COGITATIO

Article

A Computational Approach to Analyzing the Twitter Debate on Gaming Disorder

Tim Schatto-Eckrodt *, Robin Janzik, Felix Reer, Svenja Boberg and Thorsten Quandt

University of Münster, Department of Communication, 48143 Münster, Germany;
E-Mails: tim.schatto-eckrodt@uni-muenster.de (T.S.-E.), robin.janzik@uni-muenster.de (R.J.),
felix.reer@uni-muenster.de (F.R.), svenja.boberg@uni-muenster.de (S.B.), thorsten.quandt@uni-muenster.de (T.Q.)

* Corresponding author

Abstract

The recognition of excessive forms of media entertainment use (such as uncontrolled video gaming or the use of social networking sites) as a disorder is a topic widely discussed among scientists and therapists, but also among politicians, journalists, users, and the industry. In 2018, when the World Health Organization (WHO) decided to include the addictive use of digital games (gaming disorder) as a diagnosis in the International Classification of Diseases, the debate reached a new peak. In the current article, we aim to provide insights into the public debate on gaming disorder by examining data from Twitter for 11 months prior to and 8 months after the WHO decision, analyzing the (change in) topics, actors, and sentiment over time. Automated content analysis revealed that the debate is organic and not driven by spam accounts or other overly active 'power users.' The WHO announcement had a major impact on the debate, moving it away from the topics of parenting and child welfare, largely by activating actors from gaming culture. The WHO decision also resulted in a major backlash, increasing negative sentiments within the debate.

1. Introduction

When television viewing became a mass phenomenon in the 1950s, it only took a few years until the first scientific works on television addiction were published (e.g., Meerloo, 1954). Today, discussions about excessive, pathological behavior primarily concern digital forms of media entertainment, such as social networking sites and/or video games. The debate on the latter reached a new peak in 2018 when the World Health Organization (WHO) decided to include the addictive use of digital games as a diagnosis in the 11th Revision of the International Classification of Diseases (ICD-11). Gaming disorder is defined as:

A pattern of persistent or recurrent gaming behaviour…manifested by: 1) impaired control over gaming (e.g., onset, frequency, intensity, duration, termination, context); 2) increasing priority given to gaming to the extent that gaming takes precedence over other life interests and daily activities; and 3) continuation or escalation of gaming despite the occurrence of negative consequences. The behaviour pattern is of sufficient severity to result in significant impairment in personal, family, social, educational, occupational or other important areas of functioning. (WHO, 2019)

Some scholars support the idea of gaming disorder being recognized in official manuals. They argue that this

is a prerequisite to establish adequate treatment for the improvement of public health. From a societal perspective, a possible pathologization of players is of lower priority to this goal (e.g., Rumpf et al., 2018). Others, especially from the social sciences and communication science, question the scientific basis of this decision and warn against moral panics (e.g., van Rooij et al., 2018). Through a certain media presentation, gaming might be characterized as a threat, thus pathologizing normal behavior and putting scientific results in a different light (e.g., Bowman, 2016; Markey & Ferguson, 2017).

This is reflected in traditional mass media such as newspapers and television. Reports contain numerous portrayals of extreme cases: A mother starved her daughter to death because of her gaming habits (Thompson, 2011), a gamer died from thrombosis because of playing a game for 22 uninterrupted days (McCrum, 2015), and a desperate father tried to deter his son from playing by hiring other gamers to target his avatar repeatedly (Kleinman, 2013). In addition to traditional media and scientific outlets, the debate is increasingly taking place on social networking sites. A distinctive feature of these sites is that a diverse group of stakeholders are active participants in the debate; gamers, in particular, take part in it.

While the discussion of gaming disorder within academia can be understood on the basis of the many debate articles published in scientific journals, such as the *Journal of Behavioral Addictions* (e.g., Aarseth et al., 2017; Billieux et al., 2017; Griffiths, Kuss, Lopez-Fernandez, & Pontes, 2017; Rumpf et al., 2018; van den Brink, 2017; van Rooij et al., 2018), the public debate is more fragmented and less tangible. Thousands of social media posts, blog articles, and videos on the topic have been published by a very diverse set of actors. This large amount of data makes the use of traditional methods of content analysis almost impossible and requires innovative methods suitable for the analysis of large-scale datasets. In recent years, a new discipline known as 'computational social science' emerged, developing and refining the tools necessary for these kinds of large-scale analyses (Conte et al., 2012). Fields like political science (e.g., Hopkins & King, 2010), communication science (e.g., van Atteveldt & Peng, 2018), and subfields like journalism studies (e.g., Boumans & Trilling, 2016) have started to use computational methods in their research, but they have scarcely been utilized in the field of media entertainment research.

By applying automated content analysis to a Twitter dataset, the current study, on the one hand, provides insights into the public debate on gaming disorder, and on the other hand, showcases the usefulness of computational approaches in media entertainment research.

2. The Gaming Disorder Debate in Traditional Media, Science, and Beyond

Systematic analyses of the debate on gaming disorder in media coverage, scientific journals, or on other plat-forms, such as social networking sites, are rare, and their findings are fragmented. Prior studies predominantly looked at print media and identified addiction as one aspect in the general debate on gaming. Kirkpatrick (2016) examined the gaming discourse in American magazines in the 1980s and found that most articles considered games to be unsuitable for children, but did not differentiate between the addictive potential of games specifically and technology in general. In a review of Chinese historical and media sources, Szablewicz (2010) found that Internet addiction and Internet gaming are often portrayed in a sensationalistic way, suggesting the framing of the debate as a moral panic. Whitton and Maclure (2017) showed that the video game discourse in British print media is dominated by the narrative of naïve video game players becoming addicted because they cannot control the technology. In one of the few quantitative empirical studies, Jung (2019) investigated the Korean media landscape with regard to its stance on gaming regulations. By analyzing daily newspapers, digital news sites, and digital gaming magazines, he identified several frames in the debate, such as child protection, the preservation of the social order, freedom of choice, cultural consequences, and the effectiveness of therapies. Furthermore, Jung (2019) found that conservative, moderate, and specific IT news outlets differed in the extent to which they addressed these topics. While most conservative media emphasized the negative effects of gaming, IT news outlets only exhibited a positive or neutral stance; moderate media were more balanced, yet trending toward a negative opinion.

Taken together, previous works merely focused on traditional media in a national context, ignoring changes over time. However, with rise of social media platforms in the early 21st century, the media landscape as a whole has drastically changed—and with it, the dynamics of public discourse. Social media platforms have enabled almost everyone to not only observe but actively participate in ongoing debates, reaching large audiences that were previously only accessible via traditional mass media. With this change in the dynamics of the public sphere, many hopes were raised about the democratizing potential of these new platforms (Halpern & Gibbs, 2013; Levina & Arriaga, 2014). The ideal of a discursive public sphere—for instance, in the sense of the philosopher Jürgen Habermas (1991)—in which citizens and the political elite find the best solution to social problems together and at eye level, suddenly seemed to be within reach. The low entry barriers also enable actors from civil society to gain access to public discourse and reach a mass audience.

This ideal of openness and inclusivity is especially salient on Twitter, which attracts not only a sizable share of regular media users (22% of American adults use Twitter; Wojcik & Hughes, 2019), but also policy makers, celebrities, activists, and journalists of traditional and new media organizations (Groshek & Tandoc, 2017; Paulussen & Harder, 2014). The dynamics of public dis-

course on Twitter are shaped not only by the aforementioned openness of the platform and the diversity of its users' backgrounds, but also by two characteristic, technical affordances of the platform.

