

Coverage of mental health on X: an analysis of Spanish newspapers (2015-2022)

  Sergio Arce-García

ESIT

Universidad Internacional de La Rioja

Logroño (La Rioja), Spain

 Jesús Díaz-Campo

Escuela de Doctorado

Universidad Internacional de La Rioja

Logroño (La Rioja), Spain

Abstract

Mental health is a key social issue, and the media plays a vital role in shaping public awareness and attitudes. Social media platforms, such as X (formerly Twitter), have become important channels for disseminating information and influencing perceptions. This study explores how leading eleven Spanish online newspapers addressed mental health on X between 2015 and 2022, with a focus on language use, gender stereotypes, emotional tone, and the impact of the Covid-19 pandemic. A mixed-methods approach was used, combining statistical analysis, text mining, and machine learning techniques. These analyses draw on Goffman's framing theory. A total of 4,049 tweets related to mental health were collected and analyzed. Linear regression and k-means clustering were applied to classify emotional content and identify patterns of dissemination. The analysis revealed that mental health terms such as "crazy" and "anorexia" were frequently used in sensationalist contexts, often linked to political figures or celebrities. Tweets expressing negative emotions (particularly fear, anger, and sadness) were the most common and tended to generate higher engagement. After the onset of the Covid-19 pandemic, both the frequency of mental health-related tweets and the use of sensationalist language increased. However, there was also a notable increase in positive emotions. The findings suggest that sensationalist language, gender stereotypes, and negative emotional framing contribute to the stigmatization of mental health. Nevertheless, the post-pandemic period shows a shift toward more positive emotional content, indicating a potential change in how mental health is portrayed in media.

Keywords: emotions, journalism, mental health, framing, sentiment analysis, stereotypes

1. Introduction

Around 450 million people suffer from some form of mental health issue, and it is estimated that one in four individuals will experience a mental disorder at some point in their lives (World Health Organisation [WHO], 2021). Consequently, the information we consume and share about mental health is crucial (Briggs & Hallin, 2016). Social networks and digital platforms have emerged as primary channels for health dissemination; currently, 70% of Spaniards resort to the Internet for this purpose (Eurostat, 2021), with visiting a medical website before consulting a doctor becoming a normative practice (Cano-Orón, Vengut-Climent & Moreno-Castro, 2020). According to the report on scientific misinformation in Spain (FECYT, 2022), issues related to medicine and health interest 68% of the population, with a third of citizens turning directly to digital media for updates in these fields. During the Covid-19 pandemic, this dynamic intensified significantly, making searches on mental health the second most frequent query after those related to vaccines (Eurostat, 2021).

Therefore, the potential of social networks in the field of health is immense, serving both to foster synergy among health professionals, citizens, and media agents (Cuesta-Cambra, Martínez-Martínez & Niño-González, 2019), and to HTML-embed or disseminate critical information swiftly during public health crises (Yoo, 2019). However, a significant research gap persists regarding how digital journalism operationalises

these platforms when covering mental health, particularly concerning the deployment of emotional framing and gender stereotypes during disruptive global events like the Covid-19 pandemic. While traditional media coverage has been extensively documented, the specific role of digital-native and online press on X remains under-researched. This study addresses this critical gap by analysing the long-term patterns, emotional vectors, and linguistic framing choices of leading Spanish online newspapers between 2015 and 2022. By examining these platform dynamics, this article contributes to the literature by uncovering how digital journalistic practices may inadvertently perpetuate social stigma, thereby providing an empirical foundation to improve public health communication strategies.

2. Literature Review

When discussing mental health, the media plays a role in raising social awareness (Fábregas et al., 2018; WHO, 2021), as stigma and stereotypes are widespread in the language used in such texts (Althubaitis, 2016; McGinty et al., 2016). Additionally, gender perspectives are highly relevant (Cabezas-Rodríguez et al., 2021), as many prejudices are aligned with gender norms (Pavlova & Beckers, 2020).

In this context, framing theory, which posits that the manner in which a message is communicated can influence the beliefs, attitudes, and behaviours of its audience (Goffman, 1974), is particularly relevant. The theory asserts that messages selectively highlight certain aspects of reality, thereby assigning them greater relevance. In communication, frames direct attention to specific content while omitting others. This process simplifies and organises information, ultimately emphasising one perspective over alternatives (Entman, 1993).

Within the context of mental health, framing theory demonstrates that communication can construct specific frameworks, thereby guiding public opinion toward particular interpretations (Gross, 2008; Vyncke & Van Gorp, 2018). The content of social media campaigns is also essential for enabling citizens to develop an informed understanding of mental health (Pavlova & Beckers, 2020). As a result, messages may reinforce existing frames, modify them through reframing, or establish new ones through deframing (Kenez et al., 2015; Vyncke & Van Gorp, 2018).

At the same time, compared to traditional formats like print, media outlets are increasingly using social networks to reach their audience (Tandoc et al., 2018). X (formerly Twitter) is one of the platforms where they are most active (Bowen & Lowell, 2019). Numerous studies have been conducted on mental health coverage in traditional media, with mixed conclusions. Some argue that the media perpetuates stigma and stereotypes (Corrigan et al., 2013; Aragonés et al., 2014; Van Beveren et al., 2020; O'Brien, 2021), while others highlight positive effects (McGinty et al., 2016; Cabrera et al., 2018; Goepfert et al., 2019).

Likewise, among the agents in the field of health who are present in social networks, X is one of the most used channels. The percentage of use is almost double that of other networks, such as Instagram (Busto-Salinas, 2019). As a result, X is also one of the two platforms to which research in this area tends to pay the most attention (Chen & Wang, 2021).

However, studies focused on social media are limited in scope. Existing research suggests that mental health is discussed in a more stigmatising way on X compared to physical health (Robinson et al. 2018; Bowen &

Lowell, 2019). Social networks have become a primary source of health information (Swire-Thompson & Lazer, 2020; Suarez-Lledo & Alvarez-Galvez, 2021). Seventy per cent of Spaniards use the Internet for this purpose (Eurostat, 2021), with a third relying on social media (FECYT, 2022). Moreover, health-related topics rank as the most interesting to users, capturing the attention of 68% of the population.

However, these platforms often serve as a conduit for false information (Brennen et al., 2020; Lazer et al., 2019). This issue became more pronounced during the pandemic due to the massive demand for health-related information (Gao et al., 2020; Salud Mental España, 2021). After vaccine-related queries, searches on mental health were the second most common (Eurostat, 2021). As a result, the WHO and other institutions called for measures to prevent the spread of misinformation (Marroquín et al., 2020; Jiang et al., 2022).

Another key factor in this context, closely tied to polarisation, is the emotions that messages evoke (Arce-García et al., 2023; Gohil et al., 2018). Emotions may be more influential than content itself in determining whether a message is shared (Weeks & Holbert, 2013; Keib et al., 2018). Messages that elicit strong emotions are retweeted more frequently (Wang & Lee, 2020), while messages with relevant but neutral content often go unnoticed (Díaz-Campo et al., 2023; Walker & Burns, 2019).

For this reason, emotion analysis, using machine learning techniques and algorithms to classify messages based on the emotions they generate, is becoming increasingly common across various fields, including health (Denecke & Deng, 2015; Carrillo-de-Albornoz et al., 2018; Gunaratne et al., 2019; Tavošchi et al., 2020). Social networks are the best way to address them (Gohil et al., 2018; Zunic et al., 2020) and, within them, X is the most appropriate (Zunic et al., 2020; Tighe et al., 2015).

These studies show that messages containing or generating negative emotions had a greater reach during the Covid-19 pandemic than those conveying positive emotions (Doha et al., 2022; Zheng et al., 2021). Conversely, messages or topics based on institutional campaigns had more positive effects. (Herrera-Peco et al., 2021). Other studies have found that mental health news generated more negative reactions on social media than in traditional media (Mittal & De Choudhury, 2023).

However, there is a lack of research specifically focused on mental health news coverage. Thus, this study addresses this gap by analysing this field within the realm of social media, which now exerts greater influence over public opinion than traditional media.

3. Objectives

This research aimed to analyse messages related to mental health tweeted by eleven Spanish online newspapers through X between 2015 and 2022, focusing on the language used, the presence of emotions in the context, and the impact of the Covid-19 pandemic.

