, these might not represent the emotions of the general UK population, but rather a specific subset.

News organisations during a time of crisis are central to how we react, behave, feel, and understand the ever-changing situation around us; they are an important part of how society experiences, and responds to, a crisis. Therefore, it is important to generate understanding and awareness of how emotive language impacts the readership and spreads through emotion contagion. Protecting the mental health of users, particularly during a time of crisis where, across all age ranges and populations, the prevalence of mental health issues has increased [67], is a priority. This study adds to the literature and is novel in that we show correlations between emotions in news content and the emotions in the direct comments made by users. It has provided strong evidence of a relationship between the two, across various major UK news channels, thus demonstrating the impact of news reporting on how users feel in response: thus, strategies that consider the content and consumption of online news content by users could be useful for protecting or improving the mental health of populations as we move into the later phases of the pandemic. The study furthers understanding around the timelines and corelations between news and user emotional content during 2020/2021, which is essential to encourage better practice by users, news organisations, tech companies, and regulatory bodies, to mitigate the potentially harmful consequences of heavy online news consumption on societal wellbeing.

## Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgments

This research was supported by grants from the University Global Partnership Network (UGPN; <https://ugpn.org/>) and the Regional Collaborations Programme COVID-19 Digital Grants (Australian Academy of Science), awarded to DM and SE.

## Supplementary Materials

Supplementary Table 1: average emotion scores by news channel. (*Supplementary Materials*)

## References

- [1] N. O'Shea, *COVID-19 and the Nation's Mental Health: Forecasting Needs and Risks in the UK*, The Centre for Mental Health, 2020.
- [2] N. Salari, A. Hosseinian-Far, R. Jalali et al., "Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis," *Globalization and Health*, vol. 16, no. 1, p. 57, 2020.
- [3] N. Pai and S.-L. Vella, "COVID-19 and loneliness: a rapid systematic review," *The Australian and New Zealand Journal of Psychiatry*, vol. 55, no. 12, pp. 1144–1156, 2021.
- [4] Ofcom, "Online Nation: 2020 Report," 2020, [https://www.ofcom.org.uk/\\_data/assets/pdf\\_file/0027/196407/online-nation-2020-report.pdf](https://www.ofcom.org.uk/_data/assets/pdf_file/0027/196407/online-nation-2020-report.pdf).
- [5] E. Lisitsa, K. S. Benjamin, S. K. Chun, J. Skalisky, L. E. Hammond, and A. H. Mezulis, "Loneliness among young adults during COVID-19 pandemic: the mediational roles of social media use and social support seeking," *Journal of Social and Clinical Psychology*, vol. 39, no. 8, pp. 708–726, 2020.
- [6] R. Jones, D. Mougouei, and S. L. Evans, "Understanding the emotional response to COVID-19 information in news and social media: a mental health perspective," *Human Behavior and Emerging Technologies*, vol. 3, no. 5, pp. 832–842, 2021.
- [7] P. Best, R. Manktelow, and B. Taylor, "Online communication, social media and adolescent wellbeing: a systematic narrative review," *Children and Youth Services Review*, vol. 41, pp. 27–36, 2014.
- [8] N. Berry, F. Lobban, M. Belousov, R. Emsley, G. Nenadic, and S. Bucci, "#WhyWeTweetMH: understanding why people use twitter to discuss mental health problems," *Journal of Medical Internet Research*, vol. 19, no. 4, p. e107, 2017.
- [9] J. M. First, H. Shin, Y. S. Ranjit, and J. B. Houston, "COVID-19 stress and depression: examining social media, traditional media, and interpersonal communication," *Journal of Loss and Trauma*, vol. 26, no. 2, pp. 101–115, 2021.
- [10] M. Liu, H. Zhang, and H. Huang, "Media exposure to COVID-19 information, risk perception, social and geographical proximity, and self-rated anxiety in China," *BMC Public Health*, vol. 20, no. 1, p. 1649, 2020.
- [11] A. Ø. Geirdal, M. Ruffolo, J. Leung et al., "Mental health, quality of life, wellbeing, loneliness and use of social media in a time of social distancing during the COVID-19 outbreak. A cross-country comparative study," *Journal of Mental Health*, vol. 30, no. 2, pp. 148–155, 2021.
- [12] H. Liu, W. Liu, V. Yoganathan, and V.-S. Osburg, "COVID-19 information overload and generation Z's social media discontinuance intention during the pandemic lockdown," *Technological Forecasting and Social Change*, vol. 166, article 120600, 2021.
- [13] E. A. Holman, D. R. Garfin, and R. C. Silver, "Media's role in broadcasting acute stress following the Boston Marathon bombings," *Proceedings of the National Academy of Sciences*, vol. 111, no. 1, pp. 93–98, 2014.
- [14] E. A. Holman, D. R. Garfin, P. Lubens, and R. C. Silver, "Media exposure to collective trauma, mental health, and functioning:

- does it matter what you see?," *Psychological Science*, vol. 8, no. 1, pp. 111–124, 2020.
- [15] R. C. Silver, E. A. Holman, J. P. Andersen, M. Poulin, D. N. McIntosh, and V. Gil-Rivas, "Mental- and physical-health effects of acute exposure to media images of the September 11, 2001, attacks and the Iraq War," *Psychological Science*, vol. 24, no. 9, pp. 1623–1634, 2013.
- [16] R. R. Thompson, N. M. Jones, E. A. Holman, and R. C. Silver, "Media exposure to mass violence events can fuel a cycle of distress," *Advances*, vol. 5, no. 4, 2019.
- [17] K. Stainback, B. N. Hearne, and M. M. Trieu, "COVID-19 and the 24/7 news cycle: does COVID-19 news exposure affect mental health?," *Socius: Sociological Research for a Dynamic World*, vol. 6, article 237802312096933, 2020.
- [18] J. Gao, P. Zheng, Y. Jia et al., "Mental health problems and social media exposure during COVID-19 outbreak," *PLoS One*, vol. 15, no. 4, article e0231924, 2020.
- [19] A. Moghanibashi-Mansourieh, "Assessing the anxiety level of Iranian general population during COVID-19 outbreak," *Asian Journal of Psychiatry*, vol. 51, article 102076, 2020.
- [20] N. Zhao and G. Zhou, "Social media use and mental health during the COVID-19 pandemic: moderator role of disaster stressor and mediator role of negative affect," *Applied Psychology. Health and Well-Being*, vol. 12, no. 4, pp. 1019–1038, 2020.
- [21] A. Abd-Alrazaq, D. Alhuwail, M. Househ, M. Hamdi, and Z. Shah, "Top concerns of tweeters during the COVID-19 pandemic: infoveillance study," *Journal of Medical Internet Research*, vol. 22, no. 4, article e19016, 2020.
- [22] S. Boon-Itt and Y. Skunkan, "Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study," *JMIR Public Health and Surveillance*, vol. 6, no. 4, 2020.
- [23] R. Chandrasekaran, V. Mehta, T. Valkunde, and E. Moustakas, "Topics, trends, and sentiments of tweets about the COVID-19 pandemic: temporal infoveillance study," *Journal of Medical Internet Research*, vol. 22, no. 10, article e22624, 2020.
- [24] A. Chang, P. J. Schulz, S. Tu, and M. T. Liu, "Communicative blame in online communication of the COVID-19 pandemic: computational approach of stigmatizing cues and negative sentiment gauged with automated analytic techniques," *Journal of Medical Internet Research*, vol. 22, no. 11, article e21504, 2020.
- [25] S. Das and A. Dutta, "Characterizing public emotions and sentiments in COVID-19 environment: a case study of India," *Journal of Human Behavior in the Social Environment*, vol. 31, no. 1-4, pp. 154–167, 2021.
- [26] C. de Las Heras-Pedrosa, P. Sánchez-Núñez, and J. I. Peláez, "Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems," *International Journal of Environmental Research and Public Health*, vol. 17, no. 15, p. 5542, 2020.
- [27] M. Hung, E. Lauren, E. S. Hon et al., "Social network analysis of COVID-19 sentiments: application of artificial intelligence," *Journal of Medical Internet Research*, vol. 22, no. 8, article e22590, 2020.
- [28] Q. Liu, Z. Zheng, J. Zheng et al., "Health communication through news media during the early stage of the COVID-19 outbreak in China: digital topic modeling approach," *Journal of Medical Internet Research*, vol. 22, no. 4, article e19118, 2020.
