RESEARCH

Personality trait analysis during the COVID-19 pandemic: a comparative study on social media

Marcos Fernández-Pichel[1] · Mario Ezra Aragón[1] · Julián Saborido-Patiño[1] · David E. Losada[1]

© The Author(s) 2023

Abstract

The COVID-19 pandemic, a global contagion of coronavirus infection caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), has triggered severe social and economic disruption around the world and provoked changes in people's behavior. Given the extreme societal impact of COVID-19, it becomes crucial to understand the emotional response of the people and the impact of COVID-19 on personality traits and psychological dimensions. In this study, we contribute to this goal by thoroughly analyzing the evolution of personality and psychological aspects in a large-scale collection of tweets extracted during the COVID-19 pandemic. The objectives of this research are: i) to provide evidence that helps to understand the estimated impact of the pandemic on people's temperament, ii) to find associations and trends between specific events (e.g., stages of harsh confinement) and people's reactions, and iii) to study the evolution of multiple personality aspects, such as the degree of introversion or the level of neuroticism. We also examine the development of emotions, as a natural complement to the automatic analysis of the personality dimensions. To achieve our goals, we have created two large collections of tweets (geotagged in the United States and Spain, respectively), collected during the pandemic. Our work reveals interesting trends in personality dimensions, emotions, and events. For example, during the pandemic period, we found increasing traces of introversion and neuroticism. Another interesting insight from our study is that the most frequent signs of personality disorders are those related to depression, schizophrenia, and narcissism. We also found some peaks of negative/positive emotions related to specific events.

Keywords COVID-19 · Social media · Big-5 · Personality analysis · Emotion analysis

1 Introduction

The Coronavirus disease outbreak (COVID-19) began in December 2019 and rapidly spread worldwide, causing millions of deaths (Zhu et al., 2020). Most people infected with the virus

---

✉ Marcos Fernández-Pichel
  marcosfernandez.pichel@usc.es

1 Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), University of Santiago de Compostela, Rúa Jenaro de la Fuente, Santiago de Compostela 15782, A Coruña, Spain

experienced mild to moderate respiratory illness and recovered without requiring special treatment. However, some individuals developed serious symptoms and required medical attention (Organization, 2023). Older people and those with underlying medical conditions were more likely to develop severe illnesses. COVID-19 has caused a huge impact worldwide on health systems, economies, and education systems (Organization, 2020). At this point, we still do not know how long this damage will last or when the world will fully recover. The pandemic had a major effect on our lives, and many people have been facing stressful and overwhelming challenges. These stress factors can cause changes in appetite, energy, desires, and interests, and can create feelings of fear, anger, sadness, worry, numbness, or frustration. Learning to identify these symptoms can help us to cope and manage healthily.

The increasing social media activity presents an opportunity to study the evolution of behavioral and personality traits in a large proportion of the world population. For many people, social life does not only happen in their surroundings or immediate environment. In many cases, a considerable number of social interactions take place in virtual settings created by social media platforms like Facebook, Twitter, or Reddit. Since the beginning of the pandemic, social media platforms are increasingly being used as an information source (e.g., to be informed about risks and crises) and publication channel (Reuters, 2022).

Mental health has been a major global concern within the last decades, but psychological problems have aggravated since the confinement period of the year 2020 (Gupta et al., 2020). Having this in mind, the main motivation of this work lies in the necessity to further study the growth of this "psychological pandemic" and to raise public awareness. For example, by developing new screening tools able to inform authorities and the population about the impact of the pandemic on personality dimensions. It is also important to match psychological traits with specific events or circumstances that happened during this difficult period (e.g., related to harsh confinements). In addition, analyzing the evolution of feelings and emotions is a natural complement to the study of personality dimensions.

We will focus on two collections of tweets retrieved during the worst part of the pandemic period (2020-2021). The tweets come from two countries (USA and Spain), and we estimate the prevalence and evolution of specific personality traits and disorders on these two large samples of social media entries (10,234,223 and 17,395,598 tweets, respectively). To that end, we draw from recognized psychological instruments, such as the Big-5, which defines five core personality factors (extraversion, agreeableness, openness to experience, conscientiousness, and neuroticism) (Barrick and Mount, 1991). We also report on hot topics found in the published content, with a particular focus on themes that have triggered concerns, polarized emotions, or extreme feelings. We can summarize our main contributions as follows:

1. Two publicly available datasets of tweets (about general topics) extracted during the most problematic months of the COVID-19 pandemic. The collections contain tweets written in English and Spanish and the tweets are geotagged in USA or Spain, respectively. We also make available to the community a repository to facilitate the reproducibility of the experiments presented in this study.[1]

2. A thorough analysis of the prevalence and evolution of personality traits in Twitter users from two different countries (USA and Spain).

3. A study of the evolution of signs of personality disorders (Schizoid, Depressive, Avoidant, Dependent, Histrionic, Narcissistic, Compulsive, Paranoid, and Schizotypal) during the pandemic.

---

[1] https://github.com/MarcosFP97/COVID-19-Personality

4. A complementary report of emotions estimated from the tweets published during the pandemic.

The remainder of the paper is organized as follows. Section 2, presents a brief overview of the related work. Section 3, describes in detail our methodology to create the data and our approaches to analyze the publications. Section 4 reports the results and Section 5 presents a discussion of the obtained results and limitations of this work. Finally, Section 6 gives some concluding remarks.

2 Related work

In this section, we discuss some recent works related to social media analysis during the pandemic period. Different studies have performed research on the impact of COVID-19 on people's public and personal lives. One of the most popular strategies consists of mining social media data and measuring concerns or sentiments (e.g., through publications posted by people on their social media accounts). In Lyu et al. (2020), the authors searched for tweets that contained terms referring to the COVID-19 pandemic, such as "Chinese virus" or "Wuhanvirus", and then compared them with tweets that did not have these words. This study revealed differences in age, geolocation, and political views among the individuals who posted these two types of tweets.

Vohra and Garg (2023) performed sentiment analysis on Twitter data containing keywords related to working from home. A total number of 358,823 tweets were collected. The authors first labeled a subset of tweets using VADER and, next, trained a Convolutional Neural Network (CNN) to estimate sentiment. They found that the majority of posts had a positive sentiment towards working from home.

An interesting study that analyzes population-level disparities during the pandemic was presented in Zhang et al. (2021). The authors first collected a large-scale dataset from Twitter. Then, they divided the data among population groups and extracted concerns, sentiments, and emotions. This work revealed differences regarding COVID-19 topics among population groups (for example, gender, and age). One of the findings was that the population group composed of above 40-year-old women was the group most concerned about COVID-19. This segment of females was most concerned about economics and education, while males in the same age group were most concerned about politics and economics.

In another study, Alhuzali et al. (2022) collected tweets related to COVID-19 from 48 different cities in the United Kingdom. The sentiment, emotion, and topics of the tweets were examined using deep learning models like SenticNet 6 and SpanEmo, with a combined topic modeling approach. The authors found that people's attitudes and expressions were highly positive at the beginning of 2020, but started to decrease over time toward the end of 2021. This approach can potentially supply valuable cues about how public policies are perceived by people in different geographical areas.

In a slightly different direction, Umair and Masciari (2023) explored the sentiments of the citizens related to the COVID-19 vaccination. This study performed sentiment and spatial analyses of textual pieces of evidence. The authors used the TextBlob tool to estimate the polarity of the texts and categorized them. They also mapped the data obtained to geographical locations to get a global picture of people's attitudes toward vaccination. Similarly, in Borah (2023) a multimodal deep learning method for Indian tweets classification was introduced. The goal was to estimate population's hesitancy towards vaccination through social media posts. The study revealed that confidence in vaccination increased with time.

Si et al. (2021) surveyed post-traumatic stress syndrome (PTSS) and sleep quality among residents in Wuhan and nearby cities. To that end, a PTSS checklist for DSM-5 (Diagnostic and Statistical Manual of Mental Disorders) and four tools from the Pittsburgh Sleep Quality Index (PSQI) were utilized. The survey revealed a prevalence of PTSS after the COVID-19 outbreak and showed that women suffered more than men. More specifically, an adverse alteration in cognition and sleep quality was associated with females. The authors also found that participants with better sleep quality reported lower PTSS.