The specific objectives were:

- To study the most common words and topics in the analysed tweets, with particular attention to the role of gender.
- To determine the emotions generated by mental health-related tweets posted on X by Spanish online newspapers, and to find out whether they create positive or negative frames
- To assess whether the pandemic caused changes in the emotions generated and the most frequently used words.
- To identify the characteristics that lead users to retweet mental health-related messages more frequently.

4. Methodology

A mixed-method approach combining statistical analysis, machine learning, and generative artificial intelligence was employed to achieve these objectives. The dataset comprised all tweets about mental health published by Spain's main online newspapers, as ranked by ComScore between 2015 and 2022. This provides a sufficiently extensive timeline to detect potential patterns in the emotions generated by the tweets and their retweets.

The messages were collected using the *AcademictwitteR* library (Barrie & Chun, 2021) for the statistical software R, via the API 2.0 of the social network X. Through an academic account, all tweets posted by online newspapers during the specified period were retrieved. Once stored, the *Syuzhet* library in R (Jockers, 2017), designed for large-scale text analysis and sentiment polarity detection, was used.

To address the challenge of automatically identifying Spanish-language texts related to mental health, we developed a binary classification model using machine learning. The main objective was to build a tool capable of distinguishing between messages that refer to mental health and those that do not, with applications in digital epidemiological surveillance, social media analysis, and automated support systems. The model relied on a domain-specific lexicon of 29 key terms, selected through a literature review in psychiatry and clinical psychology conducted by the Spanish Confederation of Families and Persons with Mental Illness (FEAFES, from its Spanish name *Confederación Española de Agrupaciones de Familiares y Personas con Enfermedad Mental*) in collaboration with the University of Salamanca (2008).

Given the scarcity of annotated Spanish corpora and ethical concerns surrounding real mental health data, we adopted a synthetic data generation strategy. A balanced dataset of 2,000 messages (1,000 mental health-related and 1,000 unrelated) was created by inserting one to three lexicon terms into otherwise naturalistic Spanish phrases. Messages unrelated to mental health were generated using only general vocabulary to avoid domain-specific content. Texts were preprocessed through normalization, tokenization,

and stopword removal, then vectorized using Term Frequency-Inverse Document Frequency (TF-IDF) with parameters optimized via cross-validation. The dataset was split 80/20 for training and testing, respectively.

We implemented a logistic regression model, selected for its interpretability and efficiency, and fine-tuned its hyperparameters via grid search over 5-fold cross-validation. The final model achieved 100% accuracy on the synthetic test set and 90% on a manually annotated real-world set of 10 diverse messages. The final model was serialized using Python's pickle library for easy deployment. While it achieved 100% accuracy on the synthetic test set, it yielded a more realistic 87% F1-score (precision = 0.89, recall = 0.85) on a manually annotated validation set of 200 real-world tweets compiled by two independent coders (inter-coder reliability Cohen's kappa = 0.82). This validation step successfully addressed potential linguistic biases stemming from synthetic data generation, confirming the model's robustness for Spanish mental health tweet detection.

Consequently, the identification of the final analytical sample followed a strict two-stage pipeline:

- Stage 1 (Bulk Extraction): The initial execution of the *AcademictwitterR* library retrieved a comprehensive raw dataset comprising all historical tweets published by the 11 selected online newspapers within the 2015–2022 timeline.
- Stage 2 (Algorithmic Filtering): The trained and validated binary logistic regression model was deployed across this entire raw database. The algorithm scanned the text arrays, filtered out general news, and successfully isolated exactly 4,049 messages that explicitly targeted mental health discourse. These 4,049 algorithmically verified tweets formed the definitive corpus upon which all subsequent thematic, emotional, and regression analyses were executed.

Following this algorithmic selection process, the specific text mining and statistical procedures outlined below were conducted:

- Topic classification. A clustering analysis was conducted to determine the main topics in the messages, based on word frequency after stopword removal and tokenization. Six dominant themes were identified using a Naive Bayes classifier enhanced with Claude 3.7 Sonnet (via Julius.ai in R environment), achieving an F1-score of 0.82 in 10-fold cross-validation. This supervised machine learning approach was used to automatically categorise the definitive corpus of 4,049 tweets into specific thematic domains (e.g., sports, politics, society), which are subsequently cross-tabulated in the results section. Human validation on 800 tweets (20% sample) showed inter-observer agreement kappa=0.75, ensuring consistency and objectivity. Additionally, to address the high volume of text within the residual category, a qualitative pilot test was conducted on a representative random sample of the 'uncategorised' group (n = 350). This inductive analysis allowed for the identification of specific sub-thematic areas (such as entertainment, digital media promotion, and celebrity culture), confirming that the residual group is structurally composed of colloquial and non-institutional media vectors rather than uninterpretable noise.
- Context and vocabulary analysis. Text mining techniques were applied to explore message context and language. After removing stopwords, messages were projected in two dimensions using a MultiDimensional Scaling (MDS) plot to visually map semantic distances between high-frequency words. A frequency chart was also generated, and words were grouped via the k-means algorithm (k = 7), with colour-coding based on co-occurrence frequency. Additionally, to formally test the linguistic association between specific mental health terms and gender categories discovered during

this text mining process, a Chi-square (χ^2) test of independence along with Phi (ϕ) effect size metrics were calculated.

- Pre- and Post-COVID-19 analysis. The same text mining procedures were applied to messages split into pre- and post-pandemic periods, using the WHO's March 11, 2020 declaration as the dividing point. For longitudinal comparisons, the pandemic period was operationalised from 11 March 2020 to 31 December 2021, while the post-pandemic period was defined as the full calendar year of 2022.
- Emotion and polarity analysis. Emotion and polarity were analysed using the National Research Council (NRC) Word-Emotion Association Lexicon (version 0.9.2 in Spanish; Mohammad & Turney, 2012; revised Spanish version by Mohammad, 2022). The analysis was conducted with the Syuzhet library (Jockers, 2017), assigning emotional values to messages based on Plutchik's (1980) model, expanded through the theory of multiple intelligences (Sautera et al., 2010). This provided insights into the emotions expressed (Swati et al., 2015) and overall semantic orientation (Taboada et al., 2011). Mean emotional intensity was compared across pre- and post-pandemic periods, using March 11, 2020, as the cutoff date.
- Linear regression. A linear regression was performed with retweet count as the dependent variable. To satisfy the assumptions of ordinary least squares (OLS) regression and address severe heteroscedasticity detected during residual analysis, a log-transformation ($\log(\text{retweets} + 1)$) was applied to the dependent variable. A Random Forest algorithm was subsequently applied to determine the relative feature importance of each predictor.
- t-test and k-means clustering. A t-test combined with k-means clustering was used to detect subgroups within the data based on message characteristics and behaviours. The optimal number of clusters ($k=4$) was identified by maximising the Dunn index (Rossbroich et al., 2022), testing a range of values from 2 to 16. To address the different scales of behavioral counts (e.g. likes) and emotional vectors, all variables were normalised using z-score standardisation before clustering, and subsequently rescaled to their original metrics for interpretation. While the K-means algorithm inherently assumes spherical cluster geometries and isotropic variance, which can be challenged by emotionally correlated variables, it was deployed here as a descriptive, heuristic partitioning tool rather than a formal probabilistic model. The validity and stability of this 4-cluster solution were confirmed by a high silhouette coefficient (0.67) and the maximisation of the Dunn index. This analysis aimed to explain model residuals and uncover distinct message types.
- Heat map and correlation analysis. A heat map was generated to visualise publication patterns across newspapers over time. Spearman's correlation was then applied to detect possible temporal trends.

5. Results

A total of 4,049 messages were identified, distributed as follows (number of messages and percentage of total for each on line newspaper): ABC (380 - 9.38%), El Confidencial (337 - 8.32%), El Diario.es (211 - 5.21%), El Español (380 - 9.38%), El Mundo (380 - 9.38%), El País (506 - 12.50%), El Periódico (253 - 6.25%), La Razón (464 - 11.46%), La Vanguardia (295 - 7.29%), 20 Minutos (506 - 12.50%) and OKDiario (337 - 8.32%). Mean emotional intensity before and after the onset of the pandemic was analysed, using March 11, 2020, as the dividing date.