First, new accounts on Twitter are set to be public by default, that is, other users can subscribe or follow the account without asking permission. This follow relationship is asymmetrical, as the number of accounts a user follows can and often does differ from the number of accounts the user is followed by. This technical characteristic enables two types of network structures: the 'one-to-many' structure, where, similar to traditional mass media, single actors reach a large audience and the 'many-to-many' structure, where groups of users communicate among themselves. However, the potential reach of a user is not just defined by their follow network. The mechanic of the retweet lets users share a tweet of someone they follow with their own followers, thus making the barriers of the follow networks even more pervious. As a result of these network structures, even accounts with a low number of followers can potentially reach a large audience.

Second, a feature of almost all social media platforms—the hashtag—was first used on Twitter and is a central affordance of the platform. The hashtag makes it easier for Twitter users to find relevant tweets and to make their own posts easier for other users to discover. This feature also made it possible to quickly find and participate in ongoing debates, connecting different users with each other and potentially raising awareness about trending topics. In recent years, hashtags have also gained relevance in political activism. Campaigns and movements using a hashtag both as branding and a communication tool played a significant role in the context of many topics, such as feminism (e.g., #meToo, #whyIStayed), anti-racism (e.g., #blackLivesMatter, #takeAKnee), and other political movements (e.g., #ArabSpring, #UmbrellaRevolution). Hashtags have also been used by malicious actors to insert themselves into a discourse with the intent to disrupt the ongoing debate or to push their own political views, a practice called 'hashtag hijacking' (Hadgu, Garimella, & Weber, 2013; VanDam & Tan, 2016). A notable example of both hashtag hijacking and a movement using a hashtag to organize was #gamerGate, which was "spawned by individuals who purported to be frustrated by a perceived lack of ethics within gaming journalism" (Massanari, 2017, p. 330). Partly by outside agitation and hashtag hijacking by right-wing groups on 4chan, #gamerGate "became a campaign of systematic harassment of female and minority game developers, journalists, and critics and their allies" (p. 330). As evidenced by #gamerGate, Twitter, as a platform, has a history of video game-related activism.

An analysis of the debate on Twitter is particularly interesting, as the technical affordances of the platform enable dynamics of public discourse vastly different from those of the traditional media landscape. Following the theoretical considerations on Twitter as a platform for public discourse, our research questions focus on the actors, topics, and tone present in the debate, as well as potential changes in these categories arising from the decision by the WHO to include the addictive use of digital games as a diagnosis in the ICD-11. To our knowledge, prior research has not examined the debate on this level.

In order to investigate the claim of a more diverse debate based on the general heterogeneity of users on social networking sites, our first interest centered on the participating actors. We were not only interested in the opinions and background of the actors themselves, but also in determining whether their motivations were genuine. We wanted to know if they were actually interested in the debate or whether they were trying to disrupt the discourse in an orchestrated fashion (i.e., trolling). In traditional media, it is mainly scientists, politicians, and experts—or people who are presented as such— who have a say. By allowing anyone to post for the general public, social media involves a more heterogeneous group in the debate: gamers, gaming communities, and those affected by negative consequences may also contribute. Therefore, our first research question was:

RQ1: Which actors participate in the debate?

Given the range of issues discovered in prior research on traditional media, our second question asks if this also holds true for social media. One could assume that topics are being discussed that are not included in the scholarly debate and the reporting of traditional media. Therefore, we asked:

RQ2: What topics are being discussed?

Considering the two factions on gaming disorders in academia and differences in traditional media reports based on their background, we were interested in the sentiments expressed by the actors. Thus, we asked:

RQ3: How is the tone of the overall debate?

The dataset available to us also offered the interesting option of looking at the development of the debate over time. We wanted to know whether the WHO decision had an impact on the debate, and how, if at all, the answers to the previous research questions differed before and after the WHO decision. Our last research question was:

RQ4: How did the WHO decision influence the debate?

3. Data and Methods

3.1. Data

To answer these questions, a large-scale, automated content analysis of $N = 16,831$ tweets, of which 55.11%

were retweets, posted between March 16th, 2017 and November 15th, 2018 was conducted. The dataset was extracted from Twitter's Decahose stream, which represents a 10% sample of all public tweets and filtered for tweets mentioning the discourse on gaming disorder (for an extended description of the filtering process, see Supplementary File C).

At its peak on June 18th, 2018, a total of 3,308 tweets and retweets were posted, representing roughly 0.01% of the overall volume of all tweets posted during that day. The debate was clearly stimulated by both the release of the beta draft and the official version of the ICD-11, with the corresponding peaks labeled in Figure 1. Within

those two peaks, 63.99% of the overall tweet volume was posted.

As we used archival data for our analysis, we were able to analyze tweets that were deleted between their original publishing and the time of analysis. Table 1 shows the proportion of tweets that are still online and the share of tweets that were no longer available online. There are three reasons why a tweet might be offline: 1) the tweet or the account was deleted by the user (deleted), 2) the user set their account to private (protected), or 3) the user was suspended by Twitter (suspended). Compared to a random sample of tweets with a similar age, this share of 73.1% online tweets is relatively
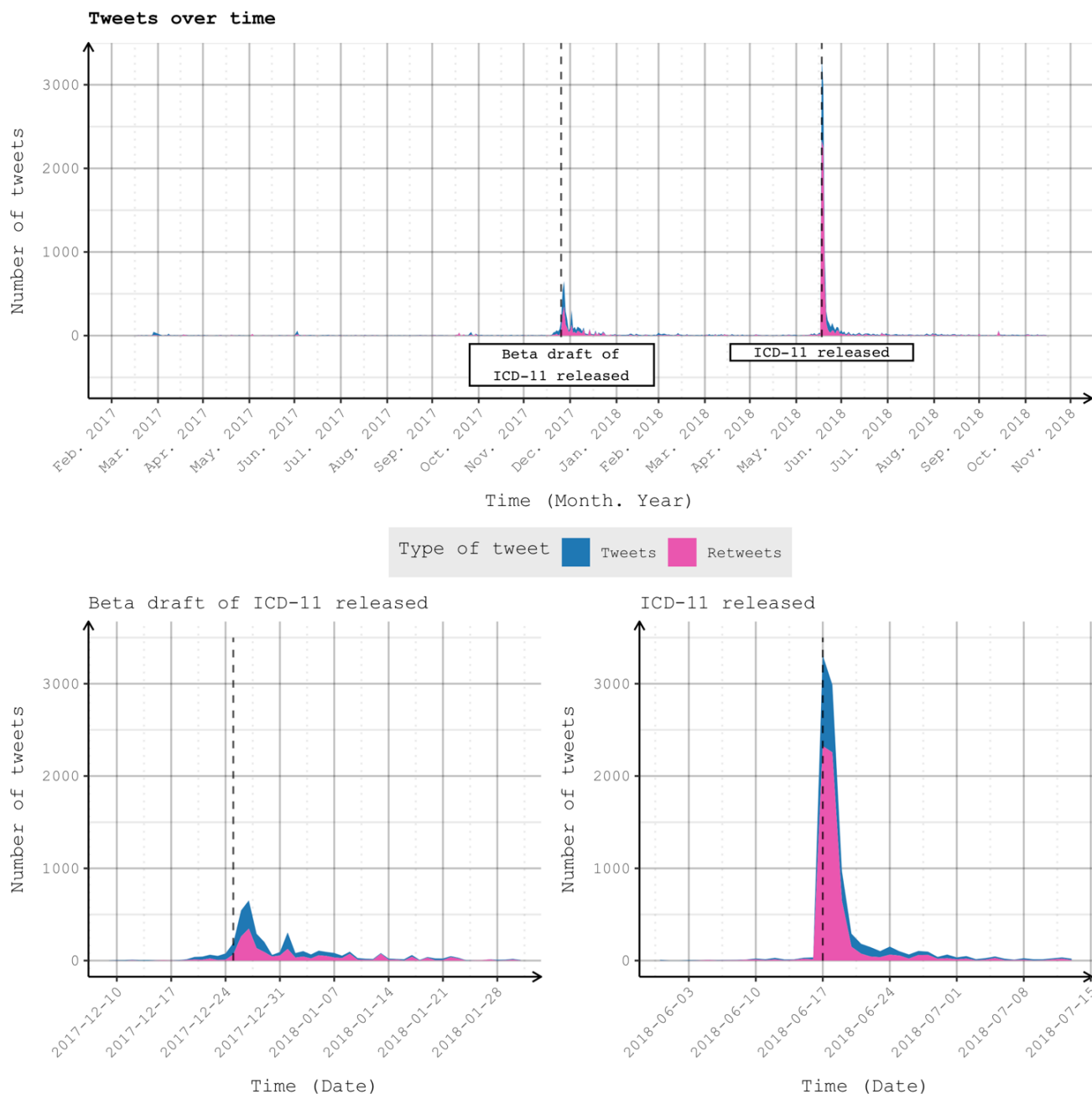
Figure 1. Number of tweets over time. Note: The vertical lines represent the release date of the ICD-11 beta draft (December 26th, 2017) and the release of ICD-11 (June 18th, 2020), respectively.