In Ahmad and Murad (2020), studied how social media affects mental health and the spread of panic about COVID-19 in the Kurdistan Region of Iraq. They found that social media has a significant impact on spreading fear and panic, and influences negatively people's mental health. Young people, aged between 18 and 35, were found to be facing psychological anxiety.

Researchers in Sert et al. (2022) investigated the topic evolution of Twitter for the Republic of Turkey in 2020. They analyzed 1.3 million tweets related to the coronavirus between February 24, 2020, and May 2, 2020. The authors found an increase in topics related to hygiene, lifestyle, and anxiety. Inspired by the World Bank's Poverty Monitoring Technical Note, they also examined the effect of income on content sharing. According to the authors' estimates, users with a lower gross domestic product per capita tend to share more news related to COVID-19 content.

In Ainley et al. (2021), the reaction of people to health and care delivery in the United Kingdom was studied. The authors identified the following main themes: access to remote care, quality of remote care, the anticipation of remote care, online booking, asynchronous communication, and publicizing changes to services or care delivery. At the beginning of the pandemic, the comments were positive but this positive trend decreased over time with the increase in restrictions and COVID cases.

In the area of personality analysis, we can also find approaches to leverage data that is shared on the Internet on a daily basis. For example, deep learning approaches show promise in personality classification when an ample amount of data is available for training (Leonardi et al., 2020; Di et al., 2018). However, their limited explanatory understanding capabilities hinder specialists from comprehending the rationale behind automatic predictions. To try to solve this, de-la Rosa et al. (2023) aimed to analyze language and discern personality traits from short communications by harnessing information extracted from a mental lexicon. The objective was to develop a valuable representation that aids specialists in understanding the key variables that contribute to the analysis of an individual's personality. This highlights the importance of explainability in this type of task.

As we can see above, former studies on social media during the pandemic have mainly focused on sentiment and emotion analysis. These works have essentially explored the concerns that people raise and their importance. Following this line of thought, we expand this type of analysis by studying personality traits and personality disorders. We estimate the prevalence and evolution of personality traits and disorders during the pandemic using Deep Linguistic technologies. Our work also differs from the previous studies by explicitly relating the results obtained to specific social events. To that end, we perform our analysis on a collection of tweets whose size is substantially larger than the sizes considered by previous works. Moreover, the data is geotagged in two different countries and the crawling/extraction process was not biased towards specific keywords. We perceive the broad scope of this resource as a strength, as it provides other research teams with the opportunity to carry out a variety of analyses. It also allows us to study cultural differences and reactions when handling a problem of the magnitude of COVID-19. The collection and the code are available for the community to test and reuse.
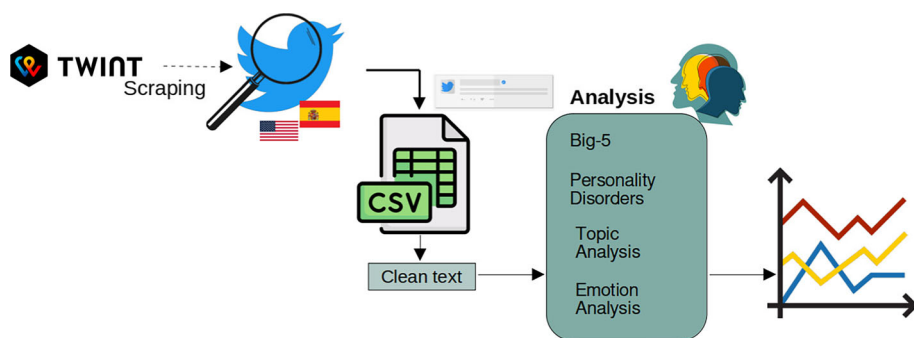
Fig. 1 General view of the system. We extracted Twitter data using Twint (scrapping tool) and, next, we performed multiple reports with complementary analytical tools

3 Methodology

In this section, we present in detail the strategy implemented for collecting tweets and we describe our methods for analyzing personality traits, disorders, and emotions. Figure 1 depicts an overall view of all the steps involved.

3.1 Data collection

For collecting data, we employed Twint.[2] Twint offers the possibility of extracting tweets in a geolocalized manner. The process consists of choosing a central geographical point, establishing a radius, and obtaining tweets from that area. For our study, the points chosen were Madrid and New York (for representing Spain and USA, respectively). After the extraction, tweets were classified by language,[3] and we only kept the English tweets from the USA dataset and the Spanish tweets from the Spain dataset. For the data processing, we performed a simple pre-processing of the texts by removing special characters like URLs, @, and hashtags.

The USA dataset contains tweets whose publication dates range from January 2020 to August 2021. The total number of tweets is 10,234,223. The Spain dataset contains 17,395,598 tweets (publication dates from January 2020 to May 2021).

A subset of this data is freely available for the community to test and use.[4] We have carefully respected Twitter's policy regarding anonymization and the maximum number of posts. More specifically, at the time of creating this collection, the limits of Twitter established that "you may only distribute up to 1,500,000 Tweet IDs to a single entity within a 30-day period".[5] Those readers interested in the entire dataset can contact the authors, since this is a generic dataset that could be exploited for other research purposes.

It is worth mentioning that we had some difficulties during the course of the extraction stage. The Twint service was shut down during the development of this project. This affected the retrieval of Spanish tweets. More specifically, for Spain's dataset, the number of collected tweets during the central period of the pandemic (around 4,561 tweets per week) are less than

---

[2] https://github.com/twintproject/twint

[3] To that end, we employed https://pypi.org/project/langdetect/.

[4] https://github.com/MarcosFP97/COVID-19-Personality/tree/master/dataset

[5] https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases

the number of collected tweets during the rest of the timeline (around 345,849 tweets/per week for Spain and 159,948 tweets/per week for the USA). This should be taken into account for the analysis, as trends during this period show often a larger variance. Still, the overall collection of tweets is substantially larger than those employed in most existing studies.

3.2 Big-5 personality analysis

The main objective of this work is to understand the personality traits and their evolution in Social Media users during the COVID pandemic. Personality has been defined (Funder, 1997) as the "individual's characteristic patterns of thought, emotion, and behavior, together with the psychological mechanism behind those patterns". However, this definition is too general and makes understanding individual differences in behavior and experience too difficult. In the area of Psychology, research on personality, led by multiple teams of researchers, reached a consensus on a general taxonomy of personality traits, named "Big Five" (John and Srivastava, 1997). Personality analysis (Barrick and Mount, 1991) using Big-5 builds from a taxonomy or classification proposed in the 1980s. This model was defined using factor analysis (Yong and Pearce, 2013), and studies the relationship between a large number of verbal descriptors and personality traits. The original Big-5 model was proposed by Digman, and later extended by Goldberg (1993). These definitions have been found to contain the majority of personality traits and establish a solid structure under which personality can be formally studied.

The five factors identified are:

1. Openness to experience: appreciation of new emotions, arts, or new and imaginative ideas.
2. Conscientiousness: responsibility, a tendency to self-discipline, and obedience.
3. Extraversion: appreciation for a breadth of activities.
4. Agreeableness: sympathy or friendliness, concern for social harmony rather than individualism.
5. Neuroticism: the tendency to experience negative emotions such as anxiety or depression.

Following Neuman and Cohen's 2014 methodology, we build one representation for describing the existence of each factor and another representation for describing the opposite of each factor. In this way, we can analyze texts written by people and, given the evidence found, estimate the level of intensity of both sides of each personality factor. For example, we might track not only the evolution of extraversion (the positive pole of the 3rd personality factor) but also the evolution of introversion, which would be represented as the negative pole of the same personality factor.

