Exploring the Public's Perception of Gambling Addiction on Twitter During the COVID-19 Pandemic: Topic Modelling and Sentiment Analysis

Emanuele Fino[1], Bishoy Hanna-Khalil[2], Mark D. Griffiths[1]

Author Note

[1]Department of Psychology, Nottingham Trent University, Nottingham, United Kingdom

[2]Queen Mary University of London, School of Medicine, London, United Kingdom

Name, mailing address and email address of author responsible for correspondence about the manuscript:

Dr Emanuele Fino

Department of Psychology

Nottingham Trent University

NG1 4FQ

Nottingham

United Kingdom

+44(0)115 84 84282

Email: emanuele.fino@ntu.ac.uk

Exploring the Public's Perception of Gambling Addiction on Twitter During the COVID-19 Pandemic: Topic Modelling and Sentiment Analysis

Abstract

The present study explored the topics and sentiment associated with gambling addiction during the COVID-19 pandemic, using topic modeling and sentiment analysis on tweets in English posted between 17-24th April 2020. The study was exploratory in nature, with its main objective consisting of inductively identifying topics embedded in user-generated content. We found that a five-topic model was the best in representing the data corpus, including: (i) the public's perception of gambling addiction amid the COVID-19 outbreak, (ii) risks and support available for those who stay at home, (iii) the users' interpretation of gambling addiction, (iv) forms of gambling during the pandemic, and (v) gambling advertising and impact on families. Sentiment analysis showed a prevalence of underlying fear, trust, sadness, and anger, across the corpus. Users viewed the pandemic as a driver of problematic gambling behaviors, possibly exposing unprepared individuals and communities to forms of online gambling, with potential long-term consequences and a significant impact on health systems. Despite the limitations of the study, we hypothesize that enhancing the presence of mental health operators and practitioners treating problem gambling on social media might positively impact public mental health and help prevent health services from being misused, in times when healthcare resources are limited.

Keywords: Gambling; Addiction; Topic Modeling; Sentiment Analysis; Twitter.

Exploring the Public's Perception of Gambling Addiction on Twitter During the COVID-19 Pandemic: Topic Modelling and Sentiment Analysis

1. Introduction

1.1 General Background

The recent, global affirmation of online social media has dramatically changed the way individual represent their health and share relevant information (Zhang et al., 2018). de Rosa et al. (2013) defined social media as contemporary "social arenas" (p. 16), online virtual communities where users gather to collectively generate and share social representations, in the attempt to make sense of the world, acquire information, and turn it into actionable knowledge. In the past decade, research has increasingly acknowledged the potential of social media to inform public and global health policy, focusing on the analysis of online user-generated content (de Martino, 2013; Sansone et al., 2019; Stellefson et al., 2015). Findings highlight the role of social media as key in health communication, allowing users to share information on prevention and treatment of diseases and engage in mutual support, ultimately to increase control over their physical and mental health (de Rosa et al., 2013; Zhang et al., 2018).

1.2 Social Media and Public Health

Social media have played a pivotal role during the COVID-19 pandemic, especially in the early phases (Hinjoy et al., 2020), allowing a large number of distanced individuals from different countries to stay connected, interact, and keep updated on the progression of the pandemic (Limaye et al., 2020). The novel coronavirus-2019 (COVID-19) is the acute respiratory disease caused by the novel Sars-CoV-2 virus. Since early cases were observed in China in November 2019, the infection spread rapidly across several countries, triggering the declaration of pandemic by the World Health Organization (WHO) on the 11th of March 2020

(Hu et al., 2020). In the attempt to contain the spread of the disease, several national and local governments introduced preventative measures, including, but not limited to, confinement, quarantine, travel and spatial distancing restrictions, commonly referred to as lockdown (Ibarra-Vega, 2020).

Studies show that lockdown measures represent a significant challenge to individual's lives and their mental health, requiring the adoption of novel strategies to cope with symptoms of stress, anxiety, depression, insomnia (Kar et al., 2020; Rossi et al., 2020), substance abuse (Mackolil & Mackolil, 2020), and addictive behaviors, including gambling addiction (Kiràly et al., 2020). Dubey et al. (2020) recently defined COVID-19 and addiction in terms of two colliding pandemics, with future impact on public health systems yet to be estimated. Similarly, Kar et al. (2020) argued that spatial distancing measures may represent a threat for a number of individuals at risk, on the one hand pushing them to engage in excessive indoor behaviors, such as consuming more media and games, while on the other, increasing the likelihood of experiencing withdrawal symptoms, especially in the light of the current unavailability of in-person mental health services.

Some specific features of the pandemic and associated restrictions might have an impact on gambling behaviors (e.g., Mallet et al., 2021). For example, recent studies indicated that lockdown policies affected the mental health of individuals, interfering with and disrupting personal, social, financial, and occupational aspects of their lives (Mallet et al., 2021). Some authors have argued that such unprecedented changes may play a significant role in consolidating or generating new patterns of addictive behaviors to overcome stress, anxiety, and preoccupations (Kiràly et al., 2020). As discussed by Mallet et al. (2021), "[s]ocial isolation during spatial distancing, and these stressors in conjunction with substantial changes in gambling markets (land-based, online) during the pandemic may significantly influence gambling behaviors" (p. 5). Moreover, spatial distancing measures and telehealth may represent a challenge for individuals affected by addictive syndromes, particularly for those unable to access online treatments (Satre et al., 2020).

1.3 Gambling Addiction

Gambling addiction is viewed as being at the extreme end on a continuous scale of disordered gambling behaviors, resulting in detrimental psychosocial, mental, and financial consequences (American Psychiatric Association, 2013; Griffiths, 2016). The past decade has seen a proliferation in various forms of online gambling due to the diffusion of computers and mobile devices (Gainsbury, 2014). These provide users with accessible, immersive gambling interfaces, and for these reasons it has been hypothesized that online gambling may be comparatively more addictive than offline gambling (Abbott, 2020; Binde et al., 2017; Gainsbury, 2015; Gainsbury et al., 2016).

However, empirical findings from a survey conducted by Wardle et al. (2011) in a representative sample of British adults, showed that most online gamblers tended to gamble offline too, displaying both high involvement with gambling activities and high rates of problem gambling. This means that the online environment may represent a threat for those with problem gambling, as they may use the internet to pursue a variety of accessible gambling opportunities, presumably even more in times when land-based gambling establishments are not available.

1.4 Gambling Addiction During the COVID-19 Pandemic

It has been argued that the lockdown measures implemented to counter the spread of the COVID-19 pandemic may represent a possible driver of gambling behaviors, enhancing risks for a large number of individuals. For example, Bowden-Jones, (2020 as cited in Davies, 2020) has argued that “many more will be tempted by boredom and financial concerns to try gambling as a way out” (p.1). However, recent studies on online gambling trends during the pandemic present contrasting evidence. Price (2020) showed that high-risk gamblers among a larger group of online gamblers had a higher likelihood of gambling during the 2020 Canadian lockdown. A study by Håkansson (2020) carried out in May 2020, using self-reports from Swedish participants, showed that only a minority of the interviewees presented increased gambling activity, and those were already at a higher risk of problem gambling and of alcohol-related problems. However, as pointed out by Mallet et al. (2021), Sweden is a country where lockdown measures were not implemented. A study by Auer et al. (2020), from a study on a behavior tracking dataset of online sports gamblers (*N* = 5,396) from a European online gambling operator in Sweden, Germany, Finland, and Norway, showed (i) a significant reduction in the amount of money spent in sport betting, compared to the pre-pandemic period, (ii) no significant conversion of sport betting into more online casino playing, and (iii) a significant decrease in online casino playing for sport bettors. Another, more recent study by Auer and Griffiths (2021) analyzed behavioral tracking data among a sample of over 133,000 online casino Swedish gamblers, with results indicating overall decreased gambling intensity. In summary, considering all the available evidence, it might be sensible to hypothesize that gambling trends and patterns in response to the pandemic may vary, and that variations may depend on the characteristics of the observed population as well as on the public health measures put in place in different cultural and national contexts.

Nevertheless, some authors have argued that guidelines to help counter the possible detrimental effects of enhanced internet use during the pandemic, including online gambling activity, are needed (Király et al., 2020; see also the review by Mallet et al., 2021). In particular, Király et al. (2020) highlighted that although several behaviors, under normal circumstances, could help individuals mitigate stress, anxiety, and preoccupations, without necessarily leading to addictions, the current spatial distancing measures and inability to perform a number of common daily activities may turn such behaviors into new, persistent practices and subsequently, into behavioral addictions, with potentially dramatic mental health implications.