- [29] R. J. Medford, S. N. Saleh, A. Sumarsono, T. M. Perl, and C. U. Lehmann, "An "infodemic": leveraging high-volume Twitter data to understand early public sentiment for the coronavirus disease 2019 outbreak," *Infectious Diseases*, vol. 7, no. 7, 2020.
- [30] A. Sesagiri Raamkumar, S. G. Tan, and H. L. Wee, "Measuring the outreach efforts of public health authorities and the public response on Facebook during the COVID-19 pandemic in early 2020: cross-country comparison," *Journal of Medical Internet Research*, vol. 22, no. 5, article e19334, 2020.
- [31] T. Tan, T. Huang, X. Wang, and Z. Zuo, "A preliminary investigation of COVID-19 transmission in the United States by incorporating social media sentiments," *Proceedings of the Association for Information Science and Technology*, vol. 57, no. 1, p. e370, 2020.
- [32] J. Xue, J. Chen, C. Chen, C. Zheng, S. Li, and T. Zhu, "Public discourse and sentiment during the COVID 19 pandemic: using latent Dirichlet allocation for topic modeling on Twitter," *PLoS One*, vol. 15, no. 9, article e0239441, 2020.
- [33] Y. Zhao, S. Cheng, X. Yu, and H. Xu, "Chinese public's attention to the COVID-19 epidemic on social media: observational descriptive study," *Journal of Medical Internet Research*, vol. 22, no. 5, article e18825, 2020.
- [34] P. P. Iglesias-Sánchez, G. F. Vaccaro Witt, F. E. Cabrera, and C. Jambriño-Maldonado, "The contagion of sentiments during the COVID-19 pandemic crisis: the case of isolation in Spain," *International Journal of Environmental Research and Public Health*, vol. 17, no. 16, p. 5918, 2020.
- [35] D. M. Low, L. Rumker, T. Talkar, J. Torous, G. Cecchi, and S. S. Ghosh, "Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on Reddit during COVID-19: observational study," *Journal of Medical Internet Research*, vol. 22, no. 10, article e22635, 2020.
- [36] W. Shi, D. Liu, J. Yang, J. Zhang, S. Wen, and J. Su, "Social bots' sentiment engagement in health emergencies: a topic-based analysis of the COVID-19 pandemic discussions on Twitter," *International Journal of Environmental Research and Public Health*, vol. 17, no. 22, p. 8701, 2020.
- [37] D. Valdez, M. ten Thij, K. Bathina, L. A. Rutter, and J. Bollen, "Social media insights into US mental health during the COVID-19 pandemic: longitudinal analysis of Twitter data," *Journal of Medical Internet Research*, vol. 22, no. 12, article e21418, 2020.
- [38] V. Basile, F. Cauteruccio, and G. Terracina, "How dramatic events can affect emotionality in social posting: the impact of COVID-19 on Reddit," *Future Internet*, vol. 13, no. 2, p. 29, 2021.
- [39] D. Li, H. Chaudhary, and Z. Zhang, "Modeling spatiotemporal pattern of depressive symptoms caused by COVID-19 using social media data mining," *International Journal of Environmental Research and Public Health*, vol. 17, no. 14, 2020.
- [40] F. Aslam, T. M. Awan, J. H. Syed, A. Kashif, and M. Parveen, "Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak," *Communications*, vol. 7, no. 1, 2020.
- [41] K. Krawczyk, T. Chelkowski, D. J. Laydon et al., "Quantifying online news media coverage of the COVID-19 pandemic: text mining study and resource," *Journal of Medical Internet Research*, vol. 23, no. 6, article e28253, 2021.
- [42] T. de Melo and C. M. S. Figueiredo, "Comparing news articles and tweets about COVID-19 in Brazil: sentiment analysis and topic modeling approach," *JMIR Public Health and Surveillance*, vol. 7, no. 2, article e24585, 2021.