Table 1. Share of offline tweets.

Online status	%
Online	73.12
Deleted	12.24
Protected	1.30
Suspended	13.34

Notes: N = 16,831, including retweets. Checked in December 2019, an average 626 days after initial publishing.

high. This can be seen as evidence for an organic, not bot- and/or spam-driven discourse, since Twitter usually deletes spam accounts soon after their participation in a trending topic (Thomas, Grier, Song, & Paxson, 2011).

As we were interested in the effect the WHO decision had on the discourse, we split the dataset into three segments: Segment 1, comprised of tweets posted before the release of the ICD-11 beta draft, Segment 2, comprised of tweets posted after the release of the ICD-11 beta draft and before the official release of the ICD-11, and Segment 3, comprised of tweets posted after the official release of the ICD-11 (see Table 2).

3.2. Methods

The tweets' contents were analyzed using a combination of structural topic modeling, sentiment analysis, and an analysis of used hashtags and present actors. The following chapter gives a detailed description of these methods in order to provide other researchers with the tools to conduct similar analyses and to build upon this framework. All analyses were conducted with R (for a full list of used packages and versions, see Supplementary File A).

3.2.1. Preprocessing

A necessary prerequisite for all kinds of automated and semi-automated content analysis is the procedure of preprocessing. This procedure includes important and impactful decisions by the researchers and is often not well documented or dealt with in a non-transparent way (Denny & Spirling, 2018; Maier et al., 2018). In the current study, we pre-processed the documents by removing non-word characters and tokenizing the documents, by removing stopwords, by stemming, and by pruning. The R code for the preprocessing pipeline used in this study can be found in the Open Science Framework (Schatto-Eckrodt, Janzik, Reer, Boberg, & Quandt, 2020) and a detailed description of the preprocessing steps can be found in the Supplementary File D.

3.2.2. Topic Modeling

Topic modeling "is a computational content-analysis technique that can be used to investigate the 'hidden' thematic structure of a given collection of texts" (Maier et al., 2018, p. 1). In the context of topic modeling, the collection of texts to be analyzed is called a 'corpus,' while each text within the corpus is called a 'document.' In this case, the corpus consisted of every tweet, excluding retweets within the dataset, while a single tweet was a document. Retweets were excluded because including them could potentially introduce a bias towards topics that were represented in often-retweeted tweets. The structural topic model (STM) introduced by Roberts, Stewart, and Tingley (2019) is an extension of other probabilistic topic models, such as the latent dirichlet allocation (Blei, Ng, & Jordan, 2003), which enables researchers to "incorporate arbitrary metadata, defined as information about each document, into the topic model" (Roberts et al., 2019, p. 2). The topics modelled by STM and other topic modeling techniques represent latent content variables that should form a comprehensive representation of the corpus. Like most topic modeling techniques, STM tries to infer these topics from recurring patterns of word occurrence in documents, while ignoring the order of words within each document (i.e., using the bag-of-words assumption; Maier et al., 2018).

STM requires the researcher to choose a number of topics before applying the model. As there is no correct answer to the question of what this number should be (Grimmer & Stewart, 2013), we applied the elbow method on the measures for semantic coherence and held-out likelihood and found a five-topics solution to be optimal. As mentioned before, STM (Roberts et al., 2019) enables researchers to add covariates for topical prevalence to the topic model, which allows the observed metadata to affect the frequency with which a topic is discussed. In this analysis, we modeled the topical prevalence as a function of the time segment matching the publishing time of each tweet, as described above.

Table 2. Time segments.

Segment	Time frame	Number of tweets	%
1	March 16th, 2017–November 30th, 2017	810	10.72
2	December 1st, 2017–June 14th, 2018	2,886	38.20
3	June 15th, 2018–November 15th, 2018	3,859	51.08

Notes: n = 7,555, excluding retweets.

Models using the raw publishing timestamp of the tweet as a covariate and no covariates at all performed slightly worse than the final model.

3.2.3. Sentiment Analysis

The sentiment analysis was conducted using the opinion lexicon by Hu and Liu (2004), which features a list of English positive (2,001 words) and negative (4,779 words) opinion or sentiment words. This opinion lexicon is widely used for the analysis of social media data (Si et al., 2013; Zhang, Ghosh, Dekhil, Hsu, & Liu, 2011) and is available online, thus enabling replication. In order to provide a robust measure of sentiment, we created a corpus of 88 documents, with each document representing a full week's worth of tweets, including retweets. Pooling the tweet texts and moving the unit of analysis from single tweets to weeks within the debate, enabled us to show the change in the overall sentiment of the debate. For each week, we calculated the share of negative, positive, and neutral sentiment words. The same preprocessing steps were taken for the sentiment analysis as were for the topic modeling, with the exception of the removal of emojis, which were replaced with their Unicode Common Locale Data Repository short names, via the sentimentr package, as including emoticons significantly improves the accuracy of sentiment classification (Hogenboom et al., 2013). As expected, most words (83.29%) fall neither into the positive nor the negative sentiment category. Overall, 14% of all words were identified as negative, and only 2.44% as positive.