In light of the contrasting evidence on the prevalence and patterns of gambling behaviors during the pandemic across different national contexts, it is of foremost importance to study the public's perception of behavioral addictions. More specifically, the unprecedented circumstances are favoring the emergence of new work-life routines, balance, and patterns of behavior, where individuals are encouraged to stay at and work from home, whenever possible (Vasudevan et al., 2021). In this scenario, user-generated content could represent a key source of information, aiming to shed light on topics and sentiment underlying the discourse that is produced and shared globally on online "social arenas" such as Twitter (de Rosa et al., 2013). This could provide a baseline to examine how the pandemic restrictions are affecting individuals and communities, in an attempt to tackle gambling addiction as a somewhat hidden mental health condition and contribute to inform public health policy.

1.5 The Role of Twitter in Public Health Communication

A growing corpus of literature has investigated the role of social media in conveying information on mental health, particularly Twitter (Charles-Smith et al., 2015; Himelboim & Han, 2013; Sugawara et al., 2012). Twitter is one of the most popular and utilized social media in a large number of countries (Sansone et al., 2019), providing a social microblogging platform where users can post and interact by means of short text messages called tweets, limited to 280 characters. As of the first quarter of 2019, Twitter had an average of 330 million monthly active users (Clement, 2019).

Recent studies by Houghton et al. (2020) and Killick and Griffiths (2020) have shown a significance presence of gambling operators on Twitter, using a variety of strategies to link tweets with major sporting events, ultimately leading to an arguably non-responsible promotion of gambling services. This poses questions regarding the possible implications for the public, particularly whether poorly controlled gambling advertising contributes to instigating a fundamental negative sentiment towards the gambling industry and the capacity of national and local institutions to regulate it (Wood & Griffiths, 2008). This may have an impact onto public mental health, especially in the current pandemic, based on the evidence that lack of trust tends to be associated to lower adherence to public health measures (Blair et al., 2017), possibly carrying a vicious cycle of "distrust, non-compliance, hardships and further distrust" (p.89).

1.6. Research Aims

The present study explored the public's perception of gambling addiction on a corpus of tweets in English posted in April 2020. We used topic modeling and sentiment analysis to analyze the data. Topic modelling allows for the exploration of semantic patterns and latent meaning emerging from text corpora, rather than testing a set of predetermined hypotheses on the data (Vallurupalli & Bose, 2020). This makes it "a tool for theory development through exploratory research, unlike traditional quantitative techniques which are mostly used for testing of theories." (p. 794). Therefore, research using topic modelling is "exploratory in nature", with its main aim consisting of identifying topics embedded in a corpus (p. 791). Consistently, we used topic modeling and sentiment analysis to analyze Twitter user-generated content and to explore the main topics and sentiment associated with gambling addiction within a specific time frame, rather than testing hypotheses on the prevalence or the characteristics of gambling trends during the pandemic. The specific, relatively short time-frame was chosen to account for approximately one month following the WHO's (2020) pandemic declaration and the beginning of spatial distancing restrictions in several countries, including the suspension of traditional land-based gambling establishments.

2. Materials and Methods

Data mining consisted of three main steps: (i) data search and collection, (ii) data pre-processing, and (iii) data analysis. The following sections include a brief explanation of each step and the inclusion and exclusion criteria applied at each level of parsing through the corpus.

2.1 Step 1: Data Search and Collection

We used Twitter's Representational State Transfer (REST) Application Programming Interface (API) and the rtweet R package (Kearney, 2019) to search and collect tweets. Twitter's REST API is a freely available user interface that is used in data mining and machine learning applications as an endpoint to search and collect Twitter data, providing information and content from individual accounts and allowing researchers to filter content by means of selected keywords (Twitter, 2021). rtweet is a package based on R statistical programming language that provides several functions to extract and filter data from Twitter's REST and streaming APIs. This is a previously peer-reviewed instrument (Kearney, 2019) which has been already used in several studies where data mining and machine learning applications were used to inform public health policy (Aguillar-Gallegos et al., 2020; Mariano et al., 2021; Ogundepo et al., 2020; Sansone et al., 2019; Trovato et al., 2020).

Specifically, we searched for tweets posted between 17-24th April 2020, in English, using the following search terms: gambl* AND addict*, excluding retweets. We chose this specific period as the majority of countries had implemented lockdown measures in a time-period between the first and the second half of March 2020 (Dunford et al., 2020). This allowed for approximately one month for observing meaningful, prevalent topics and sentiment.

Following, we ran a preliminary search and investigated results. We observed a significant number of tweets that were not relevant to our analysis, due to the fact that terms such as ‘gambling’ and ‘addiction’ are often used metaphorically in everyday language, representing a variety of meanings (Lopez-Gonzalez et al., 2018). More specifically, these tweets mentioned ‘gambling’ and ‘addiction’ as figures of speech to comment on the conduct of public individuals such as politicians and sports player, with little or no relation to the context of mental health. For these reasonswe reviewed the search options by carefully excluding a selection of terms that in all cases had led to non-meaningful tweets (i.e., -jordan -michael -mike -mikes -mjs -documentary -lastdance -mikebow -mj -trump -james -ryzin). The final search returned 371 tweets. We performed a further, manual inspection of results, and this time we observed a corpus that was more internally consistent, showing a high prevalence of tweets relevant to gambling addiction, as required. However, in one tweet the word gambling was used figuratively to describe drug abuse, and therefore we decided to remove it. After removing duplicates, we retained a final corpus of 144 unique tweets, posted by total 143 different users (Figure 1).

All data were collected according to internet and Twitter research ethical guidelines. An institutional research committee approved the use of the data for the purposes of the study.

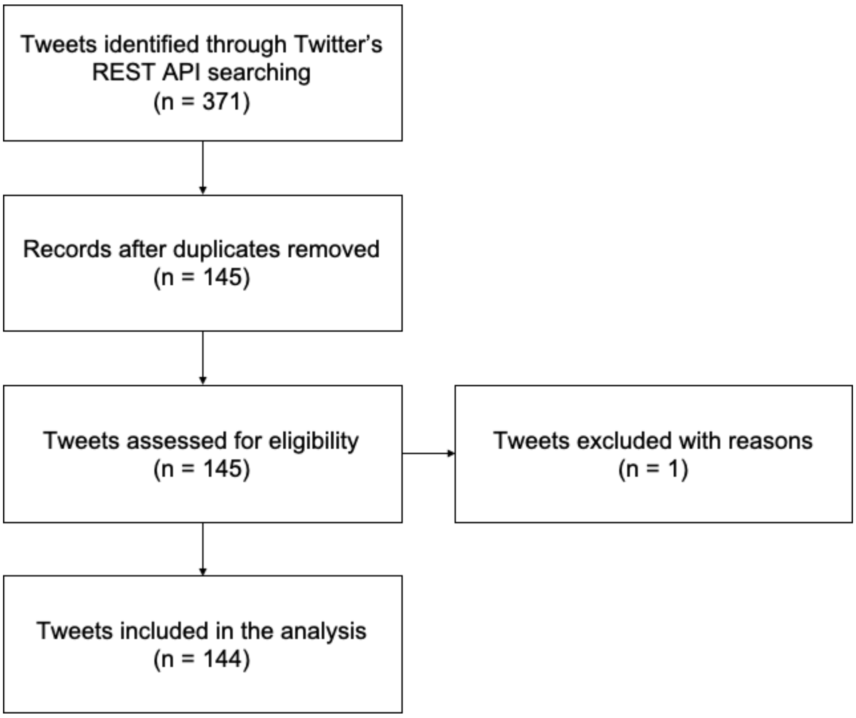

Fig 1. Data collection method.

