Recibido: 22-10-2024 / Aceptado:31-12-2024

Cross-platform *hatedom*: influencers' strategies for managing affective aversion

Hatedom multiplataforma: estrategias de los influencers para gestionar la aversión afectiva

Hatedom multi-plataforma: estratégias dos influencers para gerir a aversão afectiva

Arantxa Vizcaíno-Verdú, Universidad Internacional de la Rioja, Logroño, España (arantxa.vizcaino@unir.net)

Paloma Contreras-Pulido, Universidad Internacional de la Rioja, Logroño, España (paloma.contreras@unir.net)

ABSTRACT Amidst the rise of toxic behavior in online fan communities, this study examines the construction of cross-platform hatedom through social media and the strategies influencers use to navigate hostile digital environments. Using a mixed methods approach, this study analyzes 16,215 comments directed at six influencers using advanced natural language processing techniques and pre-trained AI models. In addition, structured qualitative interviews provided insights into the influencers' perceptions of their interactions with fans and haters and shed light on their strategies for dealing with hatedom. The results show that hatedoms, characterized by strong affective dislike, vary significantly in their levels of toxicity and emotional polarization on different platforms, including YouTube, X, Facebook, Instagram, TikTok and Twitch. Influencers employ various coping mechanisms, such as ignoring negative interactions, blocking malicious users, and using sarcasm as a defense mechanism. They also emphasize the importance of mental health and often limit their personal engagement to protect their wellbeing. This study highlights the need for a comprehensive, multi-faceted approach to understanding hatedom that combines insights into the polarization of the digital community with the implementation of advanced technological moderation. By prioritizing the mental health and resilience of influencers, it is possible to foster a healthier digital environment amidst the ever-evolving dynamics of fan behavior on social media.

KEYWORDS: fandom, hate speech, platform, social media, toxicity, followers, online community, influencers

FORMA DE CITAR

Vizcaíno-Verdú, A. & Contreras-Pulido, P. (2025). Cross-platform *hatedom*: influencers' strategies for managing affective aversion. *Cuadernos.info*, (60), 68-92. https://doi.org/10.7764/cdi.60.85872

RESUMEN Ante el crecimiento de comportamientos tóxicos en las comunidades de fans en línea, este estudio examina la construcción del hatedom en múltiples plataformas de redes sociales y las estrategias que los influencers emplean para navegar en entornos digitales hostiles. Mediante un enfoque de métodos mixtos, la investigación analiza 16.215 comentarios dirigidos a seis influencers, utilizando técnicas avanzadas de procesamiento de lenguaje natural y modelos de inteligencia artificial preentrenados. Paralelamente, por medio de entrevistas cualitativas estructuradas, se obtuvo información sobre las percepciones de estos influencers respecto de sus interacciones con sequidores y detractores, presentando las estrategias empleadas para gestionar el odio recibido. Los hallazgos revelan que el hatedom, caracterizado por una fuerte aversión afectiva, presenta variaciones significativas en sus niveles de toxicidad y polarización emocional en diferentes plataformas, como YouTube, X, Facebook, Instagram, TikTok y Twitch. Los influencers adoptan diversas estrategias de resistencia, entre las que destacan ignorar interacciones negativas, bloquear a usuarios malintencionados y utilizar el sarcasmo como mecanismo defensivo. Asimismo, subrayan la importancia de proteger su salud mental, limitando su exposición digital para salvaguardar su bienestar. Este estudio resalta la necesidad de un enfoque integral y multifacético para comprender el fenómeno del hatedom, combinando el análisis de la polarización de las comunidades digitales con la implementación de herramientas tecnológicas avanzadas para la moderación. Al priorizar la salud mental y la resiliencia de los influencers, es posible promover un entorno digital más saludable en una evolución constante del comportamiento de los fans en redes sociales.

PALABRAS CLAVE: fandom, discurso de odio, plataforma, redes sociales, toxicidad, seguidores, comunidad en línea, influencers

RESUMO Perante o aumento de comportamentos tóxicos nas comunidades de fãs online, este estudo examina a construção do hatedom em múltiplas plataformas de redes sociais e as estratégias que os influencers utilizam para navegar em ambientes digitais hostis. Utilizando uma abordagem de métodos mistos, a pesquisa analisa 16.215 comentários dirigidos a seis influencers, utilizando técnicas avançadas de processamento de linguagem natural e modelos de inteligência artificial pré-treinados. Paralelamente, foram utilizadas entrevistas qualitativas estruturadas para obter informações sobre as percepções destes influencers em relação às suas interações com seguidores e detractores, apresentando as estratégias utilizadas para gerenciar o ódio recebido. Os resultados revelam que o hatedom, caracterizado por uma forte aversão afectiva, apresenta variações significativas nos seus níveis de toxicidade e polarização emocional em diferentes plataformas, inlcuindo YouTube, X, Facebook, Instagram, TikTok e Twitch. Os influencers adoptam uma variedade de estratégias de resistência, incluindo ignorar interações negativas, bloquear utilizadores maliciosos e usar o sarcasmo como mecanismo de defesa. Salientam também a importância de proteger a sua saúde mental, limitando a sua exposição digital para salvaguardar o seu bem-estar. Este estudo realça a necessidade de uma abordagem holística e multifacetada para compreender o fenômeno do hatedom, combinando a análise da polarização das comunidades digitais com a implementação de ferramentas tecnológicas avançadas de moderação. Ao dar prioridade à saúde mental e à resiliência dos influencers, é possível promover um ambiente digital mais saudável em uma constante evolução do comportamento dos fãs nas redes sociais.

PALAVRAS CHAVE: fandom, discurso de ódio, plataforma, redes sociais, toxicidade, seguidores, comunidade online, influencers

INTRODUCTION AND THEORETICAL FRAMEWORK

The Oxford English Dictionary chose toxic as the word of the year 2018, reflecting the growing concern about harmful attitudes and behaviors in social contexts, especially in online relationships (Schuessler, 2018). This phenomenon has been analyzed by authors such as Arouh (2020), who has explored the use and implications of the concept in contemporary culture, particularly in fandoms. In these communities, toxic behaviors often manifest as harsh criticism or negative opinions that can escalate into more severe forms of aggression, sometimes meeting the criteria for hate speech (Díaz et al., 2019).

Hate speech is considered one of the most pressing problems in today's society. The recent United Nations report highlights how social media algorithms amplify online hate speech (OHCHR, 2023) by prioritizing content that elicits heightened emotional responses such as outrage, shock or anger (Maarouf et al., 2024). While hate speech involves hostile acts aimed at inciting harm or violence to individuals based on characteristics such as ethnicity, gender, or religion, the same algorithmic mechanisms extend their influence to fan communities and manifest in different forms rooted in a dynamic interplay of hostility, social validation, and group polarization (Riemer & Peter, 2021).