3.2.4. Co-Occurrence Graphs

To analyze the hashtags in our dataset, in addition to simple frequency tables, we calculated co-occurrence graphs of the used hashtags. In short, co-occurrence analysis is the analysis of the pairwise connection between elements in a set, with the connection modelled as the occurrence of two elements in a subset of all elements. In content analyses, the elements are often words, the subsets are documents, and the set of all elements is the corpus. Co-occurrence analysis can be used to construct networks (i.e., graphs) representing the connection between words, revealing thematic clusters (Buzydlowski, 2015). As a general method in content analysis, co-occurrence was used even before the introduction of computational methods (Harris, 1957) and has been used specifically for the analysis of social media data in numerous studies (e.g., Aiello et al., 2013; Pervin, Phan, Datta, Takeda, & Toriumi, 2015; Wang, Wei, Liu, Zhou, & Zhang, 2011). In the current study, we used co-occurrence graphs to gain insights into the use of hashtags within the debate. We extracted the hashtags from all tweets, excluding retweets, and built graphs with hashtags as vertices and the co-occurrence of two hashtags in the same tweet as edges. Edges were weighted by the number of co-occurrences.

4. Results

4.1. Actors

Of the 15,498 unique users that are represented in our dataset, 2.39% were verified by Twitter. According to Twitter (2020), "an account may be verified if it is determined to be an account of public interest. This includes accounts maintained by users in music, acting, fashion, government, politics, religion, journalism, media, sports, [and] business." This relatively high number (compared to less than 1% verified users in a random sample) suggests a high involvement of journalistic actors and actors who are otherwise involved in public life (Paul, Khattar, Kumaraguru, Gupta, & Chopra, 2019).

When analyzing the most important actors in a social media discourse, one must consider two characteristics: the reach of a user and the volume of their participation in the discourse. The reach of a user is defined by both their follower count and the number of times they were retweeted in the context of the debate. The higher the reach of a user, the higher the number of users getting into contact with their posts. In contrast, users who participate to a higher extent than the average user might have a lower reach than other users; yet, they are still responsible for a large share of the posts in the debate. These two characteristics are a consequence of the technical affordances of Twitter as a platform. It is possible that a user only mentions the discourse's topic in a single tweet, but—taking retweets into account—is still seen by most people participating in the discourse, thus being overrepresented in most users' timelines.

Those users of the second category, that is, users who participate more frequently than the average user, are shown in Table 3. The top-five most active users are supportive of the WHO decision and try to warn others of the dangers of gaming disorder; they have a parenting or professional education background. Users who oppose the WHO decision are also present in this group, and most of them have a background in gaming culture or technology journalism.

All of the users with the highest reach, except the CNN account, oppose the WHO decision and have a background in gaming culture (see Table 4). These accounts only began participating after the official release of the ICD-11, while the users with the larger extent of participation were part of the discourse months before the release of the beta draft on ICD-11.

To investigate whether a group of users was overrepresented in the dataset, we calculated the distribution of the tweet volume per user share. An equal distribution, that is, a distribution with the share of tweets equal to the share of users, would mean that there are no overly active 'power users.' Looking at all participating users, this kind of equal distribution can be observed. Again, this can be seen as evidence of organic discourse. However, the distribution for retweeted accounts was more skewed. In total, 10% of all tweets that

Table 3. Top users according to their extent of participation.

Username	Tweets by the user	% of all tweets	Verified	View on the WHO decision	Number of followers	Notes
MommyNooz *	32	0.19	No	Supporting	576	Parenting blog
camerondare	31	0.18	Yes	Supporting	2,545	Activist
AdvocateforEd	25	0.15	No	Supporting	26,261	Blog
Lynch39083	22	0.13	No	Supporting	45,143	Scholar/activist
techedvocate	22	0.13	No	Supporting	20,398	Tech blog
eplayuk	16	0.10	No	Opposing	389	Gaming blog
gamescosplay	14	0.08	No	Opposing	664	Gaming blog
gamingthemind	14	0.08	No	Opposing	779	NGO/activists
Pairsonnalites	13	0.08	No	Opposing	4,647	NGO
HealthyWrld*	12	0.07	No	Neutral	21,936	Health blog

Notes: *N* = 16,831. * = account suspended.

were potentially seen by other users, including retweets, were authored by six users.

4.2. Content

4.2.1. Topics

Analyzing the hashtags revealed that a large share of the discourse was neither centered around #gamingdisorder nor #gamingaddiction. Only 9.84% of all tweets included at least one hashtag. The strategy of including the terms (not hashtags) 'gaming disorder' and 'gaming addiction' for sampling the data was thus an adequate choice. Table 5 shows the 15 most frequently used hashtags. Besides the two topical hashtags used as the sample query (#gamingdisorder and #gamingaddiction) and their variations (#gaming, #addiction, #videogames), we also found hashtags referencing the WHO (#icd11, #who) hashtags related to parenting (#children, #parenting) and hashtags most likely used in journalistic reporting (#bbcbreakfast, #tech, #news).

Comparing the co-occurrence graphs of the hashtags used in Segment 1 and Segment 3 illustrates how the debate changed after the release of the ICD-11 (see

Figures 2 and 3). The debate in Segment 1 consisted of two topical groups (education and a broad discussion of gaming disorder). This distinction fades in Segment 3, as the focus shifts away from the educational debate towards a more general discussion.

This shift was also noticeable when applying topic modeling to the data (see Table 6). Topic 3, which represents the topical group of tweets discussing educational and parenting-related arguments, is overshadowed by Topics 1, 2, and 4, which arise in Segment 3. Topic 5, where the classification of gaming addiction is compared to other mental conditions as gender dysmorphia, is represented almost only in Segment 3.

4.2.2. Sentiments

A sentiment analysis revealed that the topic was generally discussed with a relatively negative sentiment (see Figure 4). The tweets in our dataset were, in comparison to a random sample of English language tweets from the same time span, significantly more negative ($t(90) = 8.00$, $p < 0.001$, $d = 1.01$). Both the release of the beta draft and the official release of the ICD-11 resulted in a slight peak of negative sentiment.

Table 4. Top users according to their reach.

Username	Times the user was retweeted	% of all tweets	Verified	View on the WHO decision	Number of followers	Notes
CNN	403	2.39	Yes	Neutral	39,159,370	Media
deadmau5	385	2.29	Yes	Opposing	3,939,972	Musician
GaijinGoombah	318	1.89	No	Opposing	65,049	YouTube CC
BrendoTGB	274	1.63	No	Opposing	349	Regular user
LEGIQN	245	1.46	Yes	Opposing	306,185	Twitch CC
Pamaj	236	1.40	Yes	Opposing	1,174,043	E-sports athlete
NoahJ456	233	1.38	Yes	Opposing	949,370	YouTube CC
TheSmithPlays	227	1.35	No	Opposing	209,071	YouTube CC
Boogie2988	182	1.08	Yes	Opposing	685,290	YouTube CC
CaptainSparklez	170	1.01	Yes	Opposing	4,934,618	YouTube CC

Notes: *N* = 16,831. CC = content creator.

Table 5. Top hashtags.

Hashtag	Number of occurrences	% of hashtag occurrences
gaming	269	8.92
gamingdisorder	123	4.08
datascience	121	4.01
addiction	113	3.75
health	102	3.38
icd11	97	3.22
mentalhealth	97	3.22
bbcbreakfast	70	2.32
videogames	61	2.02
parenting	59	1.96
who	47	1.56
gamingaddiction	46	1.52
children	40	1.33
news	36	1.19
tech	32	1.06

Note: $n = 3{,}017$ hashtag occurrences.