2.2 Step 2: Data Pre-Processing

Consistent with recommendations from the literature (Bickel, 2019; Symeonidis et al., 2018), the text corpus was prepared for the analyses by performing a sequence of pre-processing operations. In particular, at this stage, the following refinement criteria were applied to the collected tweets, in the following order: Removal of hashtags, due to their relatively small number across the whole corpus; removal of unicode strings (e.g., strings like "\u002c" and "\x06"), accidentally retained in the extraction procedure; replacement of URLs and user mentions; replacement of abbreviations (e.g., "ty" with "thank you"); replacement of contractions (e.g., "won't" and "don't" with "will not" and "do not", respectively); removal of numbers; removal of special characters (e.g., ampersands); removal of punctuation; lowercasing all words; spelling correction, where required. Subsequently, we proceeded with the semi-automated removal of English stopwords and manual lemmatization of the corpus. Stopwords are common words in English, but with little or no relevance to the subject of interest (e.g., words like "and", "or", "however"). We used a list combining stopwords from the tidyverse R package (Wickham et al., 2019) and additional custom words that we had previously identified through manual inspection. At this stage, exclusion criteria encompassed filtering for redundant, non-substantive semantic content relative to the domain of analysis (for example, but not limited to, words such as: “basically”, “woulda”, “possibly”, “couldve”, “gonna”). Lemmatization consists of the analysis of words’ morphology and the replacement of inflectional endings with their base forms (Symeonidis et al., 2018). Finally, in order to predispose the corpus for the topic analysis, we restructured it into a tokenized data frame using the tidytext R package (Silge et al., 2016).

2.3 Step 3: Data analysis: Topic Modeling and Model Selection

We used topic modelling and sentiment analysis to analyze the data. Topic modeling is a “method for unsupervised classification of such documents, similar to clustering on numeric data, which finds natural groups of items even when we’re not sure what we’re looking for” (Silge & Robinson, 2017, Chapter 6). Several methods and algorithms are available, each suiting specific types of corpora. We used Biterm Topic Modelling (BTM) (Yan et al., 2013).

Biterm is a topic modelling algorithm that classifies words and documents from a text corpus into a smaller number of latent topics, being particularly suitable for short texts such as tweets (Blei, 2012). Specially, the algorithm generates a set of topics from biterms, namely patterns of unordered word co-occurrences within the text. In doing so, BTM differs from other methods based on word-document co-occurrences (e.g., Latent Dirichlet Allocation; Blei et al., 2003). In fact, short texts are characterized by a significant degree of sparsity, due to their limited amount of information. BTM generates topics from co-occurring, unordered word-pairs (biterms), ultimately tackling the issue of sparsity. This allows for the generation of more coherent topics, and consistently, research shows that BTM outperforms other established methodologies in short texts such as tweets (Yan et al., 2013).

Aiming to identify the best-fitting model, we ran 99 different BTM analyses by progressively varying the number of topics (k) anticipated for the corpus (from 2 to 100), and eventually compared the results across models. We used the collapsed Gibbs sampling Markov chain Monte Carlo algorithm (Qiu et al., 2014; Yan et al., 2013) with five iterations per model via the BTM R package (Wijffels, 2020). The algorithm assigns biterms to topics based on the conditional distributions of word co-occurrences, using the first extracted topic as a background topic, corresponding to the empirical word distribution. In addition to k, the algorithm requires the specification of two hyperparameters, namely α (i.e., the symmetric Dirichlet prior probability of a topic), representing the granularity of topic distribution over the corpus, and β (i.e., the symmetric Dirichlet prior probability of a word given the topic), representing the specificity of topics (Bickel, 2019). The outcome of the process consists of a model identified by a topic-word probability distribution (ϕ) and an overall topic probability distribution (θ), the latter quantifying the proportion of biterms per topic (Yan et al., 2013). In summary, a given topic probability is determined by the proportion of biterms accounted by that topic over the whole corpus (e.g., a topic probability of 0.60 indicates that a given topic explains more than half of all word co-occurrences in the corpus).

In order to select and retain the best model, we used two complementary approaches: (i) We estimated a measure of topic coherence and exclusivity, and (ii) we manually inspected the model to make sure this was interpretable. In fact, evidence from previous studies suggests that automated detection alone does not allow for optimal decision-making over cut-off values, warranting expert manual inspection and evaluation in any cases (Bickel, 2019). We first estimated topic exclusivity by adapting the relevant function from the stm R package (Roberts et al., 2019) to BTM. This is a measure of relative word exclusivity and coherence across topics, originally developed for LDA (Bischof & Airoldi, 2012; Roberts et al., 2019). We used the mean to aggregate topic scores per model and normalized results to facilitate their interpretation. Second, we manually inspected the model by examining the topics extracted and their associated words.

Regarding the two hyper-parameters of BTM, commonly used values are $\alpha = 50/k$ and $\beta = 0.1$ (Griffiths & Stayvers, 2004). However, smaller values are recommended when dealing with distributions that are sparse over topics (Bickel, 2019; Sun & Yin, 2017), and for this reason we fixed values to $\alpha = 5/k$ and $\beta = 0.01$. Finally, we performed a sentiment analysis of the corpus by means of the syuzhet R package (Jockers, 2015). Sentiment analysis clusters the corpus into smaller, comprehensive categories, based on pre-coded dictionaries, with each category representing a distinct, prevalent emotion or sentiment (Sansone et al., 2019). In particular, we used sentiment analysis to detect the prevalent emotional tone associated to the discourse on gambling addiction on Twitter, aiming to improve the interpretation of results from the analysis based on the classification of information in topics (Yada et al., 2020). We used the NRC Word-Emotion Association Lexicon dictionary (Mohammad & Turney, 2013), including a list of English words and their associations with eight basic clusters of emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust, respectively). We used the probabilistic inference obtained through topic modelling to describe the semantic patterns underlying the text and to interpret the topics, and we supported the interpretation by performing a qualitative analysis of text when required.

3. Results

Table 1 shows a summary of the characteristics of the corpus of tweets. Figure 2 displays the most frequent words observed in the corpus, in decreasing order.

Table 1. Characteristics of tweets (N = 144)

Variables		
Location (count, %)	Antigua & Barbuda	1 (0.69%)
	Australia	3 (2.08%)
	Brazil	1 (0.69%)
	Canada	4 (2.78%)
	Estonia	1 (0.69%)
	India	4 (2.78%)
	Kenya	1 (0.69%)
	Mexico	1 (0.69%)
	Netherlands	1 (0.69%)
	Nigeria	3 (2.08%)
	Not available	36 (25%)
	Portugal	1 (0.69%)
	Puerto Rico	1 (0.69%)
	Republic of the Philippines	2 (1.39%)
	South Africa	3 (2.08%)
	UK	24 (16.67%)
	United Arab Emirates	1 (0.69%)
	USA	55 (38.19%)
	Zambia	1 (0.69%)
Date of creation	Earliest	17/04/2020
	Latest	24/04/2020
Width (characters)	Min	50
	Average	125.58
	Max	212
Hashtags	Min	0
	Average	0.15
	Max	7
Retweets	Min	0
	Average	13.88
	Max	981
Favorites	Min	8
	Average	30375.41
	Max	371449