Fan communities are particularly susceptible to these dynamics due to the intense emotional engagement that defines their social interactions. Fans display a high degree of affective investment (Théberge, 2006) and adhere to authoritarian rules (Luo & Li, 2024), demonstrating a deep emotional devotion to their fandoms. However, this devotion is not always positive. As Jenkins (1992) explains, the same intensity that drives affection for fandom can also lead to deviant or toxic behavior. Barnes (2022) illustrates how this affective intensity can lead to a sudden shift from admiration to hatred, especially when public figures or content creators are perceived as traitors to their fandom.

This change can be seen in the emergence of anti-fans (Duffy et al., 2022), who take an antagonistic stance towards certain creators, often driven by perceived moral or social grievances (Aburime, 2022). Anti-fans engage negatively and are characterized by an obsession with an online influencer or celebrity. Vizcaíno-Verdú and colleagues (2020) describe this relationship as fanbullying, understood as an inherently abusive bond between fans and public figures. This dynamic ensures that the roles of fictional characters are no longer believed and directs the hatred of the community towards actors, actresses and influencers. In this sense, Valenzuela-García and colleagues (2023) refer to influencers as non-ideal victims of this hatred because, unlike other victims (e.g., anonymous individuals suffering from cyberbullying), influencers are not perceived as such due to their public and celebrated status, which does not generate empathy and appreciation.

According to Ouvrein and colleagues (2023), influencers generate envy and contempt, emotions that justify contemptuous attitudes and bullying. Valenzuela-García and colleagues (2023) conclude that bullies express a dislike for the lifestyle of influencers because they perceive it as unfair, believing that they lack personal development or effort to achieve such privileges. This perception of incompetence or unfairness stems from the belief that influencers gain their wealth by selling their lives.

Furthermore, online platforms facilitate this hatred through the disinhibition effect (Suler, 2006), where the line between fiction and reality becomes blurred (Vizcaíno-Verdú et al., 2020). Added to this are the effects of algorithms on freedom of expression in social media (Riemer & Peter, 2021). In this digital world, algorithms amplify or suppress discourse in order to disseminate content more efficiently, distorting the fair and free exchange of ideas in public discourse (Martínez-Valerio & Mayagoitia-Soria, 2021; Riemer & Peter, 2021).

Riemer and Peter (2021) found that opaque algorithms shift the question of free speech from what can be said to what will be heard and who will hear it. As a result, influencers often resorted to self-censorship. In recent years, social media has enabled public figures to express their political or ideological views. However, recent studies, including Martínez-Valerio and Mayagoitia-Soria (2021), suggest a decline in this behavior due to the prevalence of cyberbullying and the need to protect brand reputation (von-Mettenheim & Wiedmann, 2022).

Influencers are therefore forced to communicate with their followers through effective digital strategies across different, potentially integrated platforms, developing different approaches to handling discourse and dealing with cyberbullying (Abidin, 2019). These strategies include feelings of guilt or shame regarding messages received, which are often hidden from the public because they have been previously blocked, filtered or deleted; not reporting or silencing potentially criminal acts; self-censorship on certain topics; or leaving platforms where they feel most attacked (Martínez-Valerio & Mayagoitia-Soria, 2021).

According to Walther's (2022) study, platforms that enable user feedback (e.g. likes on Facebook, favorites on X (Twitter) or upvotes on Reddit) reinforce the recipients' behavior. Such interactions can trigger three important processes: (1) incentivizing hostile individuals to post more hate messages, (2) intensifying hatred towards their targets as people tend to take publicly defended positions, especially when they receive positive feedback, and (3) developing deep relationships that reinforce toxic messages through continuous interactions with like-minded individuals, even if these relationships are purely virtual (Walther, 2022).

Given this, there is growing academic interest in the study of relationships between fans, followers and influencers, particularly with the proliferation of platforms and the opportunities for communities to interact and participate. One pressing issue that requires attention is the rise of toxic fandoms. While recent studies have analyzed the personal, physical, and psychological effects of this type of hate on fandom personalities (Walther, 2022; Martínez-Valerio & Mayagoitia-Soria, 2021; Vizcaíno-Verdú et al., 2020), this research examines how anti-fans construct hatedom and how influencers manage it through social media platforms. The specific aims that guided this approach were: (O1) to identify cross-platform polarity and toxicity in comments received by influencers; (O2) to explore the role and relevance of fan communities to influencers' labor; (O3) to examine how influencers interact with fan and hater communities; and (O4) to explore the strategies influencers use to manage hatred on social media. This comprehensive approach allowed for an exploratory yet in-depth understanding of the interaction dynamics of communities on social media characterized by highly polarized relationships between affection and hate towards public and online celebrities.

METHOD

We used a mixed methods approach that integrates both quantitative and qualitative techniques. First, we used natural language processing (NLP) and pre-trained AI models to analyze the likelihood of negativity and toxicity in the comments on the six influencers' social media profiles. We then conducted interviews with these influencers to gain insights into their perceptions and strategies for dealing with hostile community interactions.

NLP models for analyzing polarity and toxicity and comments' sample

In the first phase of the analysis, we examined the polarity and toxicity of comments received by influencers. Comments were collected using web scraping techniques (Mysiuk & Shuvar, 2023), and data were systematically extracted from each profile (see the influencers' sample section). A broad selection of platforms was made to ensure a random and representative comment sample for the analysis. Specifically, for each influencer, we selected the last three posts from May 2024 on all active platforms (table 2), resulting in 7,236 comments on Instagram, 5,131 on YouTube, 4,013 on TikTok, 1,664 on X, 1,380 on Facebook and 1,072 on Twitch.

Before conducting the analysis, the comments were pre-processed using the Pandas library to eliminate possible coding and data delimitation errors. The filtering function aimed to remove emoji, non-alphanumeric characters, text after the @ symbol to ensure the anonymity of commenting and tagging users, links (spam) and blank lines. Special and relevant Spanish characters, such as accents (e.g., á), the letter ñ and punctuation marks, were retained. The final sample included 5,341 comments on Instagram, 4,782 on YouTube, 3,031 on TikTok, 1,472 on X, 1,033 on Twitch and 556 on Facebook, for a total of 16,215 units of analysis.