Table 1 presents the personality traits and some words related to each factor (or to the opposite of each factor). The selection of words representing each pole comes from previous studies on vectorial semantics for personality assessment, which drew from a list of adjectives related to the Big-5 dimensions (Trapnell and Wiggins, 1990).

Additionally, for the analysis of personality disorders, we employ a similar word-based approach to represent and track each disorder. Table 2 presents the tracked personality disorders and the main adjectives used to represent them (extracted from Millon et al. (2004)).

For our first analysis, our objective was to study how different characteristics of people's personalities have evolved over the period of the pandemic. To that end, we exploit the

Table 1 Personality traits: Extraversion (E) Agreeableness (A) Conscientiousness (C) Neuroticism (N) Openness to experience (O)

traits	representative words
E-POSITIVE	dominant, assertive, authoritarian, forceful, assured, confident, firm, persistent
E-NEGATIVE	nervous, modest, quiet, forceless, afraid, shy, calm, indecisive
A-POSITIVE	tender, gentle, soft, kind, affectionate, helpful, sympathetic, friendly
A-NEGATIVE	cruel, unfriendly, negative, mean, brutal, inconsiderate, insensitive, cold
C-POSITIVE	organized, orderly, tidy, neat, efficient, persistent, systematic, straight, careful, reliable
C-NEGATIVE	distracted, unreliable, incompetent, wild, inefficient, disloyal, chaotic, confused, messy, disorganized
N-POSITIVE	worried, stressed, anxious, nervous, fearful, touchy, guilty, insecure, restless, emotional
N-NEGATIVE	balanced, stable, confident, fearless, calm, easy going, relaxed, secure, comforted, peaceful
O-POSITIVE	philosophical, abstract, imaginative, curious, reflective, literary, questioning, individualistic, unique, open
O-NEGATIVE	narrow-minded, concrete, ordinary, incurious, thoughtless, ignorant, uneducated, common, conventional, restricted

taxonomy of personality traits proposed within the Big 5, and we analyze users' publications according to the five personality factors (and their corresponding opposite poles).

The first step for analyzing user data consists of converting the user's tweets to a vectorial representation. To obtain this representation we employ Sentence-BERT (Reimers and Gurevych, 2019) a modification of the pre-trained BERT network that uses network structures to derive semantically meaningful sentence embeddings. Sentence-BERT was trained on the combination of the SNLI and Multi-genre NLI datasets. The SNLI collection (Bowman et al., 2015) contains sentence pairs annotated with the labels contradiction, entailment, and neutral. The MultiNLI dataset (Williams et al., 2018) is a collection that contains sentence pairs and covers a range of genres of spoken and written text. Sentence-BERT has been shown to be a

Table 2 Personality disorders

disorders	representative words
schizoid	indifferent, apathetic, remote, solitary
depressive	sad, depressed, hopeless, gloomy, fatalistic
avoidant	shy, reflective, embarrassed, anxious
dependent	helpless, incapable, passive, immature
histrionic	dramatic, seductive, shallow, hyperactive, vain
narcissistic	selfish, arrogant, grandiose, indifferent
compulsive	restrained, conscientious, respectful, rigid
paranoid	cautious, defensive, distrustful, suspicious
schizotypal	eccentric, alien, bizarre, absent

solid approach to represent short pieces of text. We feed each tweet to the model to obtain an embedding representation.

The vectorial representation of each personality factor (or its opposite pole) is built from the word list presented in Table 1. Each word is first passed to the language model, which produces an embedding representation of the word. This helps to formally capture the semantics of the word and internally represents the potential contexts in which the word is used. Next, the personality factor (or its opposite) is assigned a vectorial representation that is the average of the vectors of the words associated with the factor. In this way, we can capture, for example, traces of neuroticism (N-positive vector) based on the occurrence of words and expressions semantically related to worries, stress, anxiety, nerves, and so forth.

To estimate the degree of presence of a personality dimension in a tweet, we use the cosine similarity between the text and the dimension's representations. This technique has proven to be very robust in capturing semantic meaning (Laskar et al., 2020). For representing the evolution of personality traits, we plot average similarities per week.

3.3 Topic analysis

An additional level of analysis of the data can be oriented to extract topic-oriented signals. To that end, we have performed a further inspection of the collections with Empath (Fast et al., 2016). Empath is a tool developed at the University of Stanford for understanding topic signals in large collections of text. The original version of Empath was built from a collection of more than 1.8 billion words and it estimates the relationship between words and phrases by exploiting neural embeddings and sophisticated language models. Given an input text, Empath can associate it to 200 built-in, pre-validated categories.

For the extraction of topics, we worked with Empath's default categories and computed the level of presence of each of them in the tweets (for each topic, the library returns a suitability score in [0,1]). Next, we grouped the tweets by week and plotted the average value of the ten most salient categories. This approach helps to reveal the most prominent topics in the collections analyzed.

3.4 Emotion analysis

Further evidence about people's perceptions can be obtained through emotion analysis, a process that consists of identifying and analyzing the underlying emotions expressed in text. Emotions are pervasive among humans and are studied in many fields like Psychology and Neuroscience (Canales and Martínez-Barco, 2014).

Using emotions, we can extract additional clues about people's feelings and concerns. By measuring emotions during the COVID-19 pandemic, we can gain further understanding on the impact of this difficult period on people's lives.

To perform this analysis, we used EmoRoBERTa (Kim and Vossen, 2021), a transformer language model trained using a dataset of Reddit comments. The comments were labeled following a list of 28 emotions. This model has been shown to outperform alternative models in multiple emotion detection tasks and it incorporates a wider variety of emotions. In Table 3, we present some real examples of tweets classified by this model.

Again, as for previous analysis, we have grouped the tweets by week and we report the proportion of tweets per week assigned to each emotion.

Table 3 Examples of tweets classified by EmoRoBERTa

Tweet	Emotion
"Thank you cmtrm for this lovely share"	gratitude
"MidgetTamer Another reason I hate that guy"	anger
"devastating news from my alma mater tonight."	sadness
"Today is gonna be hard."	disappointment
"I can't stop laughing after talking to these guys"	amusement

4 Analysis of results

4.1 Big-5

Let us individually analyze the patterns obtained for each personality dimension. In this Section, we analyze how personality traits vary over time. We first focus on **Extraversion**, which consists of the general tendency to experience positive emotions, such as being friendly, lively, and active. In Fig. 2, we present the evolution of this trait through the pandemic time. In both countries (USA and Spain) the negative pole of this personality factor was dominant (higher line of E-NEGATIVE compared to E-POSITIVE). This result suggests that themes related to nervousness, fear, or shame (E-NEGATIVE words) were more frequent than those related to E-POSITIVE (dominance, confidence, persistence, etc). This pattern happens during the whole period and, thus, it might be a general feature of social media rather than a special feature of the COVID-19 period.

The USA lines are rather flat and, thus, do not reveal significant changes in patterns of extraversion/intraversion. The Spain lines (particularly the E-negative pattern) show a more agitated behavior. For example, we can observe higher evidence of nerves, fear, etc. in the second semester of 2020.

Another interesting outcome of this analysis is that, overall, the presence of this personality factor (regardless of the orientation towards extraversion or intraversion) seems to be higher in Spain. Note also that the difference between extraversion (E-POSITIVE) and intraversion (E-NEGATIVE) is higher in the USA. So, the USA dataset does not appear to contain too many extraversion/intraversion-related tweets but those tweets that are actually related to this personality factor tend to show evidence of nerves, fear, etc.

To further analyze this factor, we took the 1000 tweets that scored highest on E-NEGATIVE (highest similarity to the negative pole of this personality factor) and created a word cloud with the most prominent words (see Fig. 3).[6] In both countries, "nervous" is the dominant word, and the plots also show words related to anxiety and fear. In Spain, we can see that words like "calm" or "quiet" also had a high degree of occurrence. This matches with the results reported in Fig. 2, where we saw that, in Spain, the presence of the E-POSITIVE pole was not insignificant.