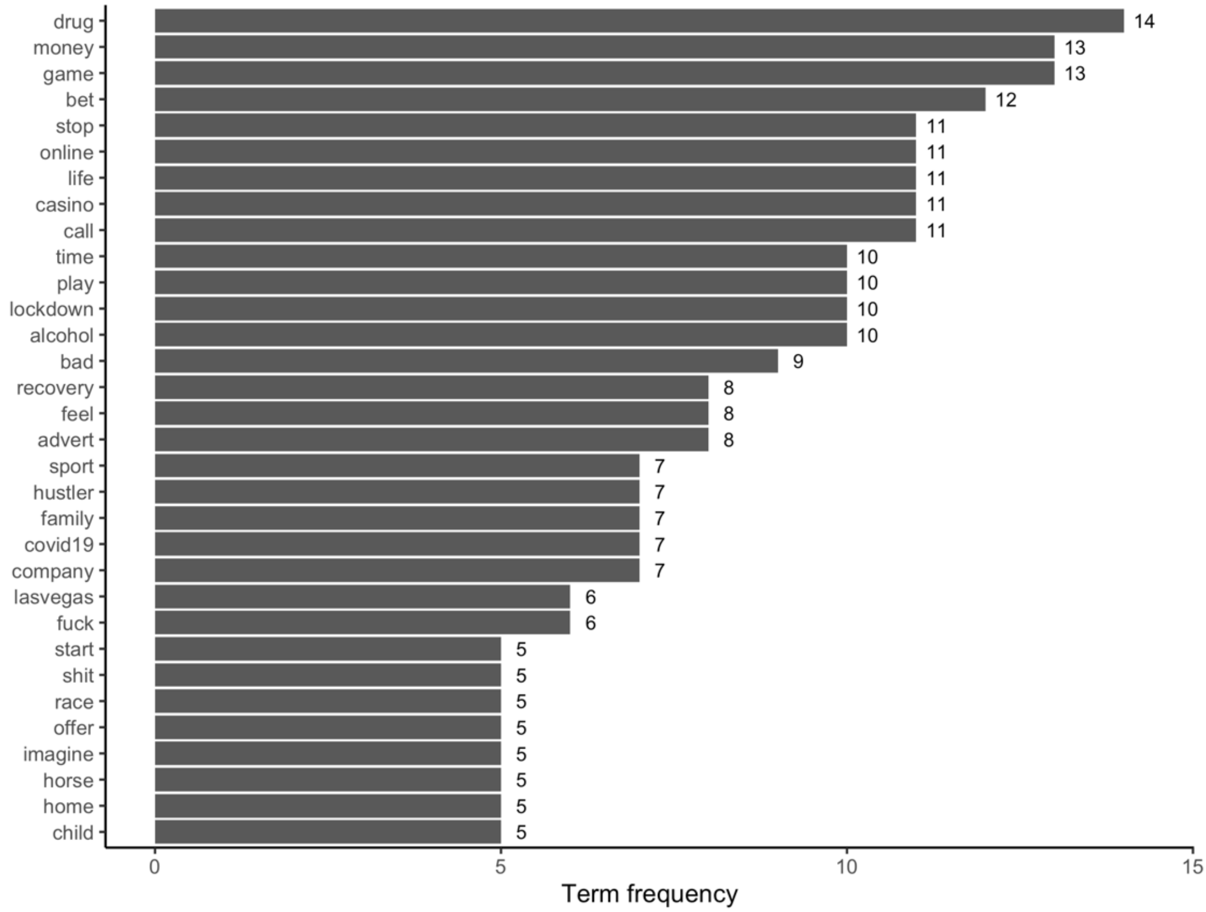

Fig 2. Most frequent words (> 4) observed in the corpus of tweets, excluding the search terms.
Note: The words "gambl*" and "addict*", occurring 174 and 171 times in the corpus, respectively, were removed from the graph to facilitate its interpretation.

The results of topic modelling and the distribution of normalized exclusivity across models are presented in Figure 3. We observed a major rise in topic exclusivity between models accounting for 10 to 20 topics, respectively, peaking at approximately 40 topics. However, the manual inspection challenged the interpretability of those models, showing high sparsity and inconsistencies within and between topics. Instead, we found that a model accounting for five topics (mean exclusivity across topics = 25.22; normalized mean exclusivity = 0.46) was the best to represent the corpus, maximizing word exclusivity, topic internal consistency, and the overall model's interpretability. Therefore, we decided to retain the five-topic model for further analyses and interpretation.

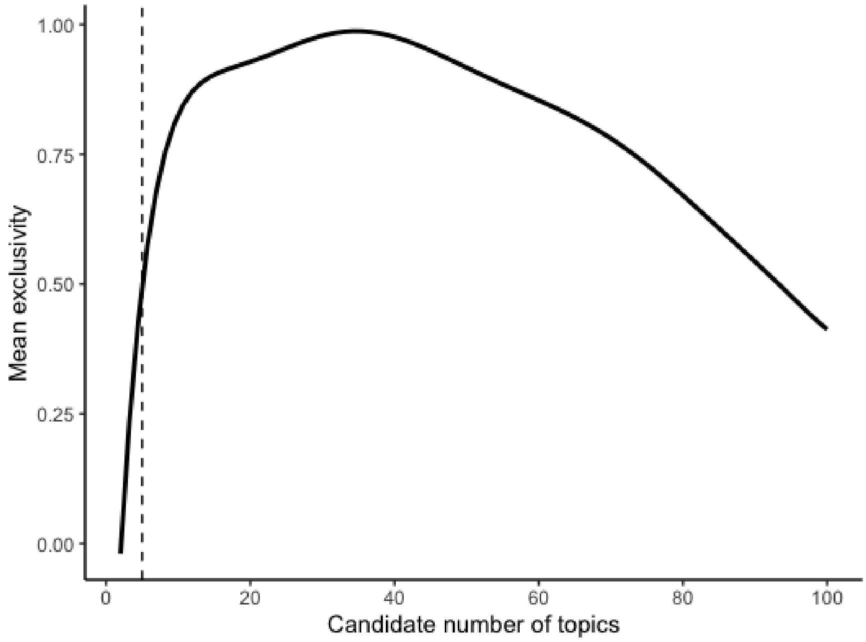

Fig. 3. Mean exclusivity computed across k = 99 topic models (2-100). The continuous line represents the smoothed exclusivity curve. The dotted line represents the selected cut-off (5 topics).

Figure 4 presents the network model of relations across words, resulting from the analysis of top 100 bigrams across the corpus of tweets, namely the most common sequences of consecutive words observed within tweets (Silge & Robinson, 2017).

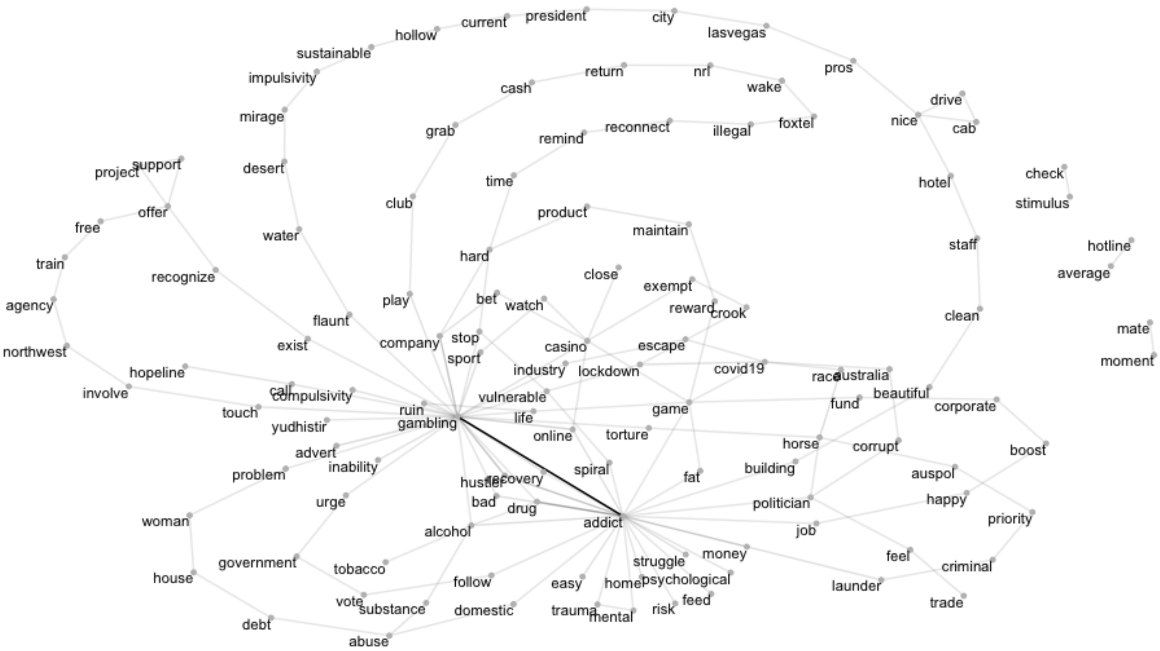

Fig. 4. Model of relations between words (bigrams) from the corpus of tweets. The width of the connecting lines is proportional to the frequency of co-occurrences between words (i.e., a thicker line indicates a more frequent word co-occurrence, as in the case of the words "gambling" and "addict").

In the following paragraphs, we present each topic individually, focusing on their top 25 words (i.e., words with highest co-occurrence probability) (Table 2). Aiming to present a meaningful interpretation of topics, we used portions of text from tweets as elementary context units (Reinert, 1983) (i.e., significant parts of text as defined by words co-occurrences).