To detect and classify the polarity and toxicity of the comments, we used two NLP models: (1) BETO for sentiment analysis and (2) Multilingual BERT for toxicity detection. These models are based on the bidirectional pre-training technique of Bidirectional Encoder Representations from Transformers (BERT), which is validated in terms of understanding the context of a word in all its positions within a sentence (Devlin & Chang, 2018). The selection of these models is based on several factors:

- 1. BETO. Specialized in Spanish, provides high precision in the classification of emotions (negative, neutral and positive), which is crucial for understanding emotions in comments from Spanish-speaking influencers.
- 2. Multilingual BERT. With its multilingual capability and specialized toxicity detection training, this model provides a robust tool for identifying and classifying offensive and aggressive language.
- 3. Transformer architecture: Both models use the transformer architecture, which captures complex contextual relationships and improves the detection of nuances in both sentiment and toxicity.

The technical implementation of the sentiment and toxicity analyzes was carried out in several steps. First, we loaded both the pre-trained models and the corresponding tokenizers into Google Colab using the transformer library. The BETO tokenizer converted the comments into token sequences that could be processed by the model, enabling sentiment analysis. Similarly, the Multilingual BERT tokenizer delimited the comments for toxicity analysis.

Logit matrices were created for each comment, representing the probability that a comment belonged to different sentiment and toxicity categories. These categories included positive, neutral and negative for sentiment and toxicity, insult, threat and obscenity for toxicity. For example, if a comment had logits of 2.0, 1.0, and 0.1 for the sentiment categories (positive, neutral and negative), the softmax function converted these values into probabilities of 0.65, 0.24 and 0.11 respectively. This indicates a 65% probability that the comment is positive, a 24% probability that it is neutral, and an 11% probability that it is negative. Similarly, in the toxicity analysis, the logits were converted to probabilities indicating the nature of a comment. For example, a comment might have a 45% probability of being considered toxic, 30% probability of containing an insult, 20% probability of expressing a threat, and 5% probability of containing obscenities.

Objectives	Framework	Open questions
O2. Exploring the role and relevance of fan communities for influencers' labor.	(Valenzuela-García et al., 2023; Abidin, 2019; Vizcaíno- Verdú et al., 2020)	(1) What do you think of your social media fans? What do they represent?
		(2) What is your perspective on the role of fans in your labor as an influencer?
O3. Examining how influencers interact with the fan community on various social media platforms.	(Aburime, 2022; Abidin, 2019; Riemer & Peter, 2021; Vizcaíno- Verdú et al., 2020; Walther, 2022).	(3) Have you noticed any significant differences between your follower communities on the different platforms? How do they vary?
		(4) Do you consider your follower community to be characterized by respect and positivity or have you experienced any controversy or conflict in your social media profiles?
O4. Delving into the strategies that influencers use to manage hatedom on social media.	Arouh, 2020; Abidin, 2019; Martínez-Valerio & Mayagoitia-Soria, 2021; Valenzuela- García et al., 2023).	(5) How do you handle negative comments and consructive criticism of your posts? Have you ever had to deal with hates and toxicity?
		(6) What measures do you take to foster a respectful and positive environment for your community?
		(7) How do you interact with your followers? Do you try to respond to all comments and messages or do you prioritize certain interactions?
		(8) Do you think that it is important to establish clear boundaries with your fans? How do you maintain a healthy balance between your personal life and social media presence?
		(9) Have you experienced conflicts among your fans due to differences in opinions or loyalty to other influencers? How do you handle these situations?
		(10) Have you noticed changes in community dynamics over time? Have there been moments when you felt you lost or gained followers owing to certain actions or statements?

Table 1. Structured interviews

Source: Own elaboration.

The results were stored in DataFrame and exported to a spreadsheet for analysis to get a comprehensive overview of the hostile interactions on the influencers' profiles.

Structured interviews to influencers

We focused on influencers' perspectives and conducted structured interviews from June 2023 to February 2024 to dive deep into these individuals' accounts of their experiences in relation to Hatedom (Sjoerd-Jeroen et al., 2024).

The interviews followed a structured model with ten open-ended questions aligned with the research objectives and inspired by the studies of Valenzuela-García and colleagues (2023), Martínez-Valerio and Mayagoitia-Soria (2021), Walther (2022), and Vizcaíno-Verdú and colleagues (2020). This standardized design (table 1) was used to try to keep the duration of the interviews as short as possible, as media professionals with a tight schedule are not able to conduct longer interviews with the researchers.

Influencers' sample

To address the high media demand for influencers and the associated low response rate (Senft, 2013), we invited 22 Spanish influencers to participate, with the aim of achieving a sufficient response rate for meaningful qualitative data (Creswell, 2013). Influencers were selected through non-probabilistic random sampling that took into account the cultural and operational context of the researchers and focused on influencers' activity and recognition on different platforms characterized by unique community dynamics (Abidin, 2018).

Six influencers accepted the invitation. Due to their professional and personal commitments and the sensitivity of dealing with hate and conflict with their communities, we guaranteed their anonymity to protect their careers and online reputations. Table 2 provides details on each influencer's main content, platforms used and number of followers (as of June 2024).

Pseudonym of the influencer	Main theme	Social media platform	Followers
		TikTok	398,000
INFLU1	Art and make- up	Instagram	162,000
	up	YouTube	54,5000
	11	TikTok	794,9000
INFLU2	Humor	Instagram	207,000
		TikTok	1.4 million
	Lifestyle, singing, and dancing	Facebook	697,000
INFLU3		Instagram	260,000
		YouTube	14,600
		Х	915
	Humor	TikTok	397,600
INFLUA		Instagram	276,000
INFLU4		Х	27,500
		YouTube	675
		TikTok	2.4 million
	Humor	YouTube	922,000
INFLU5		Instagram	378,000
		Twitch	5,786
	Scientific dissemination	YouTube	786,000
		Instagram	331,000
INFLU6		Х	256,756
		TikTok	162,400
		Twitch	61,741

Table 2. Influencers' sample

Source: Own elaboration.

Analysis of transcriptions using grounded theory

After transcribing the interviews with the automatic support of Microsoft Office and manually filtering the arguments for clarity, the answers were translated from Spanish into English for the qualitative analysis using Atlas.ti 24 software. The grounded theory method (Strauss & Corbin, 1997) guided the inductive-deductive analysis through several phases: (1) open coding, (2) axial coding, and (3) memoing.

In the open coding phase, the transcripts were broken down into meaningful text units, each of which was assigned a descriptive code. Through this process, initial patterns and concepts were identified, such as (a) employee support, (b) negative comments, (c) personal interaction, (d) management strategies and (e) positive environment.