Let us move to the second personality trait, **Agreeableness**. Agreeableness is a trait that describes a person's ability to put others' needs before their own. People with more agreeableness are more likely to be empathetic and work with people who need more help. Figure 4 shows the results for this personality factor. In this graph, we can see that, for both countries, the negative polarity of this personality factor is greater than the positive one. In Spain, we can observe an increase for both poles (positive and negative) in the period of June-July 2020,

---

[6] The Spanish tweets have been translated into English and the clouds present high-frequency words after translation.
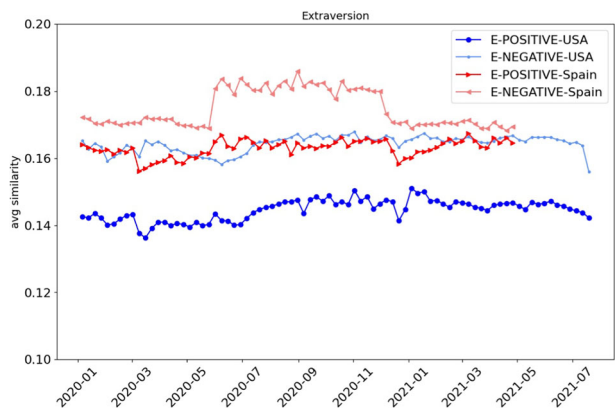
Fig. 2 Extraversion analysis during the COVID-19 pandemic

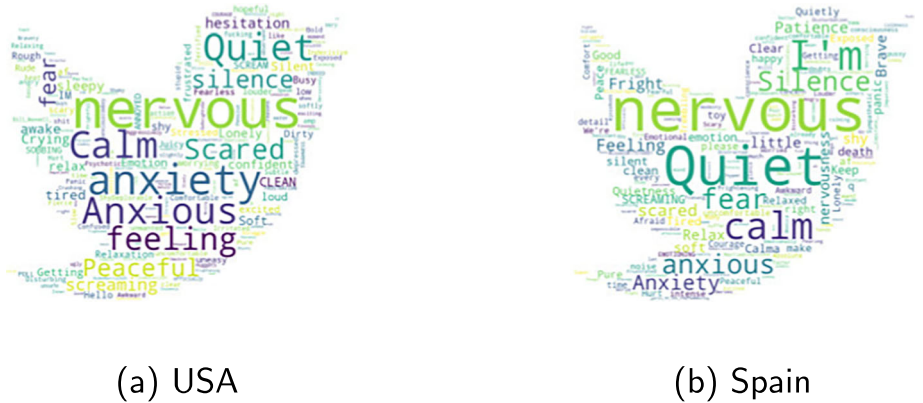(a) USA   (b) Spain

Fig. 3 Extraversion analysis. The tag clouds represent the most common words (after removing stop-words) of the 1000 tweets with the highest similarity to E-NEGATIVE
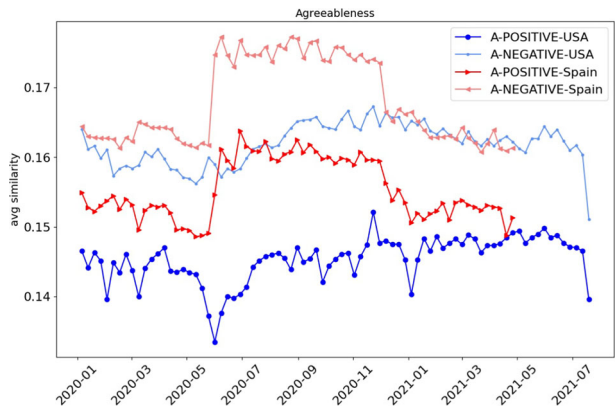

Fig. 4 Agreeableness analysis during the COVID-19 pandemic

stabilizing quickly at these values and then decreasing in January 2021. The USA plot also shows a slight increase of A-NEGATIVE in mid-2020, and this A-NEGATIVE line keeps at high values until June-July 2021. USA's positive pole had a considerable drop in June 2020 and then increases again in the following month and remains unstable throughout the season. Overall, these results suggest a higher prevalence of the negative pole of this personality factor in both countries. The A-NEGATIVE pole is associated with topics of unfriendliness, negativity, coldness, insensitivity, and so forth. In both datasets, this type of negative sentiment increased during a large part of 2020. Furthermore, in the USA we observe a large drop in A-POSITIVE in June 2020. A-POSITIVE is related to gentleness, kindness, friendliness, etc. It seems that the expression of this type of positive feelings decreased rapidly in the middle of 2020 and then it quickly got to levels similar to those of the beginning of the year.
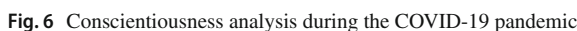
Since the A-NEGATIVE plots are dominant, we proceeded to obtain the 1000 tweets with the highest similarity to the negative polarity of agreeableness. Figure 5 presents the words with the highest frequency. In the USA, the dominant words have to do with pain, hurt, or being rude (words arguably related to inconsideration or coldness). In Spain, the words are also related to these topics, but the words used for expressing these feelings are slightly different. This is a natural outcome since people with different cultures and backgrounds express themselves in different ways.

The third personality factor that we examine is **Conscientiousness**. Conscientious individuals demonstrate a strong work ethic, are reliable and punctual, pay attention to detail, and show commitment and purpose. Figure 6 presents the evolution of this personality trait through the pandemic. Again, the negative pole is dominant in both countries. This evidence suggests that topics about distraction, unreliability, chaos, and so forth (C-NEGATIVE) are more prevalent than those associated with C-POSITIVE (order, organization, reliability, etc). This seems to be a general feature of Social Media (at least characteristic of Twitter) regardless of the period that is analyzed. Indeed, it is not surprising to observe that Twitter content is reflective of chaotic and less structured thoughts. In any case, we can identify some patterns of Conscientiousness during the COVID-19 period. For example, we observe that C-NEGATIVE grew during the second semester of 2020. And both C-NEGATIVE lines show high values at the end of 2020. In contrast, the positive pole of Conscientiousness remained relatively stable over the entire period.
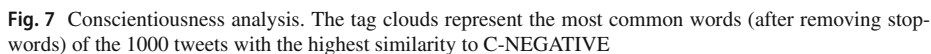
Continuing with the analysis of this personality factor, we took again the 1000 tweets with the highest similarity to C-NEGATIVE (which is the most dominant pole of Conscientiousness) and counted the frequency of occurrence of the constituting words. In Fig. 7, we
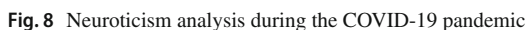

(a) USA   (b) Spain

Fig. 5 Agreeableness analysis. The tag clouds represent the most common words (after removing stop-words) of the 1000 tweets with the highest similarity to A-NEGATIVE
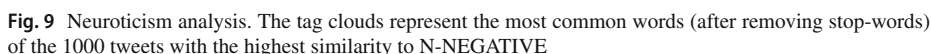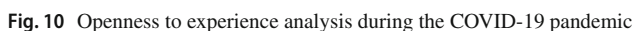
Fig. 6 Conscientiousness analysis during the COVID-19 pandemic

plot the result of this analysis. In USA's word cloud, we can observe that the words with the highest frequency are related to confusion and disorder, and Spain's word cloud also shows words semantically related to those in the USA's word cloud (e.g. referring to the chaos and unexpected events). There are some social, cultural, and language differences between both datasets but, still, both visualizations illustrate expressions about similar concerns.