Comparing the sentiment of Segment 1 against the combined sentiment in Segments 2 and 3 showed a significantly more negative sentiment in the latter ($t(66) = 4.87$, $p < 0.001$, $d = 1.09$).

In addition to this quantitative difference in sentiment, we also conducted a term frequency-inverse document frequency analysis, which is a statistical measure to determine the relative importance of a word within a document in a larger corpus, that is, words that are not only frequently used but are also used more frequently in a specific set of documents, as compared to others. Calculating the term frequency-inverse document frequency values for the tweets in our dataset and the random sample of English language tweets reveals that the terms 'irresponsible,' 'ridiculous,' and 'condemn' were the most relevant negative sentiment words associated with the topic.
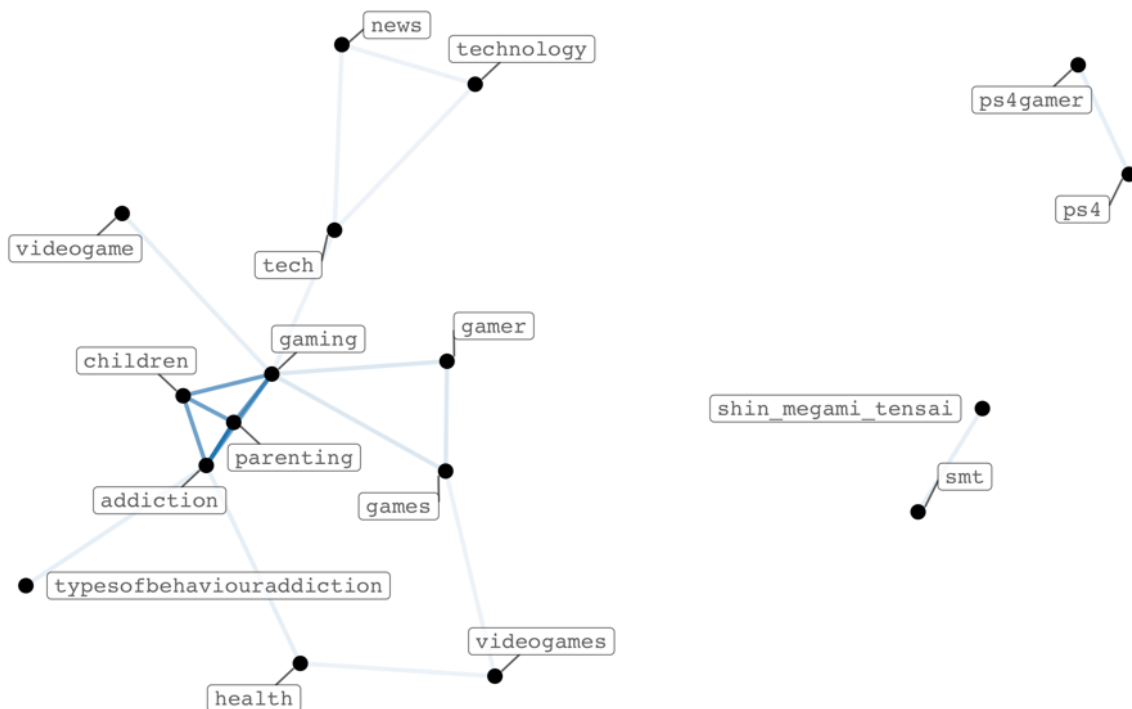
Figure 2. Co-occurrence graph for the top 30 hashtags in segment 1, force-directed layout algorithm by Fruchterman and Reingold (1991).
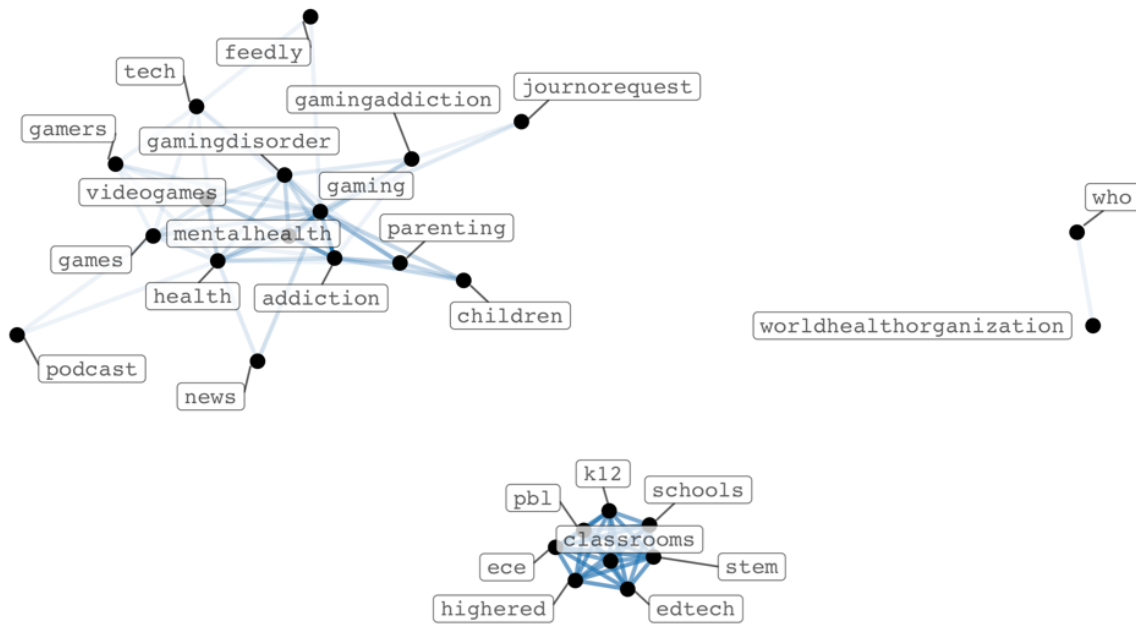
Figure 3. Co-occurrence graph for the top 30 hashtags in segment 3, force-directed layout algorithm by Fruchterman and Reingold (1991).

Table 6. Description, top terms, and representative quote of the topics.

Topic	Description	Terms	Quote
1	Breaking news: Sharing news articles on the WHO decision	Organization, recognizing, officially, mental, first time	"Gaming Disorders Officially Recognized by the World Health Organization: The World Health Organization (WHO) is recognizing 'gaming disorder' as a mental health issue in the beta draft of its upcoming 11th International Classification of Diseases"
2	Opposition: Voicing concerns on the validity of the WHO decision	Video, people, first time, mental, addiction	"'Gaming disorder' is total BS. People, especially adults, see gaming as such a taboo. It's ridiculous. Video games provide experiences that is as humans can't do in real life. Video games can connect people, video games provide a community and a sense of belonging."
3	Education: Information directed at educational professionals and parents regarding gaming addiction	Education, teachers, classroom, help, parents	"Learning Points: Parents need to wake up to gaming addiction"
4	Jokes: Mocking the WHO decision	Addict, play, video, time, mental	"PUBG is the cure to Gaming Disorder. 6 hours on that fucking thing will make you rage so hard you'll probably want to get outside for a bit."
5	Other conditions: Voicing concerns regarding the treatment of other mental afflictions in light of the WHO decision	Classification, time, video, recognizing, addiction	"So Gaming disorder is a mental health condition but gender dismorphia is not this is why government is garbage"

Note: STM of 5,378 documents, $K = 5$, time segment set as prevalence.