Table 2. Summary of topic modelling

Topic N	**Mean Topic Probability (θ)**	**Top 25 most co-occurring words (ordered decreasingly)**	**Main themes**
1	0.40	addict, gambling, money, drug, game, life, casino, stop, time, bet, play, alcohol, call, online, family, bad, lockdown, advert, feel, recovery, sport, vulnerable, covid19, ass, company	Gambling addiction amid the COVID-19 outbreak
5	0.18	gambling, addict, stop, drug, sport, play, money, game, million, online, alcohol, hard, capitalism, company, pachinko, ruin, start, follow, life, contribute, fifa, happy, inability, casino, abuse	Risks and support available for those who stay at home
2	0.15	gambling, addict, parent, bad, advert, call, fuck, offer, night, easy, house, abuse, billion, mate, agency, home, project, heroin, hurt, shit, involve, northwest, stop, gadget, train	The user's psychology of gambling addiction
3	0.14	gambling, addict, hustler, game, bet, lockdown, call, recovery, feel, online, trade, company, average, hotline, increase, support, psychological, time, casino, dopamine, assassinate, city, impulsivity, reach, secret	Forms of gambling at the time of pandemic lockdown
4	0.13	addict, gambling, covid19, casino, watch, lasvegas, lockdown, alcohol, horse, race, nice, league, tv, crook, desert, exempt, flaunt, time, water, drive, mirage, parasite, building, clean, home	Gambling advertising and impact on families

3.1. Topic Modeling

3.1.1 Topic 1: Gambling addiction amid the COVID-19 outbreak

Topic 1 showed the highest mean topic probability (0.40). It incorporated words relevant the impact of gambling addiction on individual's "lives" amid the lockdown, particularly in relation to specific forms of gambling such as "bet" and "casino". For example, users engaged in discussions on betting companies, perceived as responsible for reinforcing problem behaviors in individuals with gambling addiction during the lockdown. In this early phase of the pandemic, users reported witnessing increased "marketing" efforts by betting companies through a variety of different media such as TV, radio, and the internet.

Gambling addiction was defined in relation to words such as “money”, “drug”, and “stop”, ultimately comparing the condition to other addictive syndromes like drug abuse or alcoholism.

3.1.2. Topic 2: Risks and support available for those who stay at home

The mean topic probability for Topic 2 was 0.16. This topic included the word “advert”, referring to the detrimental impact of online gambling advertising on the mental health of those who stay at home during the lockdown. Users reported a pervasive presence of gambling adverts and concerns of the possible effects on individuals with gambling addictions. The addictive potential of gambling was compared to that of highly addictive substances such as “heroin”, in an attempt to debunk the view of gambling as a “comfort hobby”. Staying at home was described as a challenge for those who struggled with addictive behaviors, and dedicated support was advertised and promoted, for example, helplines offering advice, information or a listening ear when needed. Ultimately, users outlined a framework of risks and opportunities during the lockdown, in which “home” is perceived as not the “safest place to be”.

3.1.3. Topic 3: The user’s psychology of gambling addiction

The mean topic probability for Topic 3 was 0.14. One of the key themes underlying this topic was the “psychological” progression of gambling at the individual level, to the point of becoming a “devastating addiction”. They suggested that individuals with gambling addiction will not know they have a problem until “it’s too late”, since from a psychological point of view, the condition just seems to “take you over”. The public understanding of gambling addiction in psychological terms centers around the concept of “impulsivity”. Problem gambling was considered as a form of “impulsive behavior” leading to a “psychological addiction”. Interestingly, another word co-occurring with “gambling” and “addiction” in this topic was “hustler”, used in an attempt to correct a common misrepresentation of individuals with gambling addiction, the latter needing help to overcome their condition rather than being someone who engage into illegal activities to increase their wealth.

3.1.4. Topic 4: Forms of gambling at the time of pandemic lockdown

The mean topic probability for Topic 4 was 0.13. Words like “horse” and “race” were present in this topic, functional to questioning which forms of gambling are most common during the pandemic outbreak. Lockdown-related restrictions have affected traditional forms of gambling, triggering a reflection on what consequences they may have on individuals with gambling addiction. Users referred to the recent suspension of professional “sport” leagues and “casinos” in several countries. The most commonly represented forms of gambling during the lockdown were online games, casinos, and esports, replacing traditional sports betting and land-based establishments.

3.1.5. Topic 5: Gambling advertising and impact on families

The mean topic probability for Topic 5 was 0.18. Since the lockdown, users reported seeing a number of gambling and betting webpages posting adverts “on the few sporting fixtures left as well as esports.” Gambling “companies” and operators were criticized for their supposed unethical conduct, particularly for trying to influence individuals who might be at risk of developing gambling problems (“vulnerable”). Some users explicitly attacked sport institutions and commercial operators, viewing them as encouraging adults and children into “gambling” and “addiction” through a number of online initiatives. Addiction was associated with “family”, and gambling to escape boredom was seen as a tangible risk for those who stayed at home, potentially carrying dramatic psychosocial and financial consequences for individuals and their families (e.g., as exemplified in the following excerpt from a tweet: “[…] they’re making the most of people’s addictions and boredom during the lockdown.

'broke lads' addicted to gambling are keeping the parasites going, often at the cost to their own families who go without when these often losers go into to feed their addictions.")

3.2 Sentiment analysis

The results of the sentiment analysis showed that fear was the most common sentiment in the corpus (18.43%), followed by trust (17.20%), sadness (15.48%), and anger (12.29%), highlighting an overall prevalence of negative sentiment (Figure 5).

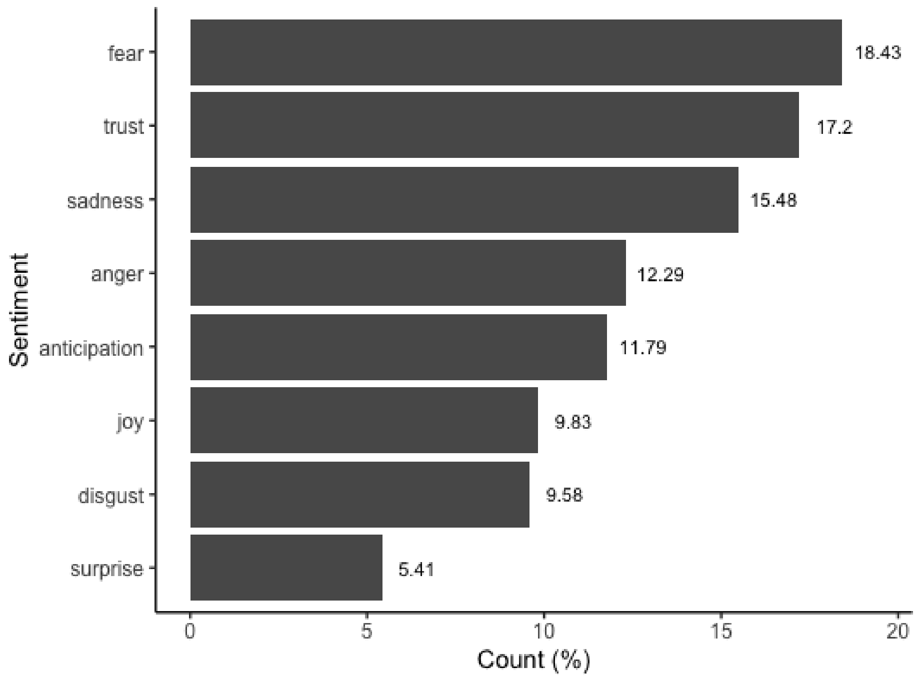

Fig 5. Sentiment analysis based on the whole corpus of tweets.

The less common sentiments observed were surprise (5.41%), disgust (9.58%), and joy (9.83%).

4. Discussion

The main aim of this study was to explore the topics and sentiment associated to gambling addiction during the COVID-19 pandemic lockdown, based on Twitter online user-generated content. We collected a corpus of tweets posted for one week, in April 2020 and we analyzed it by means of topic modelling and sentiment analysis. Results showed that a model accounting for five topics was a good fit to the corpus of tweets. Topics identified were (i) gambling addiction amid the COVID-19 outbreak; (ii) risks and support available for individuals who stay at home; (iii) the user's psychology of gambling addiction; (iv) forms of gambling at the time of the pandemic lockdown; and (v) gambling advertising and impact on families. The results of the sentiment analysis showed that fear was the most common sentiment in the corpus, followed by trust, sadness, and anger, highlighting an overall prevalence of negative sentiment. The less common sentiments observed were surprise, disgust, and joy, respectively.