Axial coding involved grouping the codes into broader categories to establish relationships between them and structure the data to explore concept connections. Axial codes included (a) relationship with follower community, (b) dealing with negative comments, (c) differences between platforms, and (d) fostering a respectful environment.

In the memoing phase, key categories were identified according to the research framework and objectives in order to develop a narrative interpretation and understand the community dynamics between influencers, fans and haters. Three key memos were extracted: (1) relationships with fan communities; (2) interactive differences with communities on different platforms; and (3) strategies for dealing with hostile comments and hate conflicts within the community.

In compliance with ethical research principles and to ensure the protection of human and civil rights in accordance with the Code of Ethics of the Association of Internet Researchers (2012), both influencers and users commenting on their content were anonymized due to the sensitive nature of the findings.

FINDINGS AND DISCUSSION

Objective 1. Polarity and toxicity in influencer profiles

Starting with the computational analysis to understand the community context of the influencers, especially in terms of polarity and toxicity, significant differences in the reception of comments were found depending on the profile of the influencer (figure 3). INFLU1, who focuses on art and make-up, received primarily neutral comments (59%), a considerably low proportion of negative comments (8%) and toxic comments (6%). INFLU2, with its humorous theme, recorded 24% negative comments and an overall toxicity of 11%. In this case, the comments often contained dark humor, sarcasm or irony, which contributed to

the perception of higher toxicity. Examples with a high percentage of negativity and toxicity include: "A group of lifeless people who get anxious about everything", "Everyone has a grudge against the neighbor, mine is kind of perverted and a drug addict", and "Why do you have boring lives and are too lazy to talk to the neighbor about your boring lives?".

INFLU3, on the other hand, with 5% negative comments and 11% toxicity, receives criticism related to his sexual orientation — "Are you gay? What a waste, you are so handsome", or his lifestyle— "What are you doing living in disgusting London? You have a gorgeous villa". INFLU4, which focuses on humor and monologs, was met with 24% negative comments, with toxicity at 9%. Comments in this case included: "Unbelievable how some people are such pleasers" or "Bring back the Caudillo, how disgusting"—referring to Francisco Franco, who ruled Spain as dictator from 1939 to 1975, expressing nostalgia for an authoritarian regime.



Figure 1. Negative and toxic comments on INFLU6.

Source: Own elaboration.

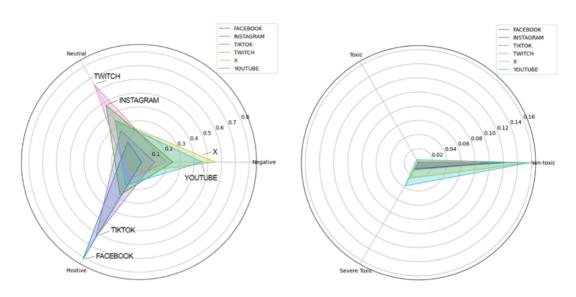


Figure 2. Cross-platform polarity and toxicity among influencers
Source: Own elaboration.

INFLU5 also showed a moderate level of negativity (21%) and toxicity (9%). Criticism was directed at the influencer on personal and content grounds— "Ungrateful daughter" and "Nobody wants to see you"; as well as on gender grounds— "It is very difficult to understand women" and "Many men are worried about savings and the future, and it seems stupid to them to spend money on unnecessary things".

In relation to INFLU6, which focuses on scientific dissemination, the sample showed the highest percentage of negative comments (50%) and toxicity (16%). The toxic forms showed minimal levels of profanity (3%) and insults (2%), highlighting the controversy and polarization inherent in the published content (figure 1).

In terms of polarization and toxicity across platforms, X and YouTube had a higher percentage of negative comments (55.5% and 46% respectively) than Instagram (24.6%), Twitch (19.9%), TikTok (11.2%) and Facebook (1.9%). Positive comments predominated on Facebook (81.4%) and TikTok (62.6%), indicating a respectful community. Neutral comments were most common on Twitch (65.2%) and Instagram (47.8%) (figure 2).

The toxicity values also differed considerably between the platforms. X and YouTube had the highest proportion of toxic comments at 15.8% and 15.7%, respectively. In contrast, Facebook, Instagram, TikTok, and Twitch had lower scores of 12.1%, 11.3%, 10.5%, and 8.7%, respectively. Obscene language occurred most frequently on YouTube (3.8%) and was minimal on Facebook (1%). In addition, threats and insults remained consistently low across all platforms, with insults ranging from 0.3% to 1.9%, and threats below 0.1%.

The high proportion of negative comments on X and YouTube indicates a more critical and/or confrontational user base, possibly influenced by the open and often anonymous communication on these platforms. In contrast, the positive comments on TikTok, Facebook and Instagram suggest that these platforms foster a more positive and cooperative community environment for the cases analyzed. In particular, the lower levels of toxicity on Twitch and TikTok (in terms of abusive language, profanity, and threats) may be due to content moderation policies and community guidelines that are more effective at curbing negative interactions and behaviors.

Objective 2. Relevance of fan communities to influencers

In this scenario, we explored the perceptions of influencers by examining their relationship with fan communities and addressing Q1 and Q2. Digital celebrities consistently view their followers as a source of support, primarily through emotional validation and constructive feedback.

For me, they signify that my content is worthwhile and appreciated, providing constant feedback and motivation for improvement (Q1-INFLU2).

They are a significant support system for me (Q1-INFLU4).

According to these figures, the community is a fundamental pillar for sustaining their work, with followers playing a co-creative role in the content. Consequently, this joint project underlines the central role of followers, especially fans. Interviewees acknowledged that without fans, their content would lose impact and relevance, suggesting that the success of their content is dependent on a positive relationship. This dependency was evident in statements such as the following,

They mean the world for me. Without them, you are no one. Humility is crucial. A content creator without followers has nothing to do (Q2-INFLUI).

Followers are the cornerstone. If there were no audience interested in my work, I would not be able to pursue it professionally (Q2-INFLU6).

The relationship between influencer and fan also includes a crucial economic component. Followers not only endorse the content or participate in its creation, but also represent an important segment for brands seeking sustainable and profitable collaboration. This dynamic transforms fans into an indirect source of revenue that is critical to the financial viability of the influencer. Influencers emphasize this personal and lucrative relationship by noting, for example, that,

The followers represent two clear things: a real connection and an economic engine. [...] You do not stop "working for them" in a way. If they do not like your videos, you will not eat that month. So it is a pretty intense relationship that one has to learn to manage (Q2-INFLU3).

If I did not have a community of committed followers who engaged with my posts regarding paid collaborations with brands, the brands would not hire me (Q2-INFLU4).