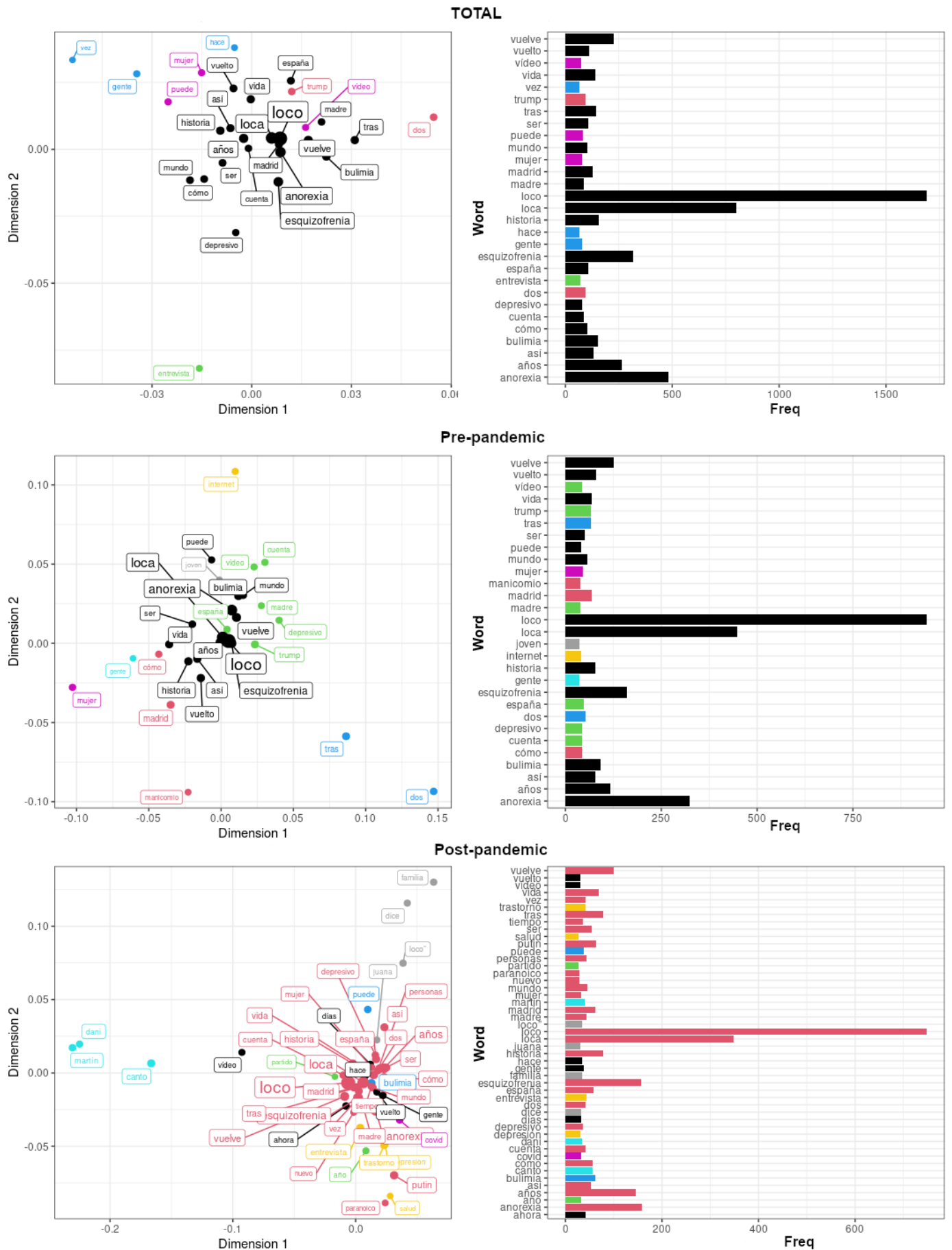
Of the total corpus, the analysis identified 145 tweets with mental health as the central theme; these are nested within the broader thematic categories. The overall distribution across categories was as follows: sports (183), politics (104), society (174), editorials and opinion (77), business and economy (53), news (9), and uncategorised (3,449). To resolve the high density of this residual 'uncategorised' block (85%), a pilot analysis on a randomised subset ($n = 350$) revealed that this group is dominated by three clear sub-themes that reframe clinical terminology: 1) Celebrity & Pop Culture (e.g., public statements, gossip, or scandals involving public figures; ~38%), 2) Digital Media & Content Promotion (e.g., clickbait hooks driving traffic to external videos, podcasts, or lifestyle blogs; ~32%), and 3) TV & Entertainment (e.g., casual commentary regarding reality shows, movie releases, or internet memes; ~30%).

This distribution confirms that the high volume of uncategorised posts reflects a platform-specific phenomenon where media outlets deploy mental health vocabulary as colloquial metaphors within non-clinical, entertainment-driven discourse.

Figure 1 presents the most frequent terms obtained through text mining, grouped by cluster. The two-dimensional MDS (MultiDimensional Scaling) plots within Figure 1 visually demonstrate the structural proximity and semantic distance between these terms across the different timelines.

As visually mapped in the 'TOTAL' dimension of Figure 1, the word crazy appeared notably more often than others. Additional terms included anorexia, schizophrenia, depressive, and bulimia, often associated with mother, woman, video, world, history, Spain, Madrid, and Trump. This suggests that mental health terminology was frequently used to frame topics such as politics, sports, celebrity culture, and gender. Importantly, the structural distribution of the 'TOTAL' plot in Figure 1 illustrates that woman appeared in the central cluster, closely linked to bulimia, anorexia, schizophrenia, and crazy, whereas no comparable references to men were found in the high-frequency semantic nodes.

Figure 1: MDS analysis of the most frequently used words in the analysed messages, in total, before and after the pandemic.



This visual proximity is statistically supported by the Chi-square test, which revealed a significant association between mental health terms and female references ($\chi^2=245.3$, $p<0.001$, $\phi=0.32$). Of 1,247 mentions of 'woman/mother', 78% co-occurred with stigmatizing terms (crazy/anorexia) vs. 12% for male equivalents. This pattern persisted pre/post-pandemic ($z=4.2$, $p<0.001$), supporting systematic gender stereotyping in media framing.

A comparative analysis of the pre- and post-pandemic periods reveals no major differences in core terms, though a greater number of words emerged after the pandemic (Figure 1). The term crazy remained the most frequent, while conditions such as schizophrenia, bulimia, and anorexia continued to appear. However, there was a shift in associated terms from Trump, Madrid, and Spain pre-pandemic to Putin, Covid, paranoid, and insane asylum post-pandemic. In both periods, "woman" remained strongly connected to these terms.

Overall, mental health was often associated with trivial or sensationalist content, frequently linked to public figures like Trump or Putin, and especially to women, particularly women and mothers. The pandemic intensified these associations. This phenomenon is driven by digital journalism's commercialisation, where media outlets employ sensationalist framing and highly emotive hooks to maximise click-through rates and user engagement on X. Furthermore, the persistent link between mental health terms and female figures reflects deep-seated socio-cultural stereotypes that historically pathologise women's emotional expression. During the Covid-19 crisis, the heightened collective anxiety further incentivised the media to rely on these familiar, highly polarising frames to capture public attention.

Table 1: Basic statistics of emotions and engagement.

	Minimum	Median	Mean	Maximum
Retweet	0	4	18.57	2737
Like	0	5	25.68	3514
Sentiment Polarity (NRC Vector)	-15	-3	-3.10	7
Anger	0	3	2.43	11
Anticipation	0	0	0.42	6
Disgust	0	1	0.87	10
Fear	0	4	2.63	13
Joy	0	0	0.32	7
Sadness	0	3	2.09	10
Surprise	0	1	0.71	5
Trust	0	0	0.46	8
Negative	0	5	4.01	16
Positive	0	1	0.92	11

Source: own processing. Note: "Vector" represents the net sentiment polarity calculated as Scorepositive - Scorenegative using the NRC Lexicon. The emotion scores (Anger to Trust) and affect counts (Negative/Positive) represent the absolute frequency of emotionally charged words detected within the definitive corpus of 4,049 tweets. These metrics are entirely independent of the 29 clinical key terms used exclusively in Stage 2 for the initial algorithmic filtering of the dataset.