The fourth personality factor in our analysis is **Neuroticism**. This trait is defined as a tendency toward anxiety, depression, self-doubt, and other negative feelings. A person with neurotic behaviors might be overly dependent on others or ask for help instead of figuring things out for themselves. Figure 8 presents the results of this analysis over time. The first thing we can notice is that the negative pole, N-NEGATIVE, is dominant in both countries. This is good news, as N-NEGATIVE is associated with balance, stability, calm, etc. However, the Spain lines show that N-POSITIVE got to N-NEGATIVE's levels during the second part of 2020. This suggests the presence of significant growth of worries, stress, and anxiety (N-POSITIVE descriptors) in this country during the second semester of 2020. In both countries, we observe a high peak of neuroticism-like expressions in April 2020 (peaks in the four lines).
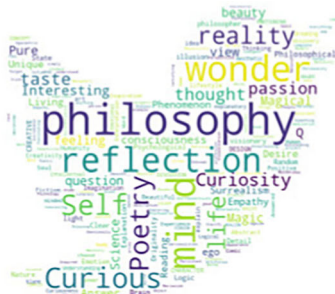

Fig. 7 Conscientiousness analysis. The tag clouds represent the most common words (after removing stopwords) of the 1000 tweets with the highest similarity to C-NEGATIVE

Fig. 8 Neuroticism analysis during the COVID-19 pandemic

This increase is sharp in Spain, but this might be related to the availability of fewer tweets from this period (higher variance of the trends because the sample is smaller). A high value of both N-POSITIVE and N-NEGATIVE is indicative of the occurrence of two main groups of publications: a group of people expressing concerns, worries, nervousness, etc, and another large group of people calling for calm or expressing relaxing and peaceful feelings. This high occurrence of both N-POSITIVE and N-NEGATIVE expressions decreased during the second quarter of 2020 (in both countries) and then it grew again during the second semester of 2020. It is also interesting to note that the two N-POSITIVE lines go down since December 2020. In late 2020 the COVID-19 vaccination process started (e.g. in December 2020 in Spain) and this milestone event may have well induced a significant decrease in concerns and stress.

We took the tweets with the highest similarity to the negative pole of this personality trait (N-NEGATIVE, which is the most prevalent neuroticism pattern) and, again, analyzed the most common words. Figure 9 shows the word clouds built from this sample of 1000 tweets.

(a) USA   (b) Spain

Fig. 9 Neuroticism analysis. The tag clouds represent the most common words (after removing stop-words) of the 1000 tweets with the highest similarity to N-NEGATIVE

Fig. 10 Openness to experience analysis during the COVID-19 pandemic

For both countries, we can notice how the words with the highest presence are related to calm, patience, and peace.

The last personality factor that we analyze is **Openness to experience**. This personality trait is associated with being imaginative and creative, inventive, open to unusual ideas, adventure, and nonconformity. Open-minded individuals enjoy learning and use their knowledge to tackle difficulties in life. Figure 10 shows the results for the Openness to experience trait. For this personality trait, its positive polarity is the one that scores the highest in both countries. This suggests that the proportion of tweets reflecting imaginative and open thoughts is higher than the proportion of tweets reflecting narrow-minded or thoughtless feelings. In general, we can observe few changes over time, the lines remain relatively flat through the entire period. The most remarkable trend is that, in Spain, there seems to be a growth in O-NEGATIVE in the second part of 2020. A plausible explanation could be related to the strict social restrictions imposed in Spain in 2020. These restrictions could have temporally led to a decrease in openness to experience. However, we would need to collect and analyze more data in order to further validate this claim.

Fig. 11 Openness to experience analysis. The tag clouds represent the most common words (after removing stop-words) of the 1000 tweets with the highest similarity to O-POSITIVE

The most prominent pole of this personality factor, O-POSITIVE, is further analyzed in Fig. 11. It represents the most frequent words in the top 1000 tweets that are the most similar to O-POSITIVE. Here, it is interesting to notice that both countries express their openness to experiences in different ways. For example, in the USA, we can see words related to thinking, living, and nature. Meanwhile, in Spain, we can observe topics related to philosophy, reflection, and curiosity.

Table 4 gives an overall view of the main trends derived from the analysis of the personality factors.

4.2 Personality disorders

As explained in Section 3.2, we measure signs of the following personality disorders: Schizoid, Depressive, Avoidant, Dependent, Histrionic, Narcissistic, Compulsive, Paranoid, and Schizotypal.

Figure 12 presents the estimated evolution of these personality disorders through the pandemic period in USA and Spain. The first thing we can observe is that, in both countries, the dominant personality disorders are depressive, schizoid, and narcissistic.

The most frequent personality disorder is the depressive personality, which is characterized by a loss of interest in normal daily activities, a lack of productivity, and a feeling of hopelessness. These symptoms usually conduct to low self-esteem and feeling of inadequacy.

Table 4 Personality factors

extraversion	Negative pole (E-NEGATIVE: nervous, afraid, ...) prevails in both countries	
	Spain: higher E-NEGATIVE during the 2nd semester of 2020	
	Spain: higher presence of extraversion/intraversion-related tweets (compared with USA)	
agreeableness	Negative pole (A-NEGATIVE: cruel, unfriendly, ...) prevails in both countries	
	Spain: sharp increase of A-POSITIVE and A-NEGATIVE in the summer of 2020	
	USA: slight increase of A-NEGATIVE in mid-2020	
conscientiousness	Negative pole (C-NEGATIVE: distracted, unreliable,...) prevails in both countries	
	C-NEGATIVE grew in the 2nd semester of 2020 in both countries	
neuroticism	Negative pole (N-NEGATIVE: balanced, stable,...) prevails in both countries	
	Spain: sharp increase of N-POSITIVE during the 2nd semester of 2020	
	N-POSITIVE goes down in both countries since Dec 2020	
openness to experience	Positive pole (O-POSITIVE: philosophical, abstract...) prevails in both countries	
	Spain: grow in O-NEGATIVE in the 2nd semester of 2020	

Main trends found

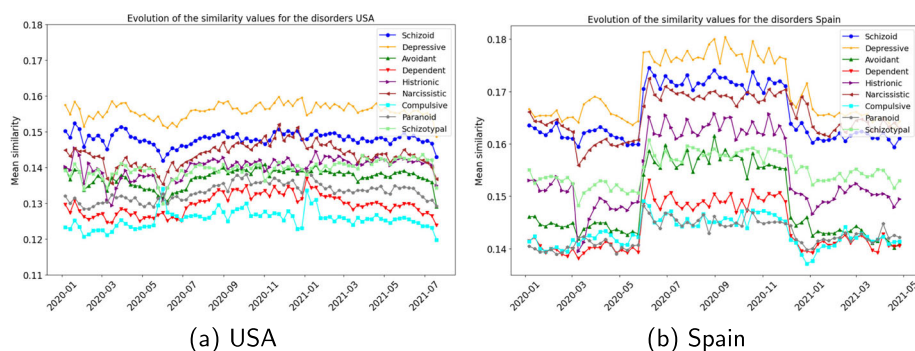(a) USA   (b) Spain

Fig. 12 Estimated Evolution of Personality Disorders for USA and Spain

The second most prevalent personality disorder is schizoid. This personality is a condition in which people avoid social activities and consistently shy away from interaction with others. This type of behavior is to be expected, as the confinement rules and regulations issued during the pandemic led to isolation among people. In fact, the blue lines (schizoid pattern) grew during a core part of the pandemic (second semester of 2020).

The third dominant personality disorder is narcissism, a personality characterized by a high sense of self-importance. In this case, individuals seek a lot of attention and a constant desire for admiration. During the pandemic, this personality may have grown due to the lack of empathy towards other people. The high prevalence of narcissism also connects with the high estimate of negative agreeableness described above (high A-NEGATIVE). A-NEGATIVE is indeed connected to being unempathetic and unhelpful to others.

In mid-2020 we can see a marked increase in most disorders and, for Spain, we can notice a sharp decline at the beginning of 2021. USA patterns also show a decrease in most disorders in late 2020/beginning of 2021, but the decline is smoother. As argued above, the Spain dataset has fewer data points for the central period of 2020 and, thus, the Spain lines are more volatile. However, the plot of the USA and the plot of Spain essentially reveal the same trends, reinforcing the insights extracted from this analysis.

The least dominant personality disorders are dependent and obsessive-compulsive. These are related to a need to have other people take care of them or a desire to have everything organized.