Followers provided meaningful support in terms of their emotional well-being. Expressing affection can have a positive impact on influencers' mental health. This is crucial in a digital environment where public pressure and scrutiny can become intense (Théberge, 2006) and often boosts the self-esteem of these online creators. Some noted that,

They always give me encouragement and help me when I share my problems (Q1-INFLU5).

The stories they tell me and my own emotional character make my relationship with them not only close but also frank; I have been moved by reading comments, and I enjoy talking to them personally, so when I create content, I have a real desire to make them smile and to get something out of consuming my content (Q2-INFLU3).

As Jenkins (1992) notes, the relationship between influencers and followers and fans is multi-layered and highly interdependent. This community provides important validation, constructive feedback and emotional support, which are crucial for influencers' motivation and continuity of their labor. However, this relationship leads to discrepancies when examining this dynamic in different social media communities.

Objective 3. Community differences between social media platforms

The differences in influencers' experiences with demographics and user behavior across platforms are notable from Q2 to Q4. For example, influencers reported that there are often more negative comments on YouTube and X, typically from non-followers. TikTok encourages a more respectful and youthful dynamic, while Instagram elicits mixed views. Some influencers associate Instagram with a mature and intimate community, while others find it critical and offensive (figure 3).

Comments on TikTok often appear immature, but not malicious. A predominantly young user base, including millennials and Generation Z, typically exhibit less sophisticated communication. Consequently, their comments, while sometimes negative, do not have the harmful intent that is common on YouTube. This distinction underscores the influence of demographic factors and platform features, including hate speech filters and reporting tools on interaction dynamics. It also highlights the increased disinhibition effect in online environments (Suler, 2006).

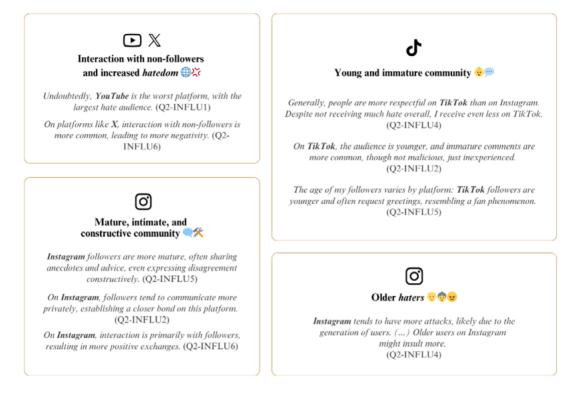


Figure 3. Influencers' perceptions of fandom and hatedom on social media

Source: Own elaboration.

In contrast, Instagram stands out in this discourse as a platform that encourages closer and more personal connections between influencers and fans. However, this interaction varied depending on the experience. Some profiles described the app as a private, intimate and respectful environment built on a foundation of bidirectional direct messages and unidirectional group chats. Others, however, describe the app as a harmful platform where adult users indiscriminately attack content and people.

This interaction highlights the importance of distinguishing the character of a non-follower (external user), which is primarily negative, from that of a follower or fan, which is perceived as a profile that contributes positively to the community. In other words, it is a fringe social group characterized by negativity, conflict and hate. For example, when it comes to feminist, progressive or controversial topics, non-followers, especially anonymous users, project their hate through comments that "even cross the boundaries of legality" (Q4-INFLU6).

The algorithms used by the platforms significantly influence the relationship between influencers and their communities, supporting the findings of Riemer and Peter (2021) and Whalter (2022). One of the interviewees noted significant differences in the way the algorithmic systems of different platforms automatically detect malicious content to prevent or restrict its reception, and acknowledged that "the algorithm filters hate differently on each platform" (Q3-INFLU3). In this context, the differences between platforms in terms of the types of comments and demographics of followers are essential to understanding the dynamics of in-groups and out-groups. Influencers adopt different strategies for organic communication and community management (Ouvrein et al., 2023), highlighting the need for a comprehensive understanding of the dynamics of hate confrontation on each social media platform beyond victimization (Abidin, 2019; Vizcaíno-Verdú et al., 2020).

Objective 4. Strategic and mediatory management of hatedom

Most influencers reported having a fan base characterized by a strong aura of respect and positivity (Q4-Q10). They emphasized that this healthy environment is essential for their emotional well-being and the sustainability of their labor and presence in the digital world.

In general, my profile has hardly any hate, mostly white humor (Q4-INFLU2).

My community of followers has always stood out with respect to any platform. That is, in terms of fans, who are loyal and like what I do (Q4-INFLU4).

My followers are always respectful even when they do not agree with something that I do (Q4-INFLU5).

I consider people who follow me constructive (Q4-INFLU6).

This type of positive interaction not only strengthens the creator's motivation, but also fosters a sense of belonging and cohesion within the community (Barnes, 2022). In terms of dealing with hostile comments, influencers reported using organic strategies that complement the hate management systems provided by the platforms themselves (figure 4). This proactive management appears to be crucial in maintaining the quality of interaction and protecting the wellbeing of both content creators and their fans.

The strategy (1) of ignoring negative comments is widespread and in many cases effective. By not making such hostile messages visible, influencers avoid stirring up hate and maintain a positive focus on both live and asynchronous conversations and interactions. Another tactic is to (2) respond sarcastically or with humor. They claimed that this approach can disarm the most harmful and critical users and turn a potentially hostile situation into an opportunity for entertainment or profiling. In this context, one influencer noted: "You can't let your followers' comments rule your life" (INFLU6-Q10). Consequently, in these cases, they resort to (3) blocking the user, filtering words or reporting comments via the platform system. These methods of defending against hatedom are consistent with the cognitive and behavioral strategies presented by Ouvrein et al. (2023).

X Ignoring negative comments	 Generally, I ignore them or respond sarcastically; my replies usually get more likes than their comments, so they end up losing. (Q4-INFLU2) I try to ignore negative comments. If I respond, it gives them importance, which I do not want. (Q4-INFLU5) 	
र्इ्ड् Replying with sarcasm or humor	 I have never insulted anyone unless they insulted me first, in which case I might use sarcasm. (Q5-INFLU6) I ignore negative comments or respond with humor, which bothers haters more. If I see a mean comment, I might respond sarcastically. (Q4-INFLU4) 	
Se Blocking the user, filtering words, and reporting comments	 If I see something negative, I block the user. Instagram's word filter option helps a lot. (Q5-INFLU1) When I notice this, I first assess the type of comments. If they are trivial, I respond and move on, but if they cross a line, I thank the user I agree with and block the other. (Q8-INFLU1) If it is not serious, I ignore it; if it is disrespectful, I delete the comment thread. (Q8-INFLU5) 	
Supporting diversity of opinions through content or feedback	 When responding to hate, I try to keep calm and use it as a teaching moment, demonstrating that compassionate responses are possible and helpful for people spreading hate. (Q5-INFLU3) If someone crosses the line, I comment myself, usually getting more likes and signaling that hate is not welcome on my profile. (Q8-INFLU2) Supporting and contributing positively is important. I follow others for knowledge, humor, entertainment, etc. As long as they keep engaging me, I support them, and when they no longer do, I simply unfollow. (Q10-INFLU2) 	

Figure 4. Strategies adopted by influencers for managing hatedom

Source: Own elaboration.