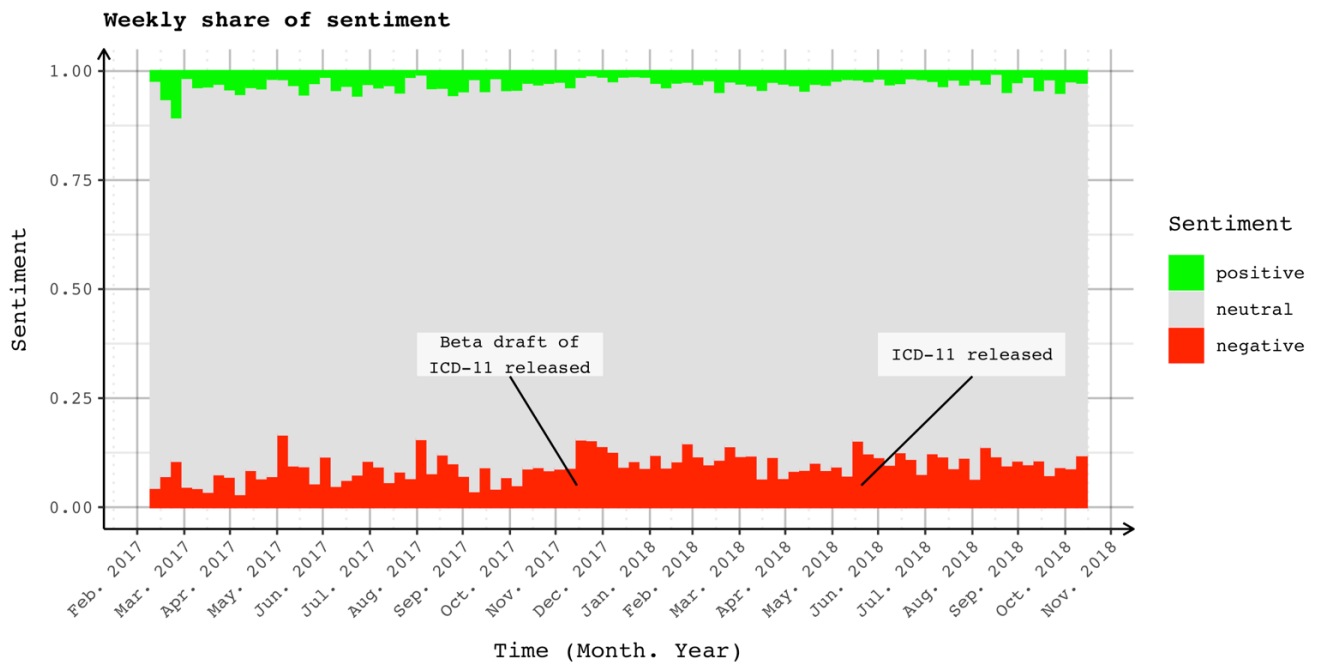**Weekly share of sentiment**

Figure 4. Share of positive and negative sentiment words per week over time.

4.2.3. Linked Websites

Users shared a total of 3,020 unique URLs with 33.92% of all (non-retweet) tweets containing at least one URL. Of these 3,020 URLs, 83.68% were shared only once. The five most often shared URLs link to Twitter's event page on the WHO decision to classify gaming disorder as mental health condition (294 shares), journalistic articles on the topic by the New York based media company Futurism (151 shares), and CNN Health (80 shares), a blog post critical of the WHO decision by the digital entertainment company Saljack Enterprises (66 shares), and a video by YouTube content creator Philip DeFranco, explaining why the WHO decision might villainize games (59 shares).

Most URLs shared by users belong to large online media outlets (e.g., CNN, ABC News) and contain factual reporting on the WHO decision. There are also multiple links to gaming-related blogs with arguments against the WHO decision. The most widely shared scientific content in the dataset was the open debate paper by Aarseth et al. (2017). Other than that, there seems to be little to no circulation of scientific studies within the debate.

5. Discussion and Conclusions

Gaming disorder is currently the most intensively discussed form of problematic entertainment media use. Our analysis shows that social media platforms, such as Twitter, are important forums where different actors (including gamers) discuss the topic. Following an explorative approach, our study was the first to examine the gaming disorder debate based on social media data. It can serve as a basis for more complex future analyses.

Overall, our results showed that the debate is organic and not driven by spam accounts or other overly active 'power users.' There is no evidence of any orchestrated campaigns for or against the decision of the WHO.

Further, we see that analyzing the social media discussion has the potential to paint a more heterogenous and balanced picture of the public perception of gaming disorder than an analysis of classical media outlets where particular actors (like politicians, psychologists, and psychiatrists) are perhaps overrepresented. While it can be seen that CNN, as a traditional news medium, has a wide reach, it is largely followed by content creators; their level of participation is also higher. This suggests that traditional news media also play a role in social media for discussion, for instance, as a source of information, but the actual discussion is led by genuine stakeholders, such as the gamers themselves. A central distinction is that the accounts of news media represent organizations, while content creators are individuals who are given the opportunity to express their own opinions. While in traditional media, only public figures appear for their role as experts, on Twitter there is the chance to express one's thoughts through a medium without this prior decision.

With regard to topics and sentiment, our results showed that the social media discussion does more than cover the spectrum that previous studies have shown when examining traditional media (e.g., Kirkpatrick, 2016; Szablewicz, 2010; Whitton & Maclure, 2017). Although negative consequences of gaming are discussed, positive aspects are also emphasized. This suggests a diversification of the debate, which is also found in the academic discussion. Nevertheless, it can be seen that the discussion's sentiment is relatively negative. On the one hand, this is in line with the picture from tradi-

tional media; on the other hand, on closer examination, the terms used might rather indicate that users express their indignation about the WHO's decision. While previous research suggests that the discussion in traditional media focuses primarily on damage control, the topics we found indicate that there is a need to mention aspects that go beyond damage control, for example, the treatment, education, and significance of the decision. Thus, the discussion here is less to be compared with a moral panic, but rather attempts to differentiate.

The release of the ICD-11 draft and official version had a major impact on the debate. The decision moved the debate away from the topics of parenting and child welfare, largely by activating actors from gaming culture. The parenting and education-related discussion still took place, but it was overshadowed by a larger discussion. After the WHO decision, most tweets opposed the classification. Despite the research boom in recent years and the ongoing debate within academia, scientific studies barely played any role in the Twitter debate and research results were hardly considered. This can be interpreted as a hint that research results are perhaps not communicated effectively and are hardly known outside of an academic context.

From a more general perspective, the current study illustrates how computational methods, especially methods of automated content analysis, can be usefully utilized in the context of entertainment research. Using the example of the gaming disorder debate, we showed that these new tools can offer interesting insights into the public perception of risks that may be connected with the use of entertainment media. Future studies may follow this route and examine other diverse topics and their societal perception, such as the discussion about violent media content and aggressiveness or the question whether the use of digital entertainment media may negatively influence users' psychosocial well-being. Furthermore, our analytic framework showcases how social media (understood as a form of entertainment media itself) can be analyzed to identify subtopics and actors within large-scale debates. An interesting approach for future studies may be the combination of computational methods with qualitative methods to gain additional in-depth knowledge of selected networks and data patterns.