To complement this analysis, in Table 5, we present the tweets with the highest values for the three most prevalent disorders.[7] One of the things we can notice about these tweets is that, usually, people express their discomfort in short texts. For example, Narcissistic disorder is usually associated with tweets referring to someone being selfish or arrogant. Moreover, Appendix A shows a complementary analysis of all traits in a single plot per country that allows visualising the relative prevalence between them.

Table 6 gives an overall view of the main trends derived from the analysis of personality disorders.

---

[7] The Spanish tweets have been translated into English and the table presents the translated tweet.

Table 5 Personality Disorders

Disorder	Tweets
USA	
depression	"Depressed 😔"
	"👆 being depressed"
	"oh dear I'm so hopelessly depressed today"
	"Hopeless world"
schizoid	"apathetic NY Awhile"
	"Me mood"
	"Feeling hateful"
	"yearning"
narcissistic	"Selfish bastard"
	"BE SELFISH"
	"Selfish. Absolutely selfish."
	"Call me selfish cause I got me and nobody else "
Spain	
depression	"Hopeless"
	"What sadness"
	"Depressed but"
	"Chronic sadness."
schizoid	"Apathetic"
	"Drama mood"
	"Feeling Alone."
	"The indifference? It hurt to me"
narcissistic	"Ayyy Selfishness"
	"Selfish, very selfish and bad."
	"An arrogant, arrogant, proud, selfish..."
	"Look who's talking about indifferent selfishness..."
	"Being indifferent nowadays"

Example of tweets

Table 6 Personality disorders analysis

The most prevalent personality disorders during the time of the pandemic were depressive, schizoid, and narcissistic in both countries
Depression is estimated to be the most salient, followed by schizophrenia and narcissism
The least dominant personality disorders are dependent and obsessive-compulsive
Mid-2020: increase in most disorders in both countries
Late 2020/beginning of 2021: decline in most disorders in both countries
Tendencies (up or down) are smoother in USA

Main trends found

4.3 Empath

Next, we continue with the presentation of the topic-based results obtained with Empath. Figure 13 reports, for both countries, the most salient topics extracted by Empath. A first observation is that emotions, either positive or negative, are the most prominent topics. Negative emotions weigh more than positive emotions in both countries.

In the USA sample, we can observe a peak of negative emotions in June 2020 and a higher presence of negative emotions during the second semester of 2020. The Spain plot shows more variance but it essentially reflects the same trend. Comparing the top 10 Empath categories in both countries we can also observe other topics that are also common to USA and Spain: friends, optimism, communication, party, and children.

In the USA plot, most of the lines are rather flat, and perhaps we can only observe a slight decrease in friends and optimism in the last quarter of 2020. The Spain lines are more chaotic and difficult to grasp. Still, there seems to be a higher presence of content related to friends, speaking, and optimism during the second semester of 2020. In any case, the most apparent conclusion from these plots is the high presence of emotions in both countries and the occurrence of some other common topics (friends, optimism, communication, party, and children).

We have further analyzed the peaks of negative emotions (see Table 7, which presents the high-frequency words in the tweets yielding negative emotions). In the USA, we observe a peak in the first week of June 2020. This is mainly related to George Floyd's death, and we observe here the presence of emotional bearing words such as "pissed" or "horrible". A second negative peak, in August 2020, is also partially related to the Black Lives Matter movement but we also find here negative expressions about 2020 and the pandemic ("fucking hate 2020", "hell 2020" or "covid").

In Spain, we observe a sharp rise in negative emotion in the first week of June 2020. According to the most frequent words, the theme of Black Lives Matter is also salient in these tweets. The second increase in negative emotions happened in October 2020 and, taking into account the most recurring words (e.g., "freedom", "cry", "virus", or "government"), these emotions seem to be related to the coronavirus and to the restrictions imposed by the government.

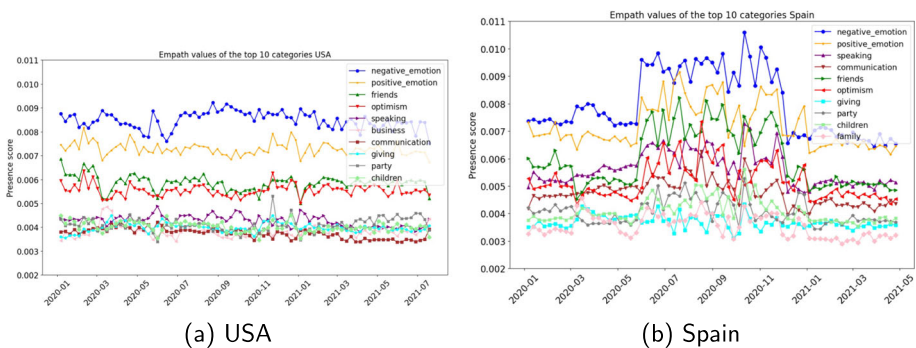Table 8 gives an overall view of the main trends derived from the analysis performed with Empath.

(a) USA   (b) Spain

Fig. 13 Empath topics for USA and Spain

Table 7 High-frequency words during the peaks of negative emotions

Date	List of words
USA	
1st week of June 2020	sad, hate, stupid, crazy, care, sick, dead, lie, death, horrible, cry, mad, scared, worried, insane, pain, pissed, shit, guilty, confused, life, trump, angry, bitch, utah, police, killing, dying, cop, accident, fuck wrong, shocked, makes sick, shot, stop crying, stupid people, super, video, need stop, story, officer, fucking crazy, fucking sick
Last week of August 2020	stop, hate, hard, trump, sad, wrong, hell, back, love, shit, hit, crazy, black, police, life, work, state, reason, today, year, dead, white, fight, stupid, death, killed, break, lie, alone, nevada, mad, worst, night, world, incompetence, regulations, hurt, hope, home, government, family, times, country, america, shot, cops, pain, 2020, covid
Spain	
1st week of June 2020	verdad, vida, mundo, miedo, eeuu, mierda, gobierno, polic, odio, 2020, usa, muerte, culpa, madre, juro, casa, violencia, fascista, triste, dormir, morir, vergüenza, feliz, locura, loco, llorando, democracia, trump, stop, pais, familia, causa, bandera, luchar, muertos, racismo, terrorista
2nd week of October 2020	mal, peor, españa, puto, mala, mundo, odio, joder, llorar, morir, triste, mienten, casa, rutina, virus, partido, poder, gentuza, libertad, locos, sola, sinvergüenza, horrible, grito, gobierno, derecha, clase, payaso, partidos, nacional, malos, madrid, lucha

4.4 Emotion analysis

Figure 14 presents the distribution of emotions. The most dominant emotions are similar in both countries. For example, the emotions of approval and admiration are the top 2 emotions in USA and Spain. Other salient emotions that are common to both countries are amusement, curiosity, love, gratitude, joy, excitement, sadness, and anger. The relative presence of these emotions varies slightly from country to country. For example, the USA sample shows larger percentages of tweets expressing amusement, love, sadness, or excitement, while the Spain sample shows larger percentages of tweets expressing gratitude or realization. In both countries, we can observe that the least frequent emotions are related to remorse, pride, embarrassment, nervousness, relief, and grief.