Despite these measures, negative comments promote a widespread opinion among influencers about the out-group, labeling them as ignorant and lacking empathy because they do not recognize the potential impact of their hate messages on the community and their own lives. Influencers engage in (4) community selfcare routines that complement those identified in previous studies (Ouvrein et al., 2023; Valenzuela-García et al., 2023). Through these routines, they claim that,

I give space to a person who is hurting to heal. It is obvious that they are the ones filled with hate, not me. I am just peacefully in my life (Q6-INFLU4).

I encourage debate and critical thinking, and lead by example. I have never insulted anyone (Q4-INFLU5).

Conflicts between followers and communities due to differing opinions are another reality among influencers, referred to as "camp battles" (Q9-INFLU3). These confrontations can arise for various reasons, such as different views on content, physical appearance, sexual orientation, loyalty to other profiles or misunderstandings. In these cases, influencers used mediation and diplomacy by (a) supporting different sensitivities and arguments with respect, (b) directly addressing hostile reactions to convey their opinion and diminish the importance of the topic, or (c) deleting the comment thread that exacerbates the discourse, as Ouvrin and colleagues (2023) and Martínez-Valerio and Mayagoitia-Soria (2021) note. An influencer explains that,

Time also taught me that it is not worth trying to convince anyone online. No one knows each other, and very few people are willing to change their minds and truly listen (Q9-INFLU3).

This work of community management is thus integrated into the daily activities of influencers (Abidin, 2019; Ouvrein et al., 2023; Valenzuela-García et al., 2023) and is based on blind hatred (Aburime, 2022). However, the overwhelming amount of comments led them to draw natural boundaries between fandom and hatedom communities.

Boundaries around expectations: balancing the intimate and the labor

Whether public or private, the messages influencers receive, especially within larger communities, make it difficult to communicate with users. These influencers reported that the fandom community has exceptionally high expectations and demands quick and constant interactions. This often leads to "feelings of guilt", as one interviewee expressed (Q7-INFLU5). They also emphasized that this dynamic can easily shift fans into the hater-out group, leading to even more aggressive and obsessive hatred. For example, some influencers explained that,

Sometimes, you try to be nice to your followers and respond to them, but they end up taking too many liberties and even get angry (Q8-INFLU1).

Some people have a hard time understanding that we are not friends. I have approximately 20 followers who talk to me every day and are regular at my shows. An 18-year-old girl has been attending my show every month since November. It is the same show that she has seen it about 20 times, and she keeps returning. She has even brought me gifts. I have tried talking to her and her mother, and I feel guilty (Q8-INFLU2).

I had a partner and I used to post a lot about us on social media. When we broke up, everyone began asking what happened and why we split. Some even messaged my new partner to get information. This affected me, but it was really my fault not to set boundaries. If I had not shared much of my private life, it would not have happened (Q8-INFLU5).

These experiences show how important it is to set clear boundaries for influencers' private lives to avoid emotional exhaustion or invasion of privacy. Some influencers reported using complementary content creation strategies to

mitigate fans' sensitive transition from affection to hate. These strategies include: (1) specific posts about their personal lives, such as Q&A content with fans where they respond specifically to questions; and (2) temporarily disconnecting from social media by establishing access routines that do not exceed the time they devote to their personal lives. The main goal, as they state, is "not to feel compelled to show every aspect of your life, but only what you feel comfortable sharing" (Q8-INFLU6).

The responsibility of the influencer: battle of influences

Influencers play a significant role in the lives of their fans and therefore bear a great responsibility for discourse and values. They are aware of the influence they exert, which can significantly affect the attitudes, behavior and decisions of their community (Jenkins, 1992, 2006). One of them explicitly noted that,

I think we need to be careful about what we show, promote, and say. We have much visibility, and I believe we need to be responsible for what we post (Q2-INFLU5).

The responsibility of influencers goes beyond the mere creation of attractive or promotional content (von-Mettenheim & Wiedmann, 2022), it also includes the promotion of values that define them as human beings. This expression of interests, opinions and lifestyles creates cognitive dissonance that often leads to the loss of followers. For example, some interviewees stated that,

I lost followers from the vegan community because I did a campaign with El Pozo, but it was a pretty silent departure. They just unfollowed me, and a few expressed their discontent but in an appropriate way (Q10-INFLU2)

I stay true what I feel like doing at any given moment and do not always create tehe same content because I prefer to do what I enjoy [...]. This does not please all my followers, who sometimes want you to do what they want to see. As a result, many leave, but I do not worry about this. If someone leaves, it is because they need to (Q10-INFLU3).

The success of an influencer's labor depends on fostering a positive environment, which requires a relentless commitment to communication, for which, as they said, "there is no instruction manual" (Q2-INFLU3). A healthy community requires "nurturing valuable content" (Q2-INFLU2), despite the influence of algorithmic recommendation systems and fluctuating polarization on connections that influencers "have to deal with" (Q10-INFLU5), as the studies by Martínez-Valerio and Mayagoitia-Soria (2021) and Riemer and Peter (2021) show. As a result, influencers are under social pressure that blurs their 'influential' role (Barnes, 2022), whereby a highly demanding and obsessive fan base can become their greatest downfall, turning into cross-platform hatred.

CONCLUSIONS

Hatedom, which encompasses (anti-)fan communities, is characterized by persistent and intense criticism of public figures, especially influencers. This phenomenon goes beyond blind hatred (Duffy et al., 2022; Aburime, 2022) and represents a community-based relationship of obsessive engagement that reacts acutely to online content posted by online celebrities and is characterized by extreme emotional polarization (Barnes, 2022). This study sheds light on the dynamics of these polarized interactions between influencers and fans on different social media platforms and highlights the complexity and underlying mechanisms at play.