- [43] S. N. B. K. Tune, M. M. H. Sarkar, M. N. Uddin, M. Z. H. Pinto, and S. M. Ahmed, "Reactions to the media coverage during the first wave of COVID-19 in Bangladesh," *Heliyon*, vol. 7, no. 10, article e08132, 2021.
- [44] C. I. Kit, J. Heyl, L. Nisha, F. Gabriel, and G. Zara, "Evaluation of Twitter data for an emerging crisis: an application to the first wave of COVID-19 in the UK," *Scientific Reports*, vol. 11, no. 1, 2021.
- [45] N. Alsaedi, P. Burnap, and O. Rana, "Can we predict a riot? Disruptive event detection using Twitter," *ACM Transactions on Internet Technology*, vol. 17, no. 2, pp. 1–26, 2017.
- [46] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, no. null, pp. 993–1022, 2003.
- [47] McCallum, "Mallet: a machine learning for language toolkit," 2002, <http://mallet.cs.umass.edu>.
- [48] S. Syed and M. Spruit, "Full-text or abstract? Examining topic coherence scores using latent Dirichlet allocation," in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Tokyo, Japan, 2017.
- [49] Y. Liu, M. Ott, N. Goyal et al., "RoBERTa: a robustly optimized BERT Pretraining approach," 2019, <http://arxiv.org/abs/1907.11692>.
- [50] S. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko, *SemEval-2018 Task 1: Affect in Tweets*, Association for Computational Linguistics, New Orleans, Louisiana, 2018.
- [51] F. Barbieri, J. Camacho-Collados, L. Espinosa Anke, and L. Neves, "TweetEval: unified benchmark and comparative evaluation for tweet classification," in *Findings of the Association for Computational Linguistics: EMNLP 2020*, Association for Computational Linguistics, 2020.
- [52] K. Takahashi, K. Yamamoto, A. Kuchiba, and T. Koyama, "Confidence interval for micro-averaged F1 and macro-averaged F1 scores," *Applied Intelligence*, vol. 52, no. 5, pp. 4961–4972, 2022.
- [53] M. Smith, *How left or right-wing are the UK's newspapers?*, YouGov, 2017, <https://yougov.co.uk/topics/politics/articles-reports/2017/03/07/how-left-or-right-wing-are-uks-newspapers>.
- [54] Y. Su, P. Wu, S. Li, J. Xue, and T. Zhu, "Public emotion responses during COVID-19 in China on social media: an observational study," *Human Behavior and Emerging Technologies*, vol. 3, no. 1, pp. 127–136, 2021.
- [55] Office for National Statistics, *Coronavirus (COVID-19) Infection Survey Technical Article: Waves and Lags of COVID-19 in England*, Office for National Statistics, 2021, June 2021.
- [56] S. M. A. Shah, D. Mohammad, M. F. H. Qureshi, M. Z. Abbas, and S. Aleem, "Prevalence, psychological responses and associated correlates of depression, anxiety and stress in a global population, during the coronavirus disease (COVID-19) pandemic," *Community Mental Health Journal*, vol. 57, no. 1, pp. 101–110, 2021.
- [57] A. Goldenberg and J. J. Gross, "Digital emotion contagion," *Trends in Cognitive Sciences*, vol. 24, no. 4, pp. 316–328, 2020.
- [58] A. K. Bellovary, N. A. Young, and A. Goldenberg, "Left- and right-leaning news organizations use negative emotional content and elicit user engagement similarly," *Affective Science*, vol. 2, no. 4, pp. 391–396, 2021.
- [59] K. Nanath and G. Joy, "Leveraging Twitter data to analyze the virality of COVID-19 tweets: a text mining approach," *Behaviour & Information Technology*, vol. 42, no. 2, pp. 196–214, 2023.
- [60] G. Lengauer, F. Esser, and R. Berganza, "Negativity in political news: a review of concepts, operationalizations and key findings," *Journalism*, vol. 13, no. 2, pp. 179–202, 2012.
- [61] H. C. Lench, T. P. Tibbett, and S. W. Bench, "Exploring the toolkit of emotion: what do sadness and anger do for us?," *Compass*, vol. 10, no. 1, pp. 11–25, 2016.
- [62] J. S. Lerner and D. Keltner, "Beyond valence: toward a model of emotion-specific influences on judgement and choice," *Cognition & Emotion*, vol. 14, no. 4, pp. 473–493, 2000.
- [63] J. S. Lerner and D. Keltner, "Fear, anger, and risk," *Journal of Personality and Social Psychology*, vol. 81, no. 1, pp. 146–159, 2001.
- [64] G. V. Bodenhausen, L. A. Sheppard, and G. P. Kramer, "Negative affect and social judgment: the differential impact of anger and sadness," *European Journal of Social Psychology*, vol. 24, no. 1, pp. 45–62, 1994.
- [65] H. Yaribeygi, Y. Panahi, H. Sahraei, T. P. Johnston, and A. Sahebkar, "The impact of stress on body function: a review," *EXCLI Journal*, vol. 16, pp. 1057–1072, 2017.
- [66] J. Mellon and C. Prosser, "Twitter and Facebook are not representative of the general population: political attitudes and demographics of British social media users," *Research & Politics*, vol. 4, no. 3, article 205316801772000, 2017.
- [67] World Health Organisation, *Mental health and COVID-19: early evidence of the pandemic's impact: scientific brief, 2 March 2022*, World Health Organisation, 2022, <https://apps.who.int/iris/handle/10665/352189>.