Table 8 Empath analysis

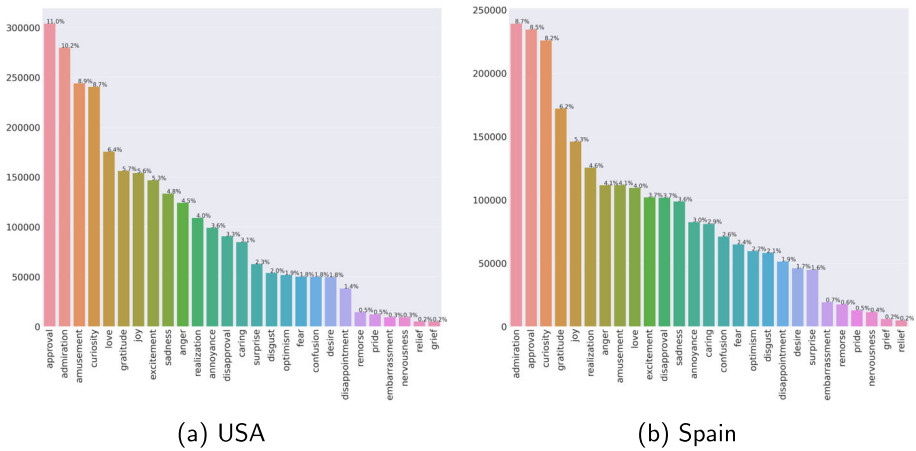| |
| --- |
| Emotions, either positive or negative, are the most prominent topics |
| Negative emotions weigh more than positive emotions in both countries |
| USA, Spain: a peak of negative emotions in June 2020 and a higher presence of negative emotions during the second semester of 2020 |
| Besides emotions, other common topics are: friends, optimism, communication, party and children |
| June 2020 negative peaks: related to George Floyd's death |
| Other negative peaks reflect citizens' concerns about covid and restrictions |
| Main trends found |

Fig. 14 Distribution of emotions, USA and Spain

Let us focus now on the most frequent emotions and their temporal analysis. Figure 15 presents the weekly evolution of the top 10 emotions. The emotions of admiration and approval were dominant throughout the entire period of the pandemic in the USA.

For example, in the USA, we observe two main peaks of anger, one of them in June 2020 and another one in January 2021. This coincides with the previously conducted topic analysis, where we saw an increase of negative feelings associated with George Floyd's death. Another emotion that shows a significant peak is gratitude, which had a large increase in both countries around the Christmas dates. This suggests that people tended to share feelings of appreciation despite the difficulties. In Spain, the emotions of curiosity and sadness dominated during the central part of the pandemic, while the emotions of gratitude and admiration tended to decrease considerably during a large part of 2020.

The emotion of excitement is related to the personality factor of "Extraversion". It is interesting to notice that, for both countries, we can observe a drop in March 2020 and, next, a slight tendency to increase during the remainder of the pandemic. This pattern matches with that of "Extraversion", where we also saw a drop at the beginning of the pandemic followed by a slow recovery.

Table 9 gives an overall view of the main trends derived from the emotion analysis.

Fig. 15 Evolution of emotions, USA and Spain

Table 9 Emotion Analysis

The two most salient emotions are approval and admiration
Other frequent emotions in both countries are amusement, curiosity, love, gratitude, joy, excitement, sadness, and anger
The least frequent emotions are related to remorse, pride, embarrassment, nervousness, relief, and grief
USA: admiration & approval dominant throughout the entire period
Spain: curiosity & sadness high and gratitude & admiration low during the 2nd semester of 2020
USA: two main peaks of anger, in June 2020 and January 2021
USA, Spain: drop in excitement at the beginning of the pandemic followed by a slow recovery
Main trends found

5 Discussion and limitations

Our study has revealed interesting trends in personality factors, disorders, emotions, and the associated topics and events. The analysis of the two poles of the 5 personality dimensions has exposed some consistent findings. For example, in both samples, we could observe the prevalence of the negative poles of extraversion (nervous, afraid, ...), agreeableness (cruel, unfriendly, ...), and conscientiousness (distracted, unreliable, ...). These three poles somehow represent the unfavorable side of these 3 Big-5 factors. On the other hand, the prevalence of the negative pole of neuroticism (balanced, stable, ...) and the positive pole of openness to experience (philosophical, abstract, ...) are better news, as these two poles record favorable aspects of personality. Regarding temporal trends, we have observed some negative changes over the period analyzed. For example, the Spain sample informs of an increase in neuroticism (N-POSITIVE goes up), introversion (E-NEGATIVE goes up), lack of conscientiousness (C-NEGATIVE goes up), and closeness to experience (O-NEGATIVE goes up) in the second semester of 2020. The USA trends also reveal some concerning evolution during 2020, with an increase in neuroticism (N-POSITIVE goes up) and lack of conscientiousness (C-NEGATIVE goes up), and a decrease in agreeableness (A-NEGATIVE goes up).

Our analysis also suggests that the most prevalent personality disorders are depression, schizophrenia, and narcissism (in both countries). The presence of signs of these disorders increased in mid-2020 and had a decline in late 2020/beginning of 2021.

The analysis of Empath topics showed a high presence of emotional publications, with a general prevalence of negative emotions over positive ones, and the occurrence of peaks of negative emotions related to COVID-19 and George Floyd's death.

By processing the data against a fine-grained categorization of 28 types of emotions, we observed that admiration and approval are the most dominant emotions. Still, we also see here some signs of concern. For example, in the USA, there were two peaks of anger (June 2020 and January 2021) and a decrease in excitement at the beginning of the pandemic. In Spain, we also observe the same decrease in excitement and, furthermore, higher signs of sadness and lower signs of gratitude and admiration during the second semester of 2020.

The study presented here has potential limitations. The estimates reported are based on two samples of tweets extracted at specific geographical locations. The two collections are large but, still, are only reflective of the preoccupations of specific segments of the USA or Spain populations. Furthermore, we cannot claim that this gives a picture of the entire population in these areas, because the samples are obviously biased toward people who are

active on a specific social media platform. In any case, we firmly believe that this type of social media analysis provides valuable insights into a significant part of the population. The Spain sample, because of technical difficulties with the extraction tool, has fewer data points during the second part of 2020. Still, we have focused our analysis on the common trends and we are aware that some strong fluctuations observed in the plots are related to the size of some weekly samples.

Another main limitation comes from the estimation methods. In these collections, we do not have a ground truth of personality labels, disorders, or emotions. We rather rely on reputed tools, such as deep linguistic models, emotion detectors, Empath, and so forth. This undoubtedly introduces noisy estimations about the presence of signs of personality, disorders, topics, or emotions. Still, these tools have been tested elsewhere and this type of social media monitoring might err for specific data points but tends to be reliable at identifying relative trends and their temporal evolution.

6 Conclusions

In this paper, we have created and made available to the community two large collections of tweets published during the agitated period of the pandemic. These sets of tweets come from two countries: USA and Spain. Given this large sample of publications, we have performed a comparative analysis of personality factors and disorders, topics, and emotions.

We have been able to observe how different events have influenced the response of the people and we have identified specific time frames when emotions or personality factors suffered substantial changes. For example, negativity arose not only because of COVID-19 but also because of other major events that happened in these two turbulent years.

We can consider social networks as a well-established source of information from which we can extract valuable knowledge and draw a partial but useful view of the feelings of the population. The synergies between health and web/social mining have been explored for a number of years (e.g., for designing new screening technologies). However, we still need to identify the specific role that technology can play in supporting mental health professionals. This type of application is still in its infancy and our paper has addressed some monitoring activities that could be part of future personality/sentiment/emotion analyzers. For example, a given public institution might be interested in tracking certain segments of the population and alert about the onset of some types of risks. In the near future, we plan to contact with relevant stakeholders, such as trained psychologists and other mental health professionals, to validate this technology and to explore future ways of exploitation.

As future work, we are also interested in expanding the study to other data sources and platforms. It would be stimulating to compare multiple information sources and social networks to extract the intensity of emotions, personality aspects, and topics and see how they differ over sources.