Although toxic fandoms and hatedoms share certain similarities, such as affective investment (Théberge, 2006) and intense reactions that blur the boundary between reality and fiction (Vizcaíno-Verdú et al., 2020), their dynamics and manifestations differ considerably. Toxic fandoms emerge within a specific fanbase where members may engage in harmful behaviors such as gatekeeping, criticism, and power struggles despite having common interests (Walther, 2022; Arouh, 2020).

In contrast, hatedoms focus their collective negative energy outward, primarily targeting public figures or influencers with whom they form intense parasocial relationships, and establish a kind of one-sided pseudo-friendship (Hoffner & Bond, 2022). These communities unite to share their disdain or hostility towards an influencer who they believe has gained undeserved fame (Ouvrein et al., 2023). Consequently, cross-platform stalking of fans to gain access to intimate life details transforms admiration into obsessive toxicity and hatred. This behavior is not only aimed at criticizing or harming, but also at breaking the boundaries that influencers try to maintain for their well-being and privacy. This means that engaging with content across multiple platforms to cultivate a broader, demographically diverse community often leads to the creation of a toxic macro-community. This broader engagement exacerbates the intensity of interactions and makes it harder for influencers to control the negative effects.

In relation to our first objective (O1), which focuses on identifying crossplatform polarity and toxicity in the comments received by influencers, our results show significant differences depending on the influencer's profile and platform. In particular, X and YouTube showed a higher percentage of negative and toxic comments, while TikTok, Facebook and Twitch fostered a more positive and collaborative environment. In contrast, on Instagram, there are discrepancies depending on the communication features used to interact with fans (chat rooms, bidirectional stories, public posts, etc.). This variance highlights the site-specific nature of platform feedback systems (Walther, 2022) to enable hatedom and emphasizes the need to review and implement targeted moderation strategies. Such differences suggest that platform architecture and community rules (Luo & Li, 2024) influence the expression of hate and toxicity, necessitating a nuanced understanding of each platform's unique demographics to develop effective interventions.

Regarding the second objective (O2), the role and relevance of fan communities for influencers' labor, our findings suggest that these communities profoundly impact influencers' personal and professional decisions (Abidin, 2019; Valenzuela-García et al., 2023). Certainly, fans provide validation and emotional support and represent a financial base for brands with which influencers collaborate. However, the affective investment in fandom (Théberge, 2006) seems to evolve into an affective aversion to influencers' daily activities. This aversion leads them to use coping strategies to mitigate negative behaviors and conflicts (Ouvrein et al., 2023). This duality of affection points to the complicated emotional labor that influencers perform by navigating a landscape where there is a delicate balance between admiration and hostility.

In relation to our third objective (O3), the way influencers engage with their fan bases, hatedom puts pressure on these individuals to prioritize certain platforms or activities in order to satisfy their fan base. Influencers' decisions are shaped by a pattern of hatedom unique to their profile, which is influenced by the age and maturity of the audience, the level of intimacy, the capabilities of the platform in question and the degree of anonymity or identity transparency of the (anti)fans involved. Consequently, influencers use different strategies to deal with online hate and toxicity (Valenzuela-García et al., 2023). They range from ignoring negative comments, responding with humor or sarcasm, to blocking and reporting offending users in some cases. Mediation and diplomacy play an important role in resolving conflicts between fans.

Indeed, we found that influencers' strategies are strongly influenced by the polarity and toxicity of the comments on their profiles. For example, influencers who received more negative and toxic comments, such as INFLU2 and INFLU6, took more proactive measures such as blocking users and using sarcasm to deflect hostility. In contrast, influencers with a predominance of neutral or positive comments, such as INFLU1, focused on promoting respectful engagement and supporting a community. This suggests that the type of feedback they receive significantly influences their approach to managing their online presence and interactions. Nevertheless, influencers need to set clear boundaries to protect their emotional wellbeing and maintain a healthy balance between their personal life and their social media presence.

In relation to our fourth aim (O4), the strategies influencers use to deal with hatedom on social media, the psychological impact of this phenomenon is consistent with previous analyzes of online harassment, including anxiety (Lehto, 2022), stress (Smith et al., 2017), and depression (Lind & Wickström, 2024). Influencers are not only seen as non-ideal victims of this hate (Valenzuela-García et al., 2023), but are also deeply dehumanized and expected to respond to the demanding and uncontrollable cross-platform hatedom. This love-hate relationship persists even within their own four walls and remains constant, viral and public.

LIMITATIONS AND FUTURE RESEARCH LINES

The study presented some limitations. The sample was limited to six Spanish influencers, which may not represent experiences in different regions or cultural contexts. In addition, the comments were collected in May 2024, which may not have taken into account temporal variations in the polarity and toxicity of the comments or improvements in the moderation of the platform. In terms of methodology, advanced natural language processing models used for quantitative analysis of comments cannot fully capture contextual and cultural nuances. Structured interviews, while qualitatively valuable, were constrained by influencers' available time, potentially limiting the depth of investigation.

Despite these limitations, our findings suggest new lines of research and recommendations to further explore and understand the dynamics of hatedom and improve the online community experience on social media platforms. For example, improved toxicity detection algorithms and content moderation can effectively manage negative and toxic comments. Longitudinal studies could provide insights into the evolving dynamics between influencers and hatedom over time. This should take into account temporal and contextual variations and include a diverse sample of influencers from different regions, cultures, content topics and gender identities. These approaches would provide a more representative understanding of the challenges faced by online celebrities and thus benefit future content creators and platform policy.

Tackling the hatedom phenomenon therefore requires a multi-layered approach that provides a deeper understanding of the polarization of digital fan communities towards influencers. This strategy can improve platforms' technological systems to moderate online hate more effectively, popularizing the term "mental health over toxic". It is imperative that influencers who demonstrate resilience, assertiveness, humanity and humility in the face of community conflict are not overshadowed by the rapid, automated and volatile evolution of social media and destroyed as empty, non-victimized beings. Such a holistic approach promises not only to protect the well-being of influencers, but also to cultivate a more respectful and supportive digital fandom.

FUNDING

This study is supported by the COYSODI research group of the Universidad Internacional de la Rioja and the project *La incidencia del marketing de influencers en la construcción de la identidad digital de los adolescentes (INFLUIDENTITY)* (The impact of influencer marketing on the construction of teenagers' digital identity (INFLUIDENTITY)) funded by UNIR. The time spent on this research is also included in the studies of the TikTok Cultures Research Network and the IERLab (Influencer Ethnography Research Lab) of Curtin University (Australia).