6. Limitations

The current study has some limitations. As we exclusively used Twitter data as the basis of our analysis, there might be a bias towards opinions shared by Twitter's relatively male and technophile userbase. The analysis is also limited to English language tweets, so there might be similar discourses in other languages, albeit using different hashtags.

The reduction in corpus size following our preprocessing procedure was relatively large, as a third of all Tweets were not considered in the topic modeling and a large share of tokens was removed due to frequency.

This loss of information might mean that some nuanced distinction between topics was not detected. The findings should thus be considered as a broad overview of the debate.

The sentiment analysis conducted in this study used a dictionary-based approach and the dictionary used only includes the binary distinction between negative and positive sentiment words (Hu & Liu, 2004). More sophisticated methods of sentiment analysis enable researchers to investigate more complex emotions like disgust, anger, or surprise either by using a dictionary that includes those categories or by using a corpus-based approach where supervised machine learning techniques are utilized (Strapparava & Mihalcea, 2008). The simple method used in the current study reveals the shift in tone caused by the WHO decision and gives insight into the potential reasoning of users behind their emotional reactions but does not reveal any more detailed information on the sentiment of the debate. Future research might address this using the methods referenced above.

Another limitation of the methods of the current study is the use of topic modeling on a corpus of documents with a relatively short length. As most traditional topic modeling techniques like the latent dirichlet allocation (Blei et al., 2003) and other probabilistic topic models like the STM (Roberts et al., 2019) used in the current study, rely on document-level word co-occurrence patterns to reveal topics. In short texts, as commonly found in social media, this co-occurrence approach may not work very well, as there is only limited word co-occurrence information available in these texts (Jipeng, Zhenyu, Yun, Yunhao, & Xindong, 2019). Future research might mitigate this issue by using methods specifically developed for shorter texts, like the biterm topic model (BTM) by Yan, Guo, Lan, and Cheng (2013).

The exclusion of emojis from the topic modeling was, on the one hand, driven by the differentiation between the topic and the tone of the debate in our research questions and, on the other hand, motivated by the need to reduce the level of noise in the already noisy and sparse data. While being an interesting question, we did not feel confident enough to address the tonality of the discussed topics in a robust way.

In general, as the methods applied in the current study are meant for the analysis of large-scale datasets, the above findings should not be seen as a complete description of every facet of the debate, but as an overview of the discourse, revealing overarching structures and topics worth investigating in greater detail.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Aarseth, E., Bean, A. M., Boonen, H., Colder Carras, M., Coulson, M., Das, D., . . . van Rooij, A. J. (2017). Scholars' open debate paper on the World Health Organization ICD-11 gaming disorder proposal. *Journal of Behavioral Addictions*, *6*(3), 267–270. https://doi.org/10/gb22vs

Aiello, L. M., Petkos, G., Martin, C., Corney, D., Papadopoulos, S., Skraba, R., . . . Jaimes, A. (2013). Sensing trending topics in Twitter. *IEEE Transactions on Multimedia*, *15*(6), 1268–1282. https://doi.org/10/f5bzrz

Billieux, J., King, D. L., Higuchi, S., Achab, S., Bowden-Jones, H., Hao, W., . . . Poznyak, V. (2017). Functional impairment matters in the screening and diagnosis of gaming disorder: Commentary on: Scholars' open debate paper on the World Health Organization ICD-11 Gaming Disorder proposal (Aarseth et al.). *Journal of Behavioral Addictions*, *6*(3), 285–289. https://doi.org/10/ggps6p

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(4/5), 993–1022.

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism*, *4*(1), 8–23.

Bowman, N. D. (2016). The rise (and refinement) of moral panic. In R. Kowert & T. Quandt (Eds.), *The video game debate: Unravelling the physical, social, and psychological effects of digital games* (pp. 22–38). London: Routledge.

Buzydlowski, J. W. (2015). Co-occurrence analysis as a framework for data mining. *Journal of Technology Research*, *6*, 1–19.

Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., . . . Helbing, D. (2012). Manifesto of computational social science. *The European Physical Journal Special Topics*, *214*(1), 325–346.

Denny, M. J., & Spirling, A. (2018). Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it. *Political Analysis*, *26*(2), 168–189.

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164.

Griffiths, M. D., Kuss, D. J., Lopez-Fernandez, O., &

Pontes, H. M. (2017). Problematic gaming exists and is an example of disordered gaming: Commentary on: Scholars' open debate paper on the World Health Organization ICD-11 Gaming Disorder proposal (Aarseth et al.). *Journal of Behavioral Addictions*, *6*(3), 296–301.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267–297.

Groshek, J., & Tandoc, E. (2017). The affordance effect: Gatekeeping and (non)reciprocal journalism on Twitter. *Computers in Human Behavior*, *66*, 201–210.

Habermas, J. (1991). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge, MA: MIT Press.

Hadgu, A. T., Garimella, K., & Weber, I. (2013). Political hashtag hijacking in the U.S. In D. Schwabe, V. Almeida, H. Glaser, R. Baeza-Yates, & S. Moon (Eds.), *Proceedings of the 22nd international conference on world wide web* (pp. 55–56). New York, NY: Association for Computing Machinery.

Halpern, D., & Gibbs, J. (2013). Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior*, *29*(3), 1159–1168.

Harris, Z. S. (1957). Co-occurrence and transformation in linguistic structure. *Language*, *33*(3), 283–340.

Hogenboom, A., Bal, D., Frasincar, F., Bal, M., de Jong, F., & Kaymak, U. (2013). Exploiting emoticons in sentiment analysis. In S. Shin & J. Maldonado (Eds.), *Proceedings of the 28th annual ACM symposium on applied computing* (pp. 703–710). New York, NY: Association for Computing Machinery.

Hopkins, D. J., & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science*, *54*(1), 229–247.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In K. Won, R. Kohavi, J. Gehrke, & W. DuMouchel (Eds.), *Proceedings of the 2004 ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 168–177). New York, NY: Association for Computing Machinery.

Jipeng, Q., Zhenyu, Q., Yun, L., Yunhao, Y., & Xindong, W. (2019, April 13). Short text topic modeling techniques, applications, and performance: A survey. *Cornell University*. Retrieved from http://arxiv.org/abs/1904.07695

Jung, C. W. (2019). Media discourse and perception of game regulatory issues. *The Communication Review*, *22*(2), 139–161.

Kirkpatrick, G. (2016). Making games normal: Computer gaming discourse in the 1980s. *New Media & Society*, *18*(8), 1439–1454.

Kleinman, Z. (2013, January 7). Gamers hired by father to 'kill' son in online games. *BBC*. Retrieved from https://www.bbc.com/news/technology-20931304

Levina, N., & Arriaga, M. (2014). Distinction and status

production on user-generated content platforms: Using Bourdieu's theory of cultural production to understand social dynamics in online fields. *Information Systems Research*, *25*(3), 468–488.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., . . . Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, *12*(2/3), 93–118.

Markey, P. M., & Ferguson, C. J. (2017). Internet gaming addiction: Disorder or moral panic? *American Journal of Psychiatry*, *174*(3), 195–196.

Massanari, A. (2017). #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society*, *19*(3), 329–346.

McCrum, K. (2015, November 3). Tragic teen gamer dies after 'playing computer for 22 days in a row.' *Mirror*. Retrieved from https://www.mirror.co.uk/news/world-news/tragic-teen-gamer-dies-after-6373887

Meerloo, J. A. M. (1954). Television addiction and reactive apathy. *The Journal of Nervous and Mental Disease*, *120*(3), 290–291.