One of the primary observations emerging from the study was that, at the time when lockdown measures were implemented in several countries, Twitter users viewed gambling mainly as an online activity emerging under spatial distancing measures, with an underlying fear, and sadness towards a number of possible consequences for public mental health. In this regard, although in the present study we did not collect data on the presence of online gambling operators, our qualitative analysis of tweets allowed us to isolate individual users' concerns about their observation of an increase in online presence of gambling adverts (e.g., "[…] mate I'm concerned about the amount of online gambling adverts on at the moment. I have mates that have been or are gambling addicts, it seems like the adverts are on constantly at the moment."), in a time when most traditional land-based gambling establishments had to close in response to the spread of the COVID-19 infection. The results from topic modeling presented in the study indicate that the pandemic lockdown in March 2020 was perceived as a possible driver of problematic gambling behaviors, exposing a large number of unprepared individuals and communities to new forms of gambling, with potential long-term consequences and a significant impact on public health systems.

Another interesting finding from the present study was the perception of the role of boredom, highlighted by Twitter users as main triggers for gambling during the pandemic lockdown, along with financial preoccupations associated with the upcoming economic crisis.

In this regard, the increased amount of time that individuals have due to staying at home, is considered as a key factor. Previous studies have shown that a lack of social interaction among problem gamblers is, in fact, associated on the one hand with a higher likelihood to engage in addictive behaviors, and on the other, to an increased need to reconnect (Gainsbury et al., 2014; 2016; Sirola et al., 2019). However, Auer et al. (2020) suggested that the reduced trends in online gambling that they had observed in the first half of 2020 may be associated with (i) individuals' financial conservatism in view of a possible upcoming economic crisis, (ii) a general reluctance to gamble in front of members of their household, and (iii) the fact that they may be engaging in more "quality time" with their families (p.7). Therefore, although our results indicate a possible misalignment between the public's perception of gambling addiction on Twitter and other empirical studies on gambling during the pandemic, we suggest caution in their interpretation, inviting future research to monitor the discourse on gambling addiction in the next stages of the pandemic, possibly including data from different media sources and multiple national contexts.

For all these reasons, despite the notable limitations of the present study, we hypothesize that enhancing the presence of mental health operators and practitioners treating problem gambling and gambling addiction on social media could positively impact patients and their families in such unprecedent times. Moreover, evidence from recent studies shows that online intervention for problem gamblers improves gambling outcomes (Bücker et al., 2018; Casey et al., 2017). Not only having online support could help addressing the demand for dedicated services, helping individuals fight such dramatic condition, but this will also prevent public health services from being misused or overwhelmed, in times when human and material resources in health care are of foremost necessity (Ransing et al., 2020).

Another interesting finding emerging from the present study was the public's general lack of trust against the gambling industry, and more generally, toward governments and institutions. In fact, trust was one of the main sentiments identified, along with fear, sadness, and anger, suggesting a widespread concern toward the conduct of gambling operators and their ethics in advertising, hitting a high number of exposed and unprepared individuals, globally, and potentially leading to decreased adherence to necessary public health measures (Blair et al., 2017). We believe that such findings could be of specific interest to policymakers, and they could be used to better address efforts and collaboration with commercial operators in the gambling sector, aiming to define clear and sustainable social responsibility guidelines, necessary to regulate online advertising and increase trust among the general public. Unsurprisingly, around the same time our data were collected, the Gambling Commission (2020) provided additional formal guidance for remote gambling operators in the UK during the COVID-19 pandemic, in an attempt to regulate the way gambling operators interact with potential customers, and aimed at minimizing the risk of harms associated with gambling. We believe this represents an important step forward, whose outcomes will deserve further evaluation in the near future, in comparison with similar initiatives in place in other national contexts.

This study has limitations. First, all tweets were collected in English, affecting the generalizability of results to other linguistic contexts. Second, the data were drawn from many countries and it is possible that some countries were in different stages of lockdown at the time the data were collected. Third, the tweets did not necessarily reflect the lived experience of gamblers but may have reflected speculation produced by several diverse users and agencies involved in online communication. Moreover, a search strategy using variations of the words “gambling” and “addiction” may have led to biased results, considering that individuals with addictive gambling behaviors may not necessarily be willing to discuss it publicly. Fourth, there was not a comparison of the data corpus with other datasets collected from a non-COVID period, which would help determine the uniqueness of the semantic associations discussed in the present study. Fifth, data were collected within a limited time, namely one week. This was in response to the need to capture topics and sentiments immediately following the outbreak of the pandemic and in the early phases of the lockdown. However, this may represent a limitation because a number of recent developments of the scenario could have determined fluctuations and changes in topics and sentiments. Finally, although biterm topic modelling was specifically designed to overcome the problem of data sparsity that affects other topic modelling techniques applied to short text (Yan et al., 2013), the number of tweets collected was limited, affecting the generalizability of results.

5. Conclusions

The results here presented suggest a possible shift in the users' perception of the forms of gambling during the lockdown, with increased concerns expressed by individual users about the presence on Twitter of online gambling operators, and the potential effect of online gambling advertising on individuals living under lockdown restrictions. The latter is arguably viewed as unethically promoting their services and potentially influencing individuals, ultimately generating fear, sadness, and lack of trust among the general public.

We consider three main implications and possible applications of these results: (i) they could prompt a reflection among policymakers and practitioners, with regards to best practices and strategies to optimize online communication concerning gambling addiction during the pandemic; (ii) as a consequence, they could inform institutional efforts to promote trust and adherence to prevention practices in the community, ideally in conjunction with evidence from cross-sectional and longitudinal research; and (iii) they could ultimately provide local and national governments with an updated view on the public's perception of gambling addiction, as evident from the analysis of online discourse, possibly suggesting changes to the mental health agenda in the attempt to address the public's demands.

Despite the limited generalizability of the findings, we believe that the results from the present study provide preliminary evidence on the development of topics and sentiments on gambling addiction during the pandemic lockdown. We recommend that future research should investigate specific risk factors of gambling addiction associated to the altered living conditions of individuals, globally, aiming to provide updated public health guidelines and prevent mental health crises during and after the COVID-19 pandemic.

6. References

Abbott, M. W. (2020). The changing epidemiology of gambling disorder and gambling-related harm: Public health implications. *Public Health*, 184, 41-45. https://doi.org/10.1016/j.puhe.2020.04.003

Aguilar-Gallegos, N., Romero-García, L. E., Martínez-González, E. G., García-Sánchez, E. I., & Aguilar-Ávila, J. (2020). Dataset on dynamics of coronavirus on Twitter. *Data in Brief*, *30*, 105684. https://doi.org/10.1016/j.dib.2020.105684

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders, 5th Edition: DSM-5 (5th edition).* American Psychiatric Publishing.

Auer, M., & Griffiths, M. D. (2021). Gambling before and during the COVID-19 pandemic among online casino gamblers: An empirical study using behavioral tracking data. *International Journal of Mental Health and Addiction*. Advance online publication. https://doi.org/10.1007/s11469-020-00462-2

Auer, M., Malischnig, D., & Griffiths, M. D. (2020). Gambling before and during the COVID-19 pandemic among European regular sports bettors: An empirical study using behavioral tracking data. *International Journal of Mental Health and Addiction.* Advance online publication. https://doi.org/10.1007/s11469-020-00327-8

Bickel, M. W. (2019a). Reflecting trends in the academic landscape of sustainable energy using probabilistic topic modeling. *Energy, Sustainability and Society*, *9*(1), 49. https://doi.org/10.1186/s13705-019-0226-z

Binde, P., Romild, U., & Volberg, R. A. (2017). Forms of gambling, gambling involvement and problem gambling: Evidence from a Swedish population survey. *International Gambling Studies*, *17*(3), 490-507. https://doi.org/10.1080/14459795.2017.1360928

Bischof, J., & Airoldi, E. M. (2012). Capturing topical content with frequency and exclusivity. *Proceedings of the 29th International Conference on Machine Learning, 2012,* 9-16. https://icml.cc/2012/papers/113.pdf

Blair, R. A., Morse, B. S., & Tsai, L. L. (2017). Public health and public trust: Survey evidence from the Ebola Virus Disease epidemic in Liberia. *Social Science & Medicine*, *172*, 89-97. https://doi.org/10.1016/j.socscimed.2016.11.016

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77-84 https://doi.org/10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993-1022.