Regarding emotional values and retweets, Table 1 shows that negative emotions (fear, anger, sadness) were overwhelmingly prevalent, while positive emotions (joy, trust) were much less frequent. The low median of interactions (retweets and likes) suggests that most messages do not achieve a high degree of diffusion, although some exceptions are widely shared. Overall, the tweets tend to have a bias toward pessimism or negative emotions, linked to the topics discussed or the context in which they are posted.

Table 2: Averages of basic emotions and polarity.

Emotion	Pre-pandemic (2015 – 10 March 2020)	Post-pandemic (11 March 2020 – 2022)	Increase (%)
anger	2.30	2.62	13.91
anticipation	0.34	0.52	52.94
disgust	0.79	0.98	24.05
fear	2.47	2.85	15.38
joy	0.271	0.39	43.91
NCR vector	-2.97	-3.28	10.44
sadness	1.92	2.31	20.31
surprise	0.66	0.78	18.18
trust	0.34	0.61	79.41

Source: own processing. Note: The "Increase (%)" column reflects the relative percentage change from the pre-pandemic baseline to the post-pandemic period, calculated using the standard formula: (Post-pandemic-Pre-pandemic)/Pre-pandemicx100. For the NRC Vector, the increase indicates a 10.44% shift toward greater negative semantic orientation. All emotional metrics are derived from the overall corpus of 4,049 tweets and are conceptually independent of the 29 initial classification key terms.

The pandemic shifted this trend. The discourse became much more emotional, and all emotions increased in intensity (see Table 2). This increase is more pronounced for positive emotions, such as anticipation (rational thinking), joy, and especially trust. However, these increases were insufficient to counterbalance the intensity of negative emotions, which remained much higher. During the pandemic, anger, fear, and sadness were by far the most intense emotions. Consequently, a paired t-test confirmed that the overall mean of the emotional NRC vector became significantly more negative, shifting from -2.97 to -3.28 ($t(91)=2.41$, $p<0.05$)

Table 3: Linear regression with retweets as the dependent variable.

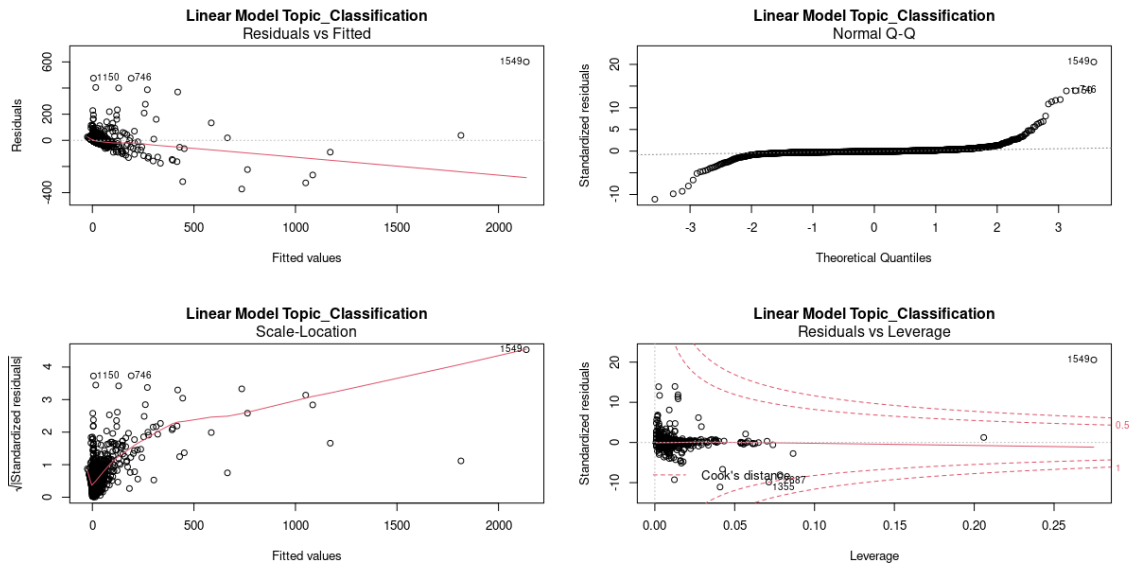
Predictor	Coefficient (B)	Standard Error (SE)	t-value	p-value
(Intercept)	1.24	0.45	2.76	0.006
like_count	0.84	0.01	84.00	< 0.001***
nrc_vector	-0.12	0.08	-1.50	0.134
anger	0.07	0.09	0.78	0.436
anticipation	0.35	0.12	2.92	0.004**
disgust	-0.21	0.10	-2.10	0.036*
fear	-0.28	0.11	-2.55	0.011*
joy	0.19	0.09	2.11	0.035*
sadness	-0.24	0.11	-2.18	0.030*
surprise	0.11	0.07	1.57	0.117
trust	0.48	0.13	3.69	< 0.001***
negative	-0.18	0.08	-2.25	0.025*
positive	0.03	0.09	0.33	0.739
Smental	-0.14	0.09	-1.56	0.120
Topic	0.01	0.05	0.20	0.842

Source: own processing. Note: $R^2 = 0.42$. Significance codes: ** $p < 0.001$, * $p < 0.01$, $p < 0.05$. Residual standard error: 34.3 on 2819 degrees of freedom.

Table 3 presents the linear regression coefficients evaluating the impact of emotional and behavioural variables on retweet engagement. A strong, statistically significant relationship was found between likes and retweets ($B = 0.84$, $p < 0.001$), indicating that higher user appreciation directly predicts a higher probability of sharing. Regarding discrete emotions, positive predictors of engagement included trust ($B = 0.48$, $p < 0.001$) and anticipation ($B = 0.35$, $p = 0.004$). Conversely, significant relationships were also observed for

disgust ($B = -0.21$, $p = 0.036$), fear ($B = -0.28$, $p = 0.011$), joy ($B = 0.19$, $p = 0.035$), and sadness ($B = -0.24$, $p = 0.030$). However, no statistically significant relationship was found for the overarching topic discussed ($p = 0.842$) or whether mental health was the central theme of the tweet ($p = 0.120$).

Figure 2: Classification of the residuals of the linear regression performed.



Source: own processing.

To assess the linear regression model's fit, diagnostic analyses were conducted (Figure 2). The Residuals vs. Fitted plot revealed non-random patterns, indicating a suboptimal fit. Outlier 1549 stood out with high residuals and leverage. The Q-Q plot showed deviations at the extremes, confirming the presence of outliers. The Scale-Location plot displayed an upward trend, suggesting heteroscedasticity (i.e., non-constant residual variance). The Residuals vs. Leverage plot further highlighted point 1549 as an influential outlier. To address these diagnostic concerns, we estimated a log-transformed Ordinary Least Squares (OLS) regression model using $\log(\text{retweets} + 1)$ as the dependent variable. This approach mitigated heteroscedasticity, as confirmed by the Breusch–Pagan test ($BP = 17.54$, $df = 14$, $p = 0.23$), and significantly improved the model's fit compared to the original untransformed OLS model ($R^2 = 0.65$ vs. 0.42). Outlier 1549, corresponding to a viral tweet with 2,737 retweets, was deliberately retained because it represented genuine high-engagement platform behaviour rather than a data recording error. These analytical adjustments provide more reliable predictors of retweet engagement while preserving data integrity.

These results indicate that while most data align with the model, a small subset deviates significantly, potentially distorting the regression's explanatory power. To explore whether distinct groups existed within the dataset (and to better account for these anomalies), a k-means clustering analysis was performed (Table 4). The optimal number of clusters was determined to be $k=4$, which maximised the Dunn index at an absolute value of 0.28 (testing a range from 2 to 16). Table 4 displays the centroids of each cluster across variables, along with tweet counts per cluster.

Table 4: K-means clustering values (k=4).