Paul, I., Khattar, A., Kumaraguru, P., Gupta, M., & Chopra, S. (2019). Elites tweet? Characterizing the Twitter verified user network. *Proceedings of the 2019 IEEE 35th international conference on data engineering workshops* (pp. 278–285). Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Paulussen, S., & Harder, R. A. (2014). Social media references in newspapers: Facebook, Twitter and YouTube as sources in newspaper journalism. *Journalism Practice*, *8*(5), 542–551.

Pervin, N., Phan, T. Q., Datta, A., Takeda, H., & Toriumi, F. (2015). Hashtag popularity on twitter: Analyzing co-occurrence of multiple hashtags. In G. Meiselwitz (Ed.), *Social computing and social media* (pp. 169–182). Cham: Springer.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(2), 1–40.

Rumpf, H.-J., Achab, S., Billieux, J., Bowden-Jones, H., Carragher, N., Demetrovics, Z., . . . Poznyak, V. (2018). Including gaming disorder in the ICD-11: The need to do so from a clinical and public health perspective: Commentary on: A weak scientific basis for gaming disorder: Let us err on the side of caution (van Rooij et al.). *Journal of Behavioral Addictions*, *7*(3), 556–561.

Schatto-Eckrodt, T., Janzik, R., Reer, F., Boberg, S., & Quandt, T. (2020). Supplementary material to "A computational approach to analyzing the Twitter debate on gaming disorder". *OSF Home*. Retrieved from https://osf.io/vzymj

Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H., & Deng, X. (2013). Exploiting topic based twitter sentiment for stock prediction. In H. Schuetze, P. Fung, & M. Poesio (Eds.), *Proceedings of the 51st annual meeting of the Association for Computational Linguistics* (pp. 24–29). Stroudsburg, PA: Association for Computational Linguistics.

Strapparava, C., & Mihalcea, R. (2008). Learning to identify emotions in text. In R. Wainwright & H. Haddad (Eds.), *Proceedings of the 2008 ACM symposium on applied computing* (pp. 1556–1560). New York, NY: Association for Computing Machinery.

Szablewicz, M. (2010). The ill effects of "opium for the spirit": A critical cultural analysis of China's Internet addiction moral panic. *Chinese Journal of Communication*, *3*(4), 453–470.

Thomas, K., Grier, C., Song, D., & Paxson, V. (2011). Suspended accounts in retrospect: An analysis of Twitter spam. In P. Thiran & W. Willinger (Eds.), *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement* (pp. 243–258). New York, NY: Association for Computing Machinery.

Thompson, P. (2011, June 7). 'Sorry' mother jailed for 25 years for allowing her daughter to STARVE to death while she played an online video game. *Daily Mail*. Retrieved from https://www.dailymail.co.uk/news/article-1394903/Rebecca-Colleen-Christie-jailed-25-years-allowing-daughter-Brandi-Wulf-STARVE-death-played-World-Warcraft.html

Twitter. (2020, March 20). About verified accounts. *Twitter Help Center*. Retrieved from https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts

VanDam, C., & Tan, P.-N. (2016). Detecting hashtag hijacking from Twitter. In W. Nejdl, W. Hall, P. Parigi, & S. Staab (Eds.), *Proceedings of the 8th ACM conference on web science* (pp. 370–371). New York, NY: Association for Computing Machinery.

van Atteveldt, W., & Peng, T.-Q. (2018). When communication meets computation: Opportunities, challenges, and pitfalls in computational communication science. *Communication Methods and Measures*, *12*(2/3), 81–92.

van den Brink, W. (2017). ICD-11 Gaming Disorder: Needed and just in time or dangerous and much too early? Commentary on: Scholars' open debate paper on the World Health Organization ICD-11 Gaming Disorder proposal (Aarseth et al.). *Journal of Behavioral Addictions*, *6*(3), 290–292.

van Rooij, A. J., Ferguson, C. J., Colder Carras, M., Kardefelt-Winther, D., Shi, J., Aarseth, E., . . . Przybylski, A. K. (2018). A weak scientific basis for gaming disorder: Let us err on the side of caution. *Journal of Behavioral Addictions*, *7*(1), 1–9.

Wang, X., Wei, F., Liu, X., Zhou, M., & Zhang, M. (2011). Topic sentiment analysis in twitter: A graph-based hashtag sentiment classification approach. In B. Berendt, A. de Vries, W. Fan, C. Macdonald, I. Ounis, & I. Ruthven (Eds.), *Proceedings of the 20th ACM international conference on information and knowledge management* (pp. 1031–1040). New York, NY: Association for Computing Machinery.

Whitton, N., & Maclure, M. (2017). Video game discourses and implications for game-based education. *Discourse: Studies in the Cultural Politics of Education*, *38*(4), 561–572.

Wojcik, S., & Hughes, A. (2019). Sizing up Twitter users. *Pew Research Center*. Retrieved from https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users

World Health Organization. (2019). 6C51 Gaming disorder. *International Classification of Diseases 11th Revision*. Retrieved from https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1448597234

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. In D. Schwabe, V. Almeida, H. Glaser, R. Baeza-Yates, & S. Moon (Eds.), *Proceedings of the 22nd international conference on world wide web* (pp. 1445–1456). New York, NY: Association for Computing Machinery.

Zhang, L., Ghosh, R., Dekhil, M., Hsu, M., & Liu, B. (2011). *Combining lexicon-based and learning-based methods for twitter sentiment analysis* [Technical report HPL-2011-89]. Palo Alto, CA: HP Laboratories.

About the Authors

Tim Schatto-Eckrodt is a Research Associate at the Department of Communication at the University of Münster, Germany. He holds an MA degree in Communication Studies from the same department. He is currently working on his PhD project about online conspiracy theories and is part of the junior research group 'Democratic Resilience in Times of Online Propaganda, Fake News, Fear and Hate Speech (DemoRESILdigital).' His further research interests include computational methods and online propaganda.

Robin Janzik is a Research Associate in the Department of Communication at the University of Münster, Germany. He holds an MA degree in Communication Studies from the University of Münster, Germany, and is currently working on his PhD project about the acceptance of virtual reality technology for private use. His further research interests include trust in technology and problematic media use.

Felix Reer is a Postdoctoral Researcher in the Department of Communication at the University of Münster, Germany. His research interests include social media and online communication, effects of digital games, and the use of highly immersive media technologies, such as virtual reality and augmented reality devices. He is Vice-Chair of the Digital Games Research Section of the European Communication Research and Education Association (ECREA).

Svenja Boberg is a Communication Scientist at the University of Münster, who researches the dynamics of online debates and social media hypes using computational methods. In her dissertation she studies the spread of outrage in social media networks. In 2016 she joined the BMBF-funded project 'Detection, Proof and Combating of Covert Propaganda Attacks via Online Media.' Beforehand, she completed her MA in Communication Studies at the University of Münster addressing news consumption on Facebook.

Thorsten Quandt is a Professor of Online Communication at the University of Münster, Germany. His research fields include online communication, digital games, VR and journalism. Quandt is particularly interested in the societal changes connected to the Internet and new media, and the question how human beings have evolved in sync with these changes. His earlier works on participatory journalism and online newsroom production have been widely cited in the field of (digital) journalism research.