Botterill, E., Gill, P. R., McLaren, S., & Gomez, R. (2016). Marital status and problem gambling among Australian older adults: The mediating role of loneliness. *Journal of Gambling Studies*, *32*(3), 1027-1038. https://doi.org/10.1007/s10899-015-9575-5

Bücker L, Bierbrodt J, Hand I, Wittekind C, Moritz S (2018) Effects of a depression-focused internet intervention in slot machine gamblers: A randomized controlled trial. *PloS One*, *13*(6), e0198859. https://doi.org/10.1371/journal.pone.0198859

Casey, L. M., Oei, T. P. S., Raylu, N., Horrigan. K., Day, J., Ireland, J., & Bonnie, A. (2017). Internet-based delivery of cognitive behaviour therapy compared to monitoring, feedback and support for problem gambling: A randomised controlled trial. *Journal of Gambling Studies*, *33*, 993-1010. https://doi.org/10.1007/s10899-016-9666-y

Charles-Smith, L. E., Reynolds, T. L., Cameron, M. A., Conway, M., Lau, E. H. Y., Olsen, J. M., Pavlin, J. A., Shigematsu, M., Streichert, L. C., Suda, K. J., & Corley, C. D. (2015). Using social media for actionable disease surveillance and outbreak management: A systematic literature review. *Plos One*, *10*(10), e0139701. https://doi.org/10.1371/journal.pone.0139701

Clement, J. (2019). Twitter: number of monthly active users 2010-2019. *Statista.* https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users

Davies, R. (2020, April 3). Impose strict curbs on gambling during Covid-19 lockdown, MPs urge. *The Guardian*. https://www.theguardian.com/society/2020/apr/03/impose-strict-curbs-on-gambling-during-covid-19-lockdown-mps-urge

De Martino, I., D'Apolito, R., McLawhorn, A. S., Fehring, K. A., Sculco, P. K., & Gasparini, G. (2017). Social media for patients: Benefits and drawbacks. *Current Reviews in Musculoskeletal Medicine*, *10*(1), 141-145. https://doi.org/10.1007/s12178-017-9394-7

de Rosa, A.S., Fino, E. & Bocci, E. (2013). Addressing healthcare on-line demand and supply relating to mental illness: Knowledge sharing about psychiatry and psychoanalysis through SN in Italy and France. In A. Kapoor & C. Kulshrestha (Eds), *Dynamics of Competitive Advantage and Consumer Perception in Social Marketing* (pp. 16-55). IGI Global.

Dubey, M. J., Ghosh, R., Chatterjee, S., Biswas, P., Chatterjee, S., & Souvik, D. (2020). COVID-19 and addiction. *Diabetes and Metabolic Syndromes*, *14*(5), 817-823. https://doi.org/10.1016/j.dsx.2020.06.008

Dunford, D., Dale, B., Stylianou, N., Lowther, E., Ahmed, M., & de la Torre Arenas, I. (2020). Coronavirus: The world in lockdown in maps and charts. *BBC News.* Retrieved December 14, 2020, from: https://www.bbc.co.uk/news/world-52103747

Eckler, P., Worsowicz, G., & Rayburn, J. W. (2010). Social media and health care: An overview. *PM&R*, *2*(11), 1046-1050. https://doi.org/10.1016/j.pmrj.2010.09.005

Gainsbury, S. M. (2015). Online gambling addiction: The relationship between internet gambling and disordered gambling. *Current Addiction Reports*, *2*(2), 185-193. https://doi.org/10.1007/s40429-015-0057-8

Gainsbury, S. M., Delfabbro, P., King, D. L., & Hing, N. (2016). An exploratory study of gambling operators' use of social media and the latent messages conveyed. *Journal of Gambling Studies, 32*(1), 125-141. https://doi.org/10.1007/s10899-015-9525-2

Gainsbury, S. M., Hing, N., & Suhonen, N. (2014). Professional help-seeking for gambling problems: Awareness, barriers and motivators for treatment. *Journal of Gambling Studies*, *30*(2), 503-519. https://doi.org/10.1007/s10899-013-9373-x

Gambling Commission (2020). Customer interaction - Additional formal guidance for remote operators during COVID-19 outbreak. Gambling Commission. Retrieved December 14, 2020, from: http://www.gamblingcommission.gov.uk/news-action-and-statistics/Statistics-and-research/Covid-19-research/Customer-interaction-%E2%80%93-Additional-formal-guidance-for-remote-operators-during-COVID-19-outbreak.aspx

Griffiths, M. (2001). Internet gambling: Preliminary results of the first U.K. prevalence study. *Journal of Gambling Issues*, 5, Article 5. https://doi.org/10.4309/jgi.2001.5.8

Griffiths, M. D. (2005). A "components" model of addiction within a biopsychosocial framework. *Journal of Substance Use*, *10*, 191-197. https://doi.org/10.1080/14659890500114359

Griffiths, M. D. (2016). Problem gambling and gambling addiction are not the same. *Journal of Addiction and Dependence*, *2*(1), 1-3. https://doi.org/10.15436/2471-061X.16.014

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, *101*(1), 5228-5235. https://doi.org/10.1073/pnas.0307752101

Håkansson, A. (2020). Changes in gambling behavior during the COVID-19 pandemic—A web survey study in Sweden. *International Journal of Environmental Research and Public Health*, *17*(11), 4013. https://doi.org/10.3390/ijerph17114013

Himelboim, I., & Han, J. Y. (2014). Cancer talk on Twitter: Community structure and information sources in breast and prostate cancer social networks. *Journal of Health Communication*, *19*(2), 210-225. https://doi.org/10.1080/10810730.2013.811321

Hinjoy, S., Tsukayama, R., Chuxnum, T., Masunglong, W., Sidet, C., Kleeblumjeak, P., Onsai, N., & Iamsirithaworn, S. (2020). Self-assessment of the Thai Department of Disease Control's communication for international response to COVID-19 in the early phase. *International Journal of Infectious Diseases*, *96*, 205-210. https://doi.org/10.1016/j.ijid.2020.04.042

Houghton, S., McNeil, A., Hogg, M. & Moss, M. (2020). Comparing the Twitter posting of British gambling operators and gambling affiliates: A summative content analysis. *International Gambling Studies*, *36*, 1-23. https://doi.org/10.1080/14459 795.2018.15619 23

Hu, Y., Sun, J., Dai, Z., Deng, H., Li, X., Huang, Q., Wu, Y., Sun, L., & Xu, Y. (2020). Prevalence and severity of corona virus disease 2019 (COVID-19): A systematic review and meta-analysis. *Journal of Clinical Virology*, *127*, 104371. https://doi.org/10.1016/j.jcv.2020.104371

Ibarra-Vega, D. (2020). Lockdown, one, two, none, or smart. Modeling containing Covid-19 infection. A conceptual model. *Science of The Total Environment*, *730*, 138917. https://doi.org/10.1016/j.scitotenv.2020.138917

Iliceto, P., Fino, E., Schiavella, M., & Candilera, G. (2020). Individual differences in interpersonal security predict suicidal ideation and problem gambling. *Personality and Individual Differences*, *162*, 110031. https://doi.org/10.1016/j.paid.2020.110031

Jockers, M. L. (2015). *Syuzhet: Extract sentiment and plot arcs from text.* https://github.com/mjockers/syuzhet

Kar, S. K., Arafat, S. M. Y., Sharma, P., Dixit, A., Marthoenis, M., & Kabir, R. (2020). COVID-19 pandemic and addiction: Current problems and future concerns. *Asian Journal of Psychiatry*, *51*, 102064. https://doi.org/10.1016/j.ajp.2020.102064

Kearney, M. (2019). rtweet: Collecting and analyzing Twitter data. *Journal of Open Source Software*, *4*(42), 1829. https://doi.org/10.21105/joss.01829

Killick, E. A., & Griffiths, M. D. (2020). A content analysis of gambling operators' Twitter accounts at the start of the English Premier League football season. *Journal of Gambling Studies*, *36*(1), 319-341. https://doi.org/10.1007/s10899-019-09879-4

Király, O., Potenza, M. N., Stein, D. J., King, D. L., Hodgins, D. C., Saunders, J. B., Griffiths, M. D., Gjoneska, B., Billieux, J., Brand, M., Abbott, M. W., Chamberlain, S. R., Corazza, O., Burkauskas, J., Sales, C. M. D., Montag, C., Lochner, C., Grünblatt, E., Wegmann, E., … Demetrovics, Z. (2020). Preventing problematic internet use during the COVID-19 pandemic: Consensus guidance. *Comprehensive Psychiatry*, 152180. https://doi.org/10.1016/j.comppsych.2020.152180