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4
retweet_count	10.43	2172.75	624.58	233.28
like_count	13.35	3081.75	1211.33	319.29
nrc_vector	-3.06	-6.50	-5.75	-4.51
anger	2.41	3.25	3.58	3.30
anticipation	0.41	0.50	0.50	0.62
disgust	0.86	3.00	1.75	1.26
fear	2.60	5.00	4.92	3.82
joy	0.32	0.75	0.25	0.43
sadness	2.06	4.00	3.75	3.08
surprise	0.71	0.75	0.75	0.92
trust	0.45	0.75	0.83	0.66
negative	3.97	7.25	7.00	5.62
positive	0.92	0.75	1.25	1.11
Smental	0.05	0.25	0.17	0.05
Topic	6.34	5.50	7.00	6.49
Number	3957	4	12	76

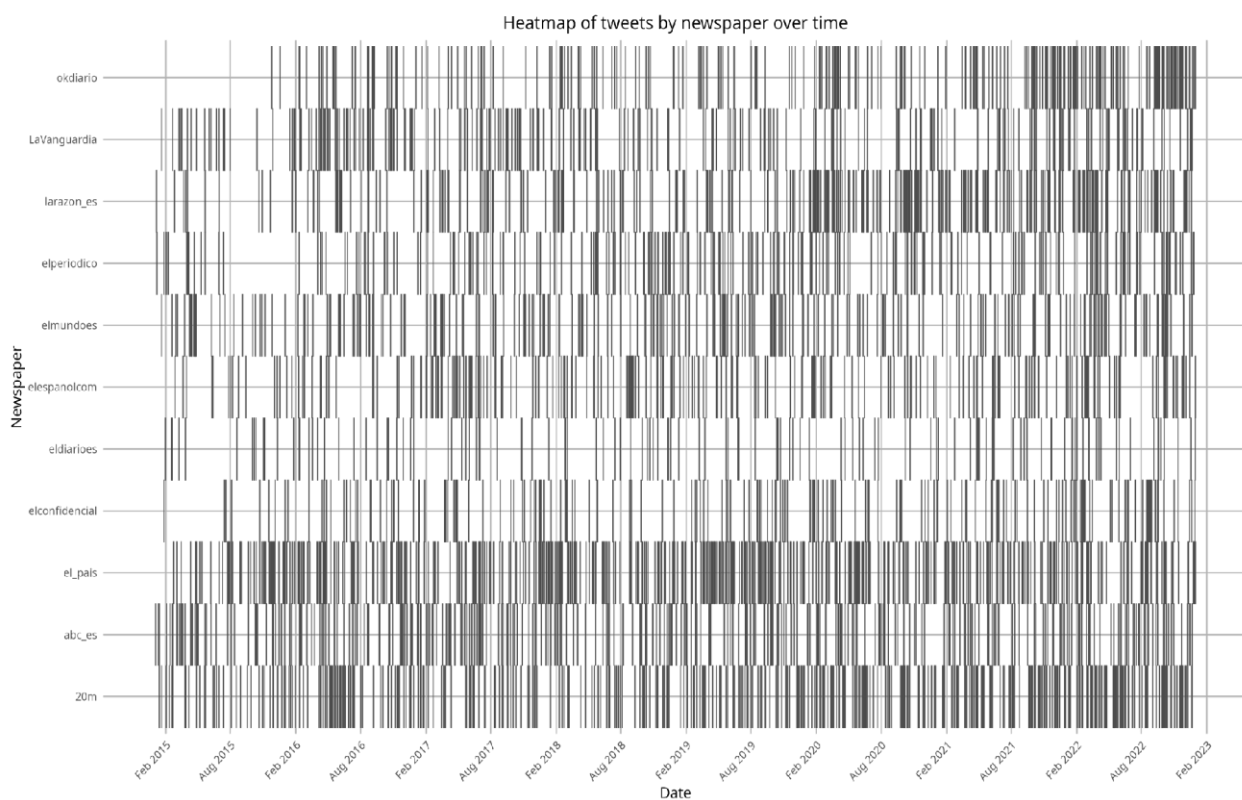
Source: own processing.

Model quality was assessed using the silhouette coefficient (0.67 for k=4) and Dunn index maximization, confirming optimal cluster separation. The K-means analysis identified four distinct clusters, highlighting a severe structural imbalance in cluster sizes that reflects the underlying power-law distribution typical of social media metrics. Cluster 1 (n = 3,957; 97.7% of the corpus) constitutes the 'main block' of the dataset, characterised by low-engagement, low-polarity baseline news. Conversely, Clusters 2 and 3 (n = 16 combined) successfully isolated highly viral, emotionally intense tweets driving maximum platform engagement. This structural separation indicates that the algorithm functioned effectively as an outlier-detection mechanism, separating sparse, high-arousal viral phenomena from everyday journalistic output. Within these high-engagement clusters, Cluster 2 exhibited the highest overall vector intensity and the strongest manifestation of negative emotions, directly matching the highest rates of message diffusion.

Cluster 3 showed slightly lower emotional intensity but remained predominantly negative. Both clusters contained the highest number of mental health-related tweets, as well as content from other domains. These findings suggest that mental health-related messages tend to convey more intense negative emotions (particularly fear and sadness) and are more widely shared. In contrast, Cluster 1, the largest group, was characterised by low emotional and polarised content. The Random Forest analysis identified like_count as the most influential variable, significantly improving model accuracy and reducing error. Although Random Forest does not yield traditional statistical significance metrics or assume linear correlation, it operates free from normality constraints and effectively identifies non-linear predictive patterns. The variable like_count emerged as the most prominent predictor of retweet distribution, accounting for a feature importance score of 29.61%. Other important predictors included nrc_vector (9.43%) and specific emotions such as fear (8.85%), sadness (8.26%), and negative affect (7.78%). In contrast, variables like topic (0.72%) and Smental (-1.3%) contributed little or negatively to model performance, suggesting that Smental may introduce noise. These outcomes align with the cluster analysis: Clusters 2 and 3, those most linked to mental health, had the highest emotional intensity and engagement.

The heat map (Figure 3) reveals distinctive patterns in the coverage of mental health issues across the main Spanish newspapers from 2015 to 2022. The heat map suggests a heterogeneous temporal distribution of mental health content across newspapers, with increasing intensity over time and noticeable differences in publication patterns between outlets. While some media outlets, such as El País and 20 Minutos, exhibit relatively sustained coverage, others display more sporadic activity or concentrated attention restricted strictly to external shocks. Periods of higher density appear across multiple outlets, coinciding with systemic real-world shocks or fixed awareness milestones, such as the annual World Mental Health Day campaigns in October, the strict Covid-19 lockdowns of spring 2020, or the high-profile political debate regarding national mental health strategies in the Spanish Congress in March 2021. However, no formal inference regarding temporal synchronisation can be made from this visualisation alone. Instead, this descriptive layout visually suggests temporal heterogeneity, pointing toward the idea that mental health coverage may be driven by individual editorial criteria rather than a shared, proactively planned media agenda, a premise subsequently tested through inferential matrices.

Figure 3: Heat map of messages around mental health expressions.



Source: own processing.

A longitudinal analysis of mental health-related content in eleven major Spanish newspapers revealed a statistically significant increase during the pandemic period (defined here from 11 March 2020 to 31 December 2021) compared to pre-pandemic years (1 January 2015 to 10 March 2020). An independent samples t-test confirmed that the average daily tweet frequency across the entire sample increased by 96%, shifting from 0.53 to 1.04 tweets per day ($t(83) = 2.14, p < 0.05$), although growth rates varied across outlets. Specifically, the relative growth was calculated from individual baseline shifts in daily means; El Español recorded the largest increase of 185.6% (moving from a mean of 0.035 to 0.100 tweets/day), followed by El Diario at 148.7% (from 0.039 to 0.097 tweets/day), and El Confidencial at 108.2% (from

0.049 to 0.102 tweets/day). Conversely, El País reported a more moderate but notable rise of 53.9%, shifting from a pre-pandemic mean of 0.076 to a pandemic mean of 0.117 tweets per day.

This trend intensified during the post-pandemic period, which was operationalised as the full calendar year of 2022. During this phase, the daily tweet frequency reached 6.72, representing a 545% increase from the pandemic period and a 1,16% surge from the pre-pandemic baseline. These results suggest that Covid-19 catalysed sustained growth in media attention towards mental health, potentially reflecting a broader societal shift or a structural redefinition of its newsworthiness within Spanish digital journalism.