Appendix A: Extra personality trait analysis

To complement the personality trait analysis and understand which factor is more prevalent, Figs. 16 and 17 present the evolution of all traits through time. We can see that the negative pole is dominant in both countries and that the traits related to conscientiousness and extraversion have the greatest presence.
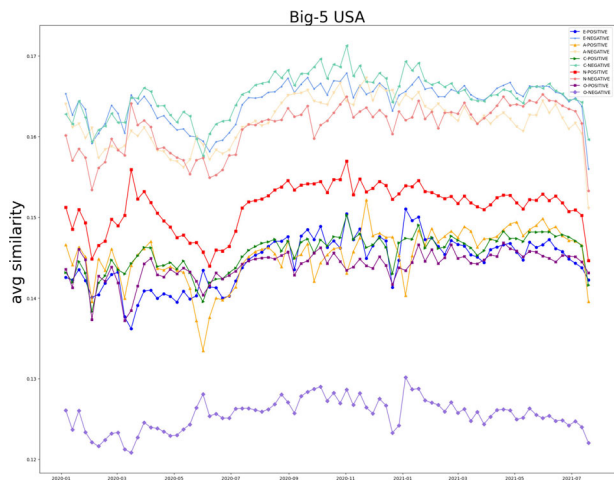
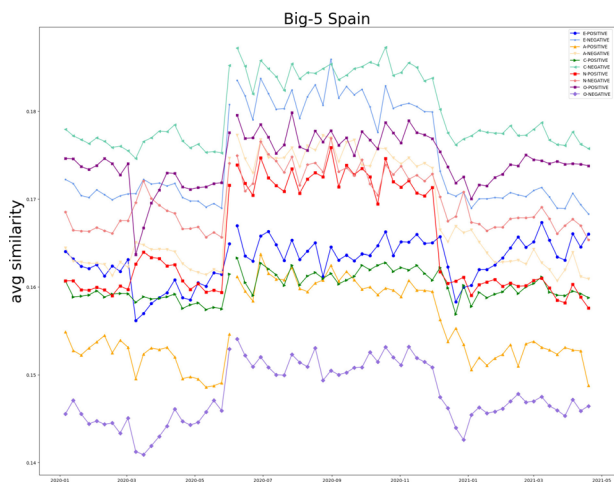**Fig. 16** Big-5 analysis during the COVID-19 pandemic for USA

Fig. 17 Big-5 analysis during the COVID-19 pandemic for Spain

Data and Code availability The datasets and codes generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval Not applicable

Consent to participate Not applicable

Consent for publication Not applicable

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

Ahmad, A. R., & Murad, H. R. (2020). The impact of social media on panic during the covid-19 pandemic in iraqi kurdistan: Online questionnaire study. *J Med Internet Res, 22*(5), e19556. https://doi.org/10.2196/19556

Ainley, E., Witwicki, C., Tallett, A., et al. (2021). Using twitter comments to understand people's experiences of uk health care during the covid-19 pandemic: Thematic and sentiment analysis. *J Med Internet Res*. https://doi.org/10.2196/31101

Alhuzali, H., Zhang, T., & Ananiadou, S. (2022). Emotions and topics expressed on twitter during the covid-19 pandemic in the united kingdom: Comparative geolocation and text mining analysis. *J Med Internet Res, 24*(10). https://doi.org/10.2196/40323

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Borah, A. (2023). Detecting covid-19 vaccine hesitancy in india: a multimodal transformer based approach. *Journal of Intelligent Information Systems, 60*(1), 157–173. https://doi.org/10.1007/s10844-022-00745-1

Bowman, S. R., Angeli, G., Potts, C., et al. (2015). A large annotated corpus for learning natural language inference. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics. Lisbon, Portugal. https://doi.org/10.18653/v1/D15-1075

Canales, L., & Martínez-Barco, P. (2014). Emotion detection from text: A survey. *Processing in the 5th Information Systems Research Working Days (JISIC)*. https://doi.org/10.3115/v1/W14-6905

Di, X., Lifa, W., Zheng, H., et al. (2018). Deep learning-based personality recognition from text posts of online social networks. *Applied Intelligence, 48*,. https://doi.org/10.1007/s10489-018-1212-4

Fast, E., Chen, B., & Bernstein, M. S. (2016). Empath: Understanding topic signals in large-scale text. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery. New York, NY, USA, CHI '16. https://doi.org/10.1145/2858036.2858535

Funder, D.C. (1997). The personality puzzle. W W Norton & Co

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*. https://doi.org/10.1037/0003-066X.48.1.26

Gupta, R. K., Vishwanath, A., & Yang, Y. (2020). COVID-19 twitter dataset with latent topics, sentiments and emotions attributes. CoRR abs/2007.06954. arXiv:2007.06954

John, O. P., & Srivastava, S. (1997). The big five trait taxonomy: History, measurement, and theoretical perspectives. L. A. Pervin & O. P. John (Eds.), Handbook of personality: Theory and research

Kim, T., & Vossen, P. (2021). Emoberta: Speaker-aware emotion recognition in conversation with roberta. https://doi.org/10.48550/ARXIV.2108.12009

Laskar, M. T. R., Huang, X., & Hoque, E. (2020). Contextualized embeddings based transformer encoder for sentence similarity modeling in answer selection task. In: Proceedings of The 12th Language Resources and Evaluation Conference

Leonardi, S., Monti, D., Rizzo, G., et al. (2020) Multilingual transformer-based personality traits estimation. *Information 11*(4). https://doi.org/10.3390/info11040179

Lyu, H., Chen, L., Wang, Y., et al. (2020). Sense and sensibility: Characterizing social media users regarding the use of controversial terms for covid-19. *IEEE Trans Big Data, 7*, 952–960. https://doi.org/10.1109/TBDATA.2020.2996401

Millon, T., Millon, C., & Meagher, S. (2004). *Brief description of the fourteen personality disorders of dsm-iii*. DSM-III-R: Tech. rep.

Neuman, Y., & Cohen, Y. (2014). A vectorial semantics approach to personality assessment. *Scientific reports, 4*(1), 1–6. https://doi.org/10.1038/srep04761

Organization WH (2020). Impact of covid-19 on people's livelihoods, their health, and our food system. https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems. Accessed 06 Jan 2023

Organization WH (2023). WHO coronavirus disease (COVID-19) dashboard. https://covid19.who.int/. Accessed 06 Jan 2023

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics. arXiv:1908.10084

Reuters (2022). Digital news report 2022. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022. Accessed 06 Jan 2023

de-la Rosa, G. R., Jiménez-Salazar, H., Villatoro-Tello, E., et al. (2023). A lexical-availability-based framework from short communications for automatic personality identification. *Cognitive Systems Research, 79*, 126–137. https://doi.org/10.1016/j.cogsys.2023.01.006

Sert, E., Okan, O., Özbilen, A., et al. (2022). Linking covid-19 perception with socioeconomic conditions using twitter data. *IEEE Transactions on Computational Social Systems, 9*(2), 394–405. https://doi.org/10.1109/TCSS.2021.3089657

Si, M. Y., Su, X. Y., Jiang, Y., et al. (2021). Prevalence and predictors of ptsd during the initial stage of covid-19 epidemic among female college students in china. INQUIRY: The Journal of Health Care Organization, Provision, and Financing 58, 00469580211059953. https://doi.org/10.1177/00469580211059953

Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the interpersonal adjective scales to include the big five dimensions of personality. *J Pers Soc Psychol, 59*,. https://doi.org/10.1037/0022-3514.59.4.781

Umair, A., & Masciari, E. (2023). Sentimental and spatial analysis of covid-19 vaccines tweets. *Journal of Intelligent Information Systems, 60*(1), 1–21. https://doi.org/10.1007/s10844-022-00699-4

Vohra, A., & Garg, R. (2023). Deep learning based sentiment analysis of public perception of working from home through tweets. *Journal of Intelligent Information Systems, 60*(1), 255–274. https://doi.org/10.1007/s10844-022-00736-2

Williams, A., Nangia, N., & Bowman, S. (2018). A broad-coverage challenge corpus for sentence understanding through inference. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers).

Association for Computational Linguistics. New Orleans, Louisiana. https://doi.org/10.18653/v1/N18-1101

Yong, A., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*. https://doi.org/10.20982/tqmp.09.2.p079

Zhang, C., Xu, S., Li, Z., et al. (2021). Understanding concerns, sentiments, and disparities among population groups during the covid-19 pandemic via twitter data mining: Large-scale cross-sectional study. *J Med Internet Res, 23*(3), e26482. https://doi.org/10.2196/26482

Zhu, N., Zhang, D., Wang, W., et al. (2020). A novel coronavirus from patients with pneumonia in china, 2019. *New England Journal of Medicine, 382*(8), 727–733. https://doi.org/10.1056/NEJMoa2001017