Limaye, R. J., Sauer, M., Ali, J., Bernstein, J., Wahl, B., Barnhill, A., & Labrique, A. (2020). Building trust while influencing online COVID-19 content in the social media world. *The Lancet Digital Health*, *2*(6), e277-e278. https://doi.org/10.1016/S2589-7500(20)30084-4

Lopez-Gonzalez, H., Guerrero-Solé, F., Estévez A., & Griffiths, M. D. (2018). Betting is loving and bettors are predators: A conceptual metaphor approach to online sports betting advertising. *Journal of Gambling Studies*, *34*, 709-726. https://doi.org/10.1007/s10899-017-9727-x

Mackolil, J., & Mackolil, J. (2020). Addressing psychosocial problems associated with the COVID-19 lockdown. *Asian Journal of Psychiatry*, *51*, 102156. https://doi.org/10.1016/j.ajp.2020.102156

Mallet, J., Dubertret, C., & Le Strat, Y. (2021). Addictions in the COVID-19 era: Current evidence, future perspectives a comprehensive review. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *106*, 110070. https://doi.org/10.1016/j.pnpbp.2020.110070

Mariano, E. R., Kou, A., Stiegler, M. A., & Matava, C. (2021). The rise and fall of the COVID-19 aerosol box through the lens of Twitter. *Journal of Clinical Anesthesia*, *69*, 110145. https://doi.org/10.1016/j.jclinane.2020.110145

Mohammad, S., & Turney, P. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, *29*. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Ogundepo, E., Folorunso, S., Adekanmbi, O., Akinsande, O., Banjo, O., Ogbuju, E., Oladipo, F., Abimbola, O., Oseghale, E., & Babajide, O. (2020). An exploratory assessment of a multidimensional healthcare and economic data on COVID-19 in Nigeria. *Data in Brief*, *33*, 106424. https://doi.org/10.1016/j.dib.2020.106424

Price, A. (2020). Online gambling in the midst of COVID-19: A nexus of mental health concerns, substance use and financial stress. *International Journal of Mental Health and Addiction*. https://doi.org/10.1007/s11469-020-00366-1

Qiu, Z., Wu, B., Wang, B., Shi, C., & Yu, L. (2014). Collapsed Gibbs sampling for latent Dirichlet allocation on spark. *Proceedings of the 3rd International Conference on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications*, *36*, 17-28.

Ransing, R., Adiukwu, F., Pereira-Sanchez, V., Ramalho, R., Orsolini, L., Teixeira, A. L. S., Gonzalez-Diaz, J. M., Pinto da Costa, M., Soler-Vidal, J., Bytyçi, D. G., El Hayek, S., Larnaout, A., Shalbafan, M., Syarif, Z., Nofal, M., & Kundadak, G. K. (2020). Mental health interventions during the COVID-19 pandemic: A conceptual framework by early career psychiatrists. *Asian Journal of Psychiatry*, *51*, 102085. https://doi.org/10.1016/j.ajp.2020.102085

Reinert, A. (1983). Une méthode de classification descendante hiérarchique: Application à l'analyse lexicale par contexte [A hierarchical descendent classification method: Application to lexical analysis by context]. *Cahiers de l'Analyse Des Données*, *8*(2), 187-198.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(1), 1-40. https://doi.org/10.18637/jss.v091.i02

Rossi, R., Socci, V., Talevi, D., Mensi, S., Niolu, C., Pacitti, F., Marco, A. D., Rossi, A., Siracusano, A., & Lorenzo, G. D. (2020). COVID-19 pandemic and lockdown measures impact on mental health among the general population in Italy. An N=18147 web-based survey. *MedRxiv*, 2020.04.09.20057802. https://doi.org/10.1101/2020.04.09.20057802

Sansone, A., Cignarelli, A., Ciocca, G., Pozza, C., Giorgino, F., Romanelli, F., & Jannini, E. A. (2019). The sentiment analysis of tweets as a new tool to measure public perception of male erectile and ejaculatory dysfunctions. *Sexual Medicine*, *7*(4), 464-471. https://doi.org/10.1016/j.esxm.2019.07.001

Satre, D. D., Hirschtritt, M. E., Silverberg, M. J., & Sterling, S. A. (2020). Addressing problems with alcohol and other substances among older adults during the COVID-19 pandemic. *The American Journal of Geriatric Psychiatry*, *28*(7), 780–783. https://doi.org/10.1016/j.jagp.2020.04.012

Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach* (1st ed.). O'Reilly Media.

Silge, J., Robinson, D., & Hester, J. (2016). *Tidytext: Text mining using Dplyr, Ggplot2, and other Tidy tools*. Retrieved December 14, 2020, from: https://doi.org/10.5281/zenodo.56714

Sinnenberg, L., Buttenheim, A. M., Padrez, K., Mancheno, C., Ungar, L., & Merchant, R. M. (2016). Twitter as a tool for health research: A systematic review. *American Journal of Public Health*, *107*(1), e1-e8. https://doi.org/10.2105/AJPH.2016.303512

Sirola, A., Kaakinen, M., Savolainen, I., & Oksanen, A. (2019). Loneliness and online gambling-community participation of young social media users. *Computers in Human Behavior*, *95*, 136-145. https://doi.org/10.1016/j.chb.2019.01.023

Stellefson, M., Chaney, B., Ochipa, K., Chaney, D., Haider, Z., Hanik, B., Chavarria, E., & Bernhardt, J. M. (2014). YouTube as a source of COPD patient education: A social media content analysis. *Chronic Respiratory Disease*, *11*(2), 61-71. https://doi.org/10.1177/1479972314525058

Sugawara, Y., Narimatsu, H., Hozawa, A., Shao, L., Otani, K., & Fukao, A. (2012). Cancer patients on Twitter: A novel patient community on social media. *BMC Research Notes*, *5*(1), 699. https://doi.org/10.1186/1756-0500-5-699

Sun, L., & Yin, Y. (2017). Discovering themes and trends in transportation research using topic modeling. *Transportation Research Part C: Emerging Technologies*, *77*, 49-66. https://doi.org/10.1016/j.trc.2017.01.013

Symeonidis, S., Effrosynidis, D., & Arampatzis, A. (2018). A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analysis. *Expert Systems with Applications*, *110*, 298-310. https://doi.org/10.1016/j.eswa.2018.06.022

Trovato, C. M., Montuori, M., Oliva, S., Cucchiara, S., Cignarelli, A., & Sansone, A. (2020). Assessment of public perceptions and concerns of celiac disease: A Twitter-based sentiment analysis study. *Digestive and Liver Disease*, *52*(4), 464–466. https://doi.org/10.1016/j.dld.2020.02.004

Twitter (2021). *Twitter Developer - API reference index.* Retrieved February 15, 2021, from: https://developer.twitter.com/en/docs/api-reference-index

Vallurupalli, V., & Bose, I. (2020). Exploring thematic composition of online reviews: A topic modeling approach. *Electronic Markets*, *30*(4), 791–804. https://doi.org/10.1007/s12525-020-00397-5

Vasudevan, M., Mehrolia, S., Alagarsamy, S., & Balachandran, A. K. (2021). Work from home in the pandemic era: Loss of mental equilibrium? *Asian Journal of Psychiatry*, *55*, 102490. https://doi.org/10.1016/j.ajp.2020.102490

Wardle, H., Moody, A., Griffiths, M. D., Orford, J., & Volberg, R. (2011). Defining the online gambler and patterns of behavior integration: Evidence from the British Gambling Prevalence Survey 2010. *International Gambling Studies*, *11*(3), 339-356. https://doi.org/ 10.1080/14459795.2011.628684

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wijffels, J. (2020). BTM: *Biterm Topic Models for Short Text. R package version 0.3.* Retrieved December 14, 2020, from: https://CRAN.R-project.org/package=BTM

World Health Organization (2020). *Novel coronavirus*. Retrieved December 14, 2020, from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Yada, A., Jha, C. K., Sharan, A., & Vaish, V. (2020). Sentiment analysis of financial news using unsupervised approach. *Procedia Computer Science*, *167,* 589-598. https://doi.org/10.1016/j.procs.2020.03.325

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. *Proceedings of the 22nd international conference on World Wide Web (WWW '13)*, 1445-1456. https://doi.org/10.1145/2488388.2488514

Zhang, L., Hall, M., & Bastola, D. (2018). Utilizing Twitter data for analysis of chemotherapy. *International Journal of Medical Informatics*, *120*, 92-100. https://doi.org/10.1016/j.ijmedinf.2018.10.002