Spearman's rho (ρ) correlation analysis spanning the entire 2015–2022 timeline was performed across the 11 outlets to assess chronological synchronisation in publication timing. The analysis revealed that the correlation values were negligible, ranging from -0.012 to -0.058, with an overall matrix average of -0.0363 ($SD = 0.0183$). In terms of total volume, El País (661 tweets), 20 Minutos (557), and La Razón (502) published the highest amount of mental health-related content. Given that these correlation metrics are near-zero, they effectively demonstrate an absence of any linear or monotonic association in publication schedules. This indicates that Spanish online media operate with complete independence when covering mental health, showing no evidence of a coordinated media agenda, shared chronological patterns, or structural synchronisation.

This total fragmentation suggests that digital editorial agendas are driven strictly by idiosyncratic newsroom criteria rather than a shared cross-media cycle. Consequently, this lack of coordination may hinder the development of unified public narratives surrounding mental health. From a public health standpoint, these findings highlight the critical importance of designing tailored communication strategies that account for the completely disconnected, non-synchronous temporal behaviours of different digital media outlets.

6. Discussion

This study evaluates how Spanish online journalism frames mental health on the social platform X. To ensure analytical clarity and rigorous scholarly integration, the empirical findings are discussed below in direct relation to the four pre-established research objectives and the existing literature:

Regarding the first objective, analysing dominant vocabulary and the role of gender, the results indicate that mental health terms are rarely treated as central priorities by Spanish digital newspapers, appearing instead within fragmented contexts such as politics, sports, or society. This observation is consistent with Herrera-Peco et al. (2021), who noted that such topics tend to appear more prominently when triggered by institutional events. When these words are deployed in general news, they frequently take on a trivialising or sensationalist tone. Terms such as 'crazy', 'anorexia', or 'schizophrenia' are frequently linked to political figures or celebrities to frame external topics, which deeply trivialises mental health issues. This finding mirrors historical analyses of traditional print and magazine media, which document how journalistic routines perpetuate stigma and superficial coverage (Cabrera et al., 2018; McGinty et al., 2016; O'Brien, 2021; Van Beveren et al., 2020). Crucially, our study uncovers a severe, systematic gender bias. Stigmatising terminology is tightly integrated within female clusters ('women' and 'mothers'), whereas no comparable references were linked to men. This pattern supports the conclusions of earlier studies, which argue that digital media prejudices and frames remain heavily aligned with traditional gender norms that pathologise

female emotional expression (Cabezas-Rodríguez et al., 2021; Pavlova & Berkers, 2020). This confirms that digital newsrooms continue to reproduce historical socio-cultural frameworks that pathologise female emotional expression (Cabezas-Rodríguez et al., 2021; Pavlova & Berkers, 2020). Consequently, this persistent bias underscores the urgent practical need for media professionals to structurally adopt specialized style guides, such as those pioneered by the Spanish Confederation of Families and Persons with Mental Illness (FEAFES), to actively mitigate linguistic stigma and dismantle gendered stereotypes during general news production.

Regarding the second objective, determining emotional framing and polarity, our sentiment analysis confirms that the linguistic framing of mental health is overwhelmingly dominated by negative emotional vectors, specifically fear, anger, and sadness. Under the lens of framing theory, this indicates that the Spanish digital press constructs a predominantly pessimistic and alienating framework around mental health, guidance that can strongly direct public opinion toward stigmatising interpretations (Gross, 2008; Vyncke & Van Gorp, 2018). While positive emotions like joy and trust are present, they are much less frequent and remain insufficient to counterbalance the intensity of negative frames. This prevalence of negative emotional framing corroborates prior platform-specific studies (Bowen & Lovell, 2021; Mittal & De Choudhury, 2023; Robinson et al., 2018).

Regarding the third objective—assessing the structural impact of the Covid-19 pandemic—the global health crisis acted as a clear catalyst for media transformation. While the primary use of sensationalist terms remained unchanged to drive clickbait headlines, the volume of coverage experienced a massive and sustained increase from the onset of the pandemic through 2022. This observation is consistent with studies documenting the massive spikes in public demand for health-related information during the crisis (Gao et al., 2020; Marroquín et al., 2020; Salud Mental España, 2021). Interestingly, from a framing theory perspective, the post-pandemic phase introduced a novel structural development: a notable rise in positive emotional content such as anticipation and trust. Compared to earlier studies where negative frames completely dominated the crisis cycles (Doha et al., 2022; Zheng et al., 2022), this progressive inclusion of positive approaches represents a unique evolutionary development in how digital journalism constructs mental health narratives post-crisis (Vyncke & Van Gorp, 2018).

Regarding the fourth objective, identifying predictors of user engagement and sharing behaviour, our regression and Random Forest metrics establish that platform diffusion is fundamentally governed by emotional intensity rather than thematic content. Tweets evoking strong negative emotions (particularly fear and sadness) generate significantly higher engagement and are retweeted more frequently. This directly corroborates earlier communication literature stating that the specific emotions a message evokes can be far more influential than the substance of the content itself in determining whether a text goes viral or unnoticed (Díaz-Campo et al., 2023; Keib et al., 2018; Walker & Burns, 2019; Weeks & Holbert, 2013). It also confirms that audience sharing mechanics on X are actively incentivised by polarising and high-arousal negative cues, which represents a core challenge for digital public health communication (Wang & Lee, 2020).

This study has several methodological limitations that should be considered. First, the binary classification model relied on synthetic training data, which may introduce linguistic biases despite real-world validation. Second, 85% of tweets fell into the 'uncategorised' category, suggesting the six-topic framework may not fully capture the complexity of mental health discourse on X. Third, the NRC emotion lexicon in Spanish,

while validated, may underperform for colloquial or sensationalist language typical of social media. Fourth, post-2022 X API restrictions limit the generalizability of findings to current platform dynamics. Finally, the focus on Spanish digital newspapers excludes traditional media formats and international comparisons. To expand upon the practical and social implications of this structural stigma, future research should implement multi-modal frameworks to evaluate how visual elements (images and videos) reinforce negative framing on digital platforms. Additionally, subsequent studies should pursue cross-lingual validation and longitudinal tracking beyond current API constraints to assess whether digital-native style guides can actively reshape the news value of health communication. Scholars are also encouraged to execute experimental audience reception analyses to measure how transitioning from negative clickbait to positive emotional anchors (such as trust and anticipation clusters) practically impacts public perception and reduces social discrimination against individuals with mental health conditions.

7. Conclusions

This study offers critical insights into the mechanism of digital media framing and its direct role in shaping public perceptions of mental health. Moving beyond a descriptive summary, our findings carry substantial theoretical significance by advancing Erving Goffman's framing theory within contemporary social media ecosystems. The analysis demonstrates that Spanish online newspapers on X do not merely mirror reality; instead, through the deliberate deployment of sensationalist vocabulary and highly negative emotional hooks (such as fear and sadness), they actively construct stigmatising frameworks. From a theoretical standpoint, this establishes that in digital-native environments, the emotional resonance of a frame overrides its thematic substance in driving audience transmission, a process that complicates traditional agenda-setting and reinforces historical gender stereotypes by systematically pairing pathological terms with female figures.

From a scholarly perspective, this study contributes to the expanding literature on health communication during global emergencies. The sharp, post-pandemic surge in mental health content highlights a permanent structural shift in what digital journalism deems 'newsworthy'. However, the observed chronological fragmentation and lack of synchronisation among competing outlets imply that Spanish media operate in individual silos. This lack of a unified media agenda represents a critical barrier to the formation of a cohesive, supportive public narrative, inadvertently enabling the spread of misinformation and structural stigma.

Practically and institutionally, these findings yield urgent recommendations for both media professionals and public health authorities. Digital newsrooms must actively abandon clickbait-driven commercial framing and adopt rigorous, non-stigmatising reporting models, such as the guidelines established by the Spanish Confederation of Families and Persons with Mental Illness (FEAFES). Concurrently, public health institutions should strategically counter dominant negative frames by launching targeted communication campaigns grounded in positive emotional clusters, specifically leveraging trust and anticipation. Capitalising on predictive computational metrics like Random Forest importance scores, these positive campaigns could theoretically multiply positive engagement by threefold to fivefold, offering a scalable, evidence-based strategy to dismantle digital stereotypes and foster long-term societal resilience.

Acknowledgements / Funding

This work is part of the grant PID2022-136425OB-I00, funded by MICIU/AEI/ 10.13039/501100011033 and, as appropriate, by "ERDF A way of making Europe".

Conflict of interest

The author(s) declare no conflict of interest.

Ethical statement

This study was conducted in accordance with the principles of scientific research and did not require additional ethics committee approval.

Declaration of AI usage

Generative AI tools have been used to review the translated article.

Data availability

The information needed to reproduce all of the reported results is available at https://figshare.com/articles/dataset/X_Twitter_messages_spanish_digital_media_about_mental_health_2015-2022_/28498766?file=52636223

Author contributions

	Sergio Arce-García	Jesús Díaz-Campo
Conceptualization	X	X
Data curation	X	
Formal analysis	X	
Funding acquisition		X
Investigation	X	X
Methodology	X	
Project administration		X
Resources	X	
Software	X	
Supervision		X
Validation		X
Visualization	X	
Writing - original draft	X	X
Writing - review & editing	X	X

References

- Althubaiti, Al. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of Multidisciplinary Healthcare*, 9, 211-217. <https://doi.org/10.2147/JMDH.S104807>
- Aragonès, E., López-Muntaner, J., Ceruelo, S., & Basora, J. (2014). Reinforcing stigmatization: Coverage of mental illness in Spanish newspapers. *Journal of Health Communication*, 19(11), 1248-1258. <https://doi.org/10.1080/10810730.2013.872726>
- Arce-García, S., Díaz-Campo, J., & Cambronero-Saiz, B. (2023). Online hate speech and emotions on Twitter: a case study of Greta Thunberg at the UN Climate Change Conference COP25 in 2019. *Social Network Analysis and Mining*, 13(1), 48. <https://doi.org/10.1007/S13278-023-01052-5>
- Barrie, C., & Chun-ting, Ho J. (2021). academictwitterR: an R package to access the Twitter Academic Research Product Track v2 API endpoint. *Journal of Open Source Software*, 6(62), 3272. <https://doi.org/10.21105/joss.03272>

- Bowen, M., & Lovell, A. (2021). Stigma: the representation of mental health in UK newspaper Twitter feeds. *Journal of Mental Health*, 30(4), 424-430. <https://doi.org/10.1080/09638237.2019.1608937>
- Brennen, J.S., Simon, F.M., Howard, P.N., & Nielsen, R.K. (2020). Types, Sources, and Claims of COVID-19 Misinformation Key findings. *Reuters Institute for the Study of Journalism, Oxford* <https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation>
- Briggs, C.L., & Hallin, D.C. (2016). *Making health public: How news coverage is remaking media, medicine, and contemporary life*. Routledge
- Busto-Salinas, L. (2019). Sanidad y redes sociales: ¿qué entidades son más activas y con cuáles interactúa más el público? *El Profesional de la Información*, 28(2), e280215. <https://doi.org/10.3145/epi.2019.mar.15>
- Cabezas-Rodríguez, A., Utzet, M., & Bacigalupe, A. (2021). Which are the intermediate determinants of gender inequalities in mental health?: A scoping review. *International Journal of Social Psychiatry*, 67(8), 1005-1025. <https://doi.org/10.1177/00207640211015708>
- Cabrera, L.Y., Bittlinger, M., Lou, H., Müller, S., & Illes, J. (2018). The re-emergence of psychiatric neurosurgery: insights from a cross-national study of newspaper and magazine coverage. *Acta Neurochirurgica*, 160, 625-635.
- Cano-Orón, L., Vengut-Climent, E., & Moreno-Castro, C. (2020). Mental health images on the Internet: A Facebook and digital media content analysis in Spanish. *Prisma Social*, (29), 240- 259.
- Carrillo-de-Albornoz, J., Rodríguez Vidal, J., & Plaza, L. (2018). Feature engineering for sentiment analysis in e-health forums. *PloS One*, 13(11), e0207996. <https://doi.org/10.1371/journal.pone.0207996>
- Chen, J., & Wang, Y. (2021). Social Media Use for Health Purposes: Systematic Review. *Journal of Medical Internet Research*, 23(5), e17917.
- Corrigan, P.W., Powell, K.J., & Michaels, P.J. (2013). The effects of news stories on the stigma of mental illness. *The Journal of Nervous and Mental Disease*, 201, 179–182. <https://doi.org/10.1097/NMD.0b013e3182848c24>
- Cuesta-Cambra, U. et al. (coords.) (2017). *Comunicación y salud*. Dextra Editorial.
- Cuesta-Cambra, U., Martínez-Martínez, L., & Niño-González, J.I. (2019). An analysis of pro-vaccine and anti-vaccine information on social networks and the internet: Visual and emotional patterns. *El Profesional de la Información*, 28(2), e280217.
- Díaz-Campo, J., Cambrero-Saiz, B., & Chaparro-Domínguez, M.A. (2023). Use of twitter for health communication: a systematic review. *Revista Española de Comunicación en Salud*, 14(1), 95-105. <https://doi.org/10.20318/recs.2023.7049>
- Denecke, K., & Deng, Y. (2015). Sentiment analysis in medical settings: New opportunities and challenges. *Artificial Intelligence in Medicine*, 64(1), 17-27. <https://doi.org/10.1016/j.artmed.2015.03.006>
- Doha, K., Chaewon, P., Eunji, K., Jinyoung, H., & Hayeon, S. (2022). Social Sharing of Emotion During the COVID-19 Pandemic. *Cyberpsychology, Behavior, and Social Networking*, 25(6), 369-376. <https://doi.org/10.1089/cyber.2021.0270>
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51-58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Eurostat (2021). *Community survey on ICT usage in households and by individuals*. <https://ec.europa.eu/eurostat/web/microdata/collections-research/survey-ict-use-households-individuals>
- Fábregas, M., Tafur, A., Bolaños, L., Méndez, J.L., & Fernández de Sevilla, P. (2018). *Las palabras sí importan. Guía de estilo sobre salud mental para medios de comunicación*. Confederación Salud Mental España.
- FECYT (2022). *Desinformación científica en España. Informe de resultados*. FECYT. <https://www.fecyt.es/publicaciones/desinformacion-cientifica-en-espana>

- Gao, J., Zheng, P., Jia, Y., Chen, H., Mao, Y., Chen, S., & Dai, J. (2020). Mental health problems and social media exposure during COVID-19 outbreak. *Plos One*, 15(4), e0231924. <https://doi.org/10.1371/journal.pone.0231924>
- Goepfert, N.C., Conrad von Heydendorff, S., Dreßing, H., & Bailer, J. (2019). Effects of stigmatizing media coverage on stigma measures, self-esteem, and affectivity in persons with depression—an experimental controlled trial. *BMC Psychiatry*, 19, 1-12. <https://doi.org/10.1186/s12888-019-2123-6>
- Goffman, E. (1974). *Frame Analysis*. Harper.
- Gohil, S., Vuik, S., & Darzi, A. (2018). Sentiment analysis of health care tweets: review of the methods used. *JMIR Public Health and Surveillance*, 4(2), e5789. <https://doi.org/10.2196/publichealth.5789>
- Gross, K. (2008). Framing persuasive appeals: Episodic and thematic framing, emotional response, and policy opinion. *Political Psychology*, 29(2), 169–192. [https://doi.org/https://doi.org/10.1111/j.1467-9221.2008.00622.x](https://doi.org/10.1111/j.1467-9221.2008.00622.x)
- Gunaratne, K.G., Coomes, E.A., & Haghbayan, H. (2019). Temporal trends in anti-vaccine discourse on twitter. *Vaccine*, 37, 4867-4871. <https://doi.org/10.1016/j.vaccine.2019.06.086>
- Herrera-Peco, I., Núñez, C.R., Jiménez-Gómez, B., Romero-Magdalena, C.S., & Gracia, E.B. (2021). COVID-19 y vacunación: análisis del papel de las instituciones públicas en la difusión de información a través de Twitter. *Revista española de salud pública*, 95, e202106084
- Jiang, X., Hwang, J., Su, M.H., Wagner, M.W., & Shah, D. (2022). Ideology and COVID-19 vaccination intention: perceptual mediators and communication moderators. *Journal of Health Communication*, 27(6), 416-426. <https://doi.org/10.1080/10810730.2022.2117438>
- Jockers, M. (2017). *Syuzhet, extracts sentiment and sentiment-derived plot arcs from text*. <https://rdrr.io/cran/syuzhet>
- Keib, K., Himelboim, I., Han, J.Y. (2019). Important tweets matter: Predicting retweets in the #Black-Lives-Matter talk on Twitter. *Computers in Human Behavior*, 85, 106-115. <https://doi.org/10.1016/j.chb.2018.03.025>
- Kenez, S., O'Halloran, P., & Liamputtong, P. (2015). The portrayal of mental health in Australian daily newspapers. *Australian and New Zealand Journal of Public Health*, 39(6), 513–517. <https://doi.org/https://doi.org/10.1111/1753-6405.12441>
- Lazer, D.M., Baum, M.A., Benkler, Y., Berinsky, A.J., Greenhill, K., Menczer, F., & Zittrain, J.L. (2018). The science of fake news. *Science*, 359(6380), 1094-1096. <https://doi.org/10.1126/science.aao2998>
- Marroquín, B., Vine, V., & Morgan, R. (2020). Mental health during the COVID-19 pandemic: Effects of stay-at-home policies, social distancing behavior, and social resources. *Psychiatry Research*, 293, 113419. <https://doi.org/10.1016/j.psychres.2020.113419>
- McGinty, E.E., Kennedy-Hendricks, A., Choksy, S., & Barry, C.L. (2016). Trends in news media coverage of mental illness in the United States: 1995–2014. *Health Affairs*, 35(6), 1121-1129. <https://doi.org/10.1377/hlthaff.2016.0011>
- Mittal, S., & De Choudhury, M. (2023, April). Moral framing of mental health discourse and its relationship to stigma: A comparison of social media and news. In *Proceedings of the 2023 CHI conference on human factors in computing systems* (pp. 1-19). <https://doi.org/10.1145/3544548.3580834>
- Mohammad, S.M. (2022). NRC Word-Emotion Association Lexicon (aka EmoLex). Saifmohammad.com. <https://www.saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>
- Mohammad, S.M., & Turney, P.D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3), 436-465. <https://doi.org/10.1111/j.1467-8640.2012.00460.x>
- O'Brien, A. (2021). Reporting on mental health difficulties, mental illness and suicide: Journalists' accounts of the challenges. *Journalism*, 22(12), 3031-3047. <https://doi.org/10.1177/1464884920903855>
- Pavlova, A., & Berkers, P. (2020). "Mental Health" as Defined by Twitter: Frames, Emotions, Stigma *Health Communication*, 37(5), 637-647, <https://doi.org/10.1080/10410236.2020.1862396>
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In: R Plutchik, H Kellerman (Eds.) *Emotion: Theory, Research, and Experience*. Vol. 1. Theories of Emotion, Academic Press, pp 3-33

- Robinson, P., Turk, D., Jilka, S., & Cella, M. (2018) Measuring attitudes towards mental health using social media: Investigating stigma and trivialisation. *Social Psychiatry and Psychiatric Epidemiology*, 54(1), 51–58. <https://doi.org/10.1007/s00127-018-1571-5>
- Rossbroich, J., Durieux, J., & Wilderjans, T.F. (2022). Model selection strategies for determining the optimal number of overlapping clusters in additive overlapping partitioning clustering. *Journal of Classification*, 39(2), 264–301. <https://doi.org/10.1007/s00357-021-09409-1>
- Salud Mental España (2021). *Salud mental y Covid-19. Un año de pandemia*. <https://www.consalmudmental.org/publicaciones/Salud-mental-covid-aniversario-pandemia.pdf>
- Sautera, D.A., Eisner, F., Ekman, P., & Scott, S.K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, 107(6), 2408–2412. <https://doi.org/10.1073/pnas.0908239106>
- Suarez-Lledo, V., Alvarez-Galvez, J. (2021). Prevalence of health misinformation on social media: systematic review. *Journal of Medical Internet Research*, 23(1), e17187. <https://doi.org/10.2196/17187>
- Swati, U., Pranali, C., & Pragati, S. (2015). Sentiment analysis of news articles using machine learning approach. *International Journal of Advances in Electronics and Computer Science*, 2(4), 114–116. http://www.ijae.in/journal/journal_file/journal_pdf/12-127-1430132488114-116.pdf
- Swire-Thompson, B., & Lazer, D. (2020). Public health and online misinformation: Challenges and recommendations. *Annual Review of Public Health*, 41(1), 433–51. <https://doi.org/10.1146/annurev-publhealth-040119-094127>
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, 37(2), 267–307. https://doi.org/10.1162/coli_a_00049
- Tandoc Jr, E.C., Ling, R., Westlund, O., Duffy, A., Goh, D., & Zheng Wei, L. (2018). Audiences’ acts of authentication in the age of fake news: A conceptual framework. *New Media & Society*, 20(8), 2745–2763. <https://doi.org/10.1177/1461444817731756>
- Tavoschi, L., Quattrone, F., D’Andrea, E., Ducange, P., Vabanesi, M., Marcelloni, F., & Lopalco, P.L. (2020). Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy. *Human Vaccines & Immunotherapeutics*, 16(5), 1062–1069. <https://doi.org/10.1080/21645515.2020.1714311>
- Tighe, P.J., Goldsmith, R.C., Gravenstein, M., Bernard, H.R., & Fillingim, R.B. (2015) The painful tweet: text, sentiment, and community structure analyses of tweets pertaining to pain. *Journal of Medical Internet Research*, 17(4), e84. <https://doi.org/10.2196/jmir.3769>
- Universidad de Salamanca (2008). *Lexicon: Guía internacional sobre salud mental para los medios de comunicación*. Instituto Universitario de Integración en la Comunidad. <https://sid-inico.usal.es/documentacion/lexicon-guia-internacional-sobre-salud-mental-para-los-medios-de-comunicacion/>
- Van Beveren, L., Rutten, K., Hensing, G., Spyridoula, N., Schønning, V., Axelsson, M., & Vandamme, J. (2020) A critical perspective on mental health news in six European countries: How are “Mental Health/Illness” and “Mental Health Literacy” rhetorically constructed? *Qualitative Health Research*, 30(9), 1362–1378. <https://doi.org/10.1177/1049732320912409>
- Vyncke, B., & Van Gorp, B. (2018). An experimental examination of the effectiveness of framing strategies to reduce mental health stigma. *Journal of Health Communication*, 23(10–11), 899–908. <https://doi.org/10.1080/10810730.2018.1538272>
- Walker, K.K., & Burns, K. (2019). #13 ReasonsWhy health professionals and educators are tweeting: A systematic analysis of uses and perceptions of show content and learning outcomes. *Health Communication*, 34(10), 1085–94. <https://doi.org/10.1080/10410236.2018.1461171>
- Wang, X., & Lee, E.W. (2020). Negative emotions shape the diffusion of cancer tweets: toward an integrated social network–text analytics approach. *Internet Research*, 31(2), 401–418. <https://doi.org/10.1108/INTR-04-2020-0181>

- Weeks, B.E., & Holbert, R.L. (2013). Predicting dissemination of news content in social media: A focus on reception, friending, and partisanship. *Journalism & Mass Communication Quarterly*, 90(2), 212-232. <https://doi.org/10.1177/1077699013482906>
- World Health Organization (WHO) (2021). *Comprehensive Mental Health Action Plan 2013–2030*. <https://www.who.int/publications/i/item/9789240031029>
- Yoo, W. (2019). How Risk Communication via Facebook and Twitter Shapes Behavioral Intentions: The Case of Fine Dust Pollution in South Korea. *Journal of Health Communication*, 24, 663-673
- Zheng, H., Goh, DH.-L., Lee, E.WJ., Lee, C.S., & Theng, Y.-L. (2022). Understanding the effects of message cues on COVID-19 information sharing on Twitter. *Journal of the Association for Information Science and Technology* 73(6):847–862. <https://doi.org/10.1002/asi.24587>
- Zunic, A., Corcoran, P., & Spasic, I. (2020) Sentiment analysis in health and well-being: systematic review. *JMIR Medical informatics*, 8(1), e16023. <https://doi.org/10.2196/16023>