REFERENCES

- Abidin, C. (2018). Internet celebrity: Understanding fame online. Emerald Publishing Limited. https://doi.org/10.1108/9781787560765
- Abidin, C. (2019). Victim, rival, bully: Influencers' narrative cultures around cyberbullying. In H. Vandebosch & L. Green (Eds.), Narratives in research and interventions on cyberbullying among young people (pp. 199-212). Springer. <u>https://doi.org/10.1007/978-3-030-04960-7_13</u>
- Aburime, S. (2022). Hate narratives, conditioned language and networked harassment: A new breed of anti-shipper and anti-fan–antis. *Journal of Fandom Studies*, 10(2-3), 135-155. https://doi.org/10.1386/jfs_00060_1
- Association of Internet Researchers. (2012). Ethical decision-making and internet research. Recommendations from the AoIR Ethics Working Committee (version 2.0). https://aoir.org/reports/ethics2.pdf
- Arouh, M. (2020). Toxic fans: Distinctions and ambivalence. Ex-centric Narratives: Journal of Anglophone Literature, Culture and Media, (4), 67-82. https://ejournals.lib.auth.gr/ExCentric/article/view/7917
- Barnes, R. (2022). Loving to hate: Fandom fueling polarized behavior. In R. Barnes (Ed.), Fandom and polarization in online political discussion (pp. 61-86). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-14039-6_4
- Creswell, J. W. (2013). Qualitative inquiry and research design: Choosing among five approaches. SAGE.
- Devlin, J. & Chang, M. W. (2018, November 2). Open sourcing BERT: State-of-the-art pretraining for natural language processing. *Google Research*. <u>https://research.google/blog/</u> open-sourcing-bert-state-of-the-art-pre-training-for-natural-language-processing/
- Díaz, J. E., Cortés, M., & Hinojosa, C. (2019). Detecting toxic interactions on the Twitter Star Wars fandom using machine learning. Tecnológico de Monterrey.

- Duffy, B. R., Miltner, K. M., & Wahlstedt, A. (2022). Policing "fake" femininity: Authenticity, accountability, and influencer antifandom. *New Media & Society*, 247, 1657-1676. https://doi.org/10.1177/14614448221099234
- Hoffner, C. A. & Bond, B. J. (2022). Parasocial relationships, social media, & well-being. *Current Opinion in Psychology*, 45, 101306. <u>https://doi.org/10.1016/j.copsyc.2022.101306</u>
- Jenkins, H. (1992). Textual Poachers: Television fans and participatory cuture. Routledge
- Jenkins, H. (2006). Fans, bloggers, and gamers. Exploring participatory culture. New York University Press.
- Lehto, M. (2022). Ambivalent influencers: Feeling rules and the affective practice of anxiety in social media influencer work. *European Journal of Cultural Studies*, 25(1), 201-216. https://doi.org/10.1177/1367549421988958
- Lind, J. & Wickström, A. (2024). Representations of mental health and mental health problems in content published by female social media influencers. *International Journal of Cultural Studies*, 27(2), 217-233. https://doi.org/10.1177/13678779231210583
- Luo, Z. & Li, M. (2024). Participatory censorship: How online fandom community facilitates authoritarian rule. New Media & Society, 26(7), 4236-4254. https://doi.org/10.1177/14614448221113923
- Maarouf, A., Pröllochs, N., & Feuerriegel, S. (2024). The virality of hate speech on social media. In J. Nichols (Ed.), Proceedings of the ACM on Human-Computer Interaction (Vol. 8) (pp. 1-22). https://doi.org/10.1145/3641025
- Martínez-Valerio, L. & Mayagoitia-Soria, A. M. (2021). Influencers y mensajes de odio: jóvenes y consumo de contenidos autocensurados. Revista Prisma Social (Influencers and hate messages: Young people and consumption of self-censored content). Prisma Social, (34), 4-39. https://hdl.handle.net/10115/30379
- Mysiuk, I. & Shuvar, R. (2023). Automating web scraping of user comments for sentiment analysis in social networks. In 2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT) (pp. 77-81). IEEE. <u>https://doi.org/10.1109/ELIT61488.2023.10310867</u>
- OHCHR. (2023). Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia, and related intolerance on combating glorification of Nazism, neo-Nazism, and other practices that contribute to fueling contemporary forms of racism, racial discrimination, xenophobia, and related intolerance. United Nations.
- Ouvrein, G., Jorge, A., Cabral, J., & Vandebosch, H. (2023). Coping comes with the job: An exploratory study into the selection and use of coping strategies for online aggression among social media influencers. *Telematics and Informatics Reports, 10,* 100052. https://doi.org/10.1016/j.teler.2023.100052
- Riemer, K. & Peter, S. (2021). Algorithmic audiencing: Why we need to rethink free speech on social media. *Journal of Information Technology*, 36(4), 409-426. <u>https://doi.org/10.1177/02683962211013358</u>
- Schuessler, J. (2018, November 14). 'Toxic' is Oxford's word of the year. No, we're not gaslighting you. *The New York Times*. https://bit.ly/3WdwxGh

- Senft, T. M. (2013). Microcelebrity and the branded self. In J. Hartley, J. Burgess, & A. Bruns (Eds.), A companion to new media dynamics (pp. 346-354). Blackwell Publishing Ltd. https://doi.org/10.1002/9781118321607.ch22
- Sjoerd-Jeroen, M., Basten, F., Taran, G., Panagoulia, A., Coughlan, G., & Duarte, J. (2024). The structured narrative interview. *Narrative Inquiry*, 34(2), 307-334. <u>https://doi.org/10.1075/ni.23079.moe</u>
- Smith, M. M., Speth, T. A., Sherry, S. B., Saklofske, D. H., Stewart, S. H., & Glowacka, M. (2017). Is socially prescribed perfectionism veridical? A new take of the stressfulness of perfectionism. *Personality and Individual Differences*, 110, 115-118. <u>https://doi.org/10.1016/j.paid.2017.01.031</u>
- Strauss, A. & Corbin, J. M. (1997). Grounded theory in practice. SAGE.
- Suler, J. (2006). The online disinhibition effect. International Journal of Applied Psychoanalytic Studies, 2(2), 184-188. <u>https://doi.org/10.1002/aps.42</u>
- Théberge, P. (2006). Everyday fandom: Fan clubs, blogging, and the quotidian rhythms of the internet. *Canadian Journal of Communication*, 30(4), 485-502. https://doi.org/10.22230/cjc.2005v30n4a1673
- Valenzuela-García, N., Maldonado-Guzmán, D. J., García-Pérez, A., & Del-Real, C. (2023). Too lucky to be a victim? An exploratory study of online harassment and hate messages faced by social media influencers. *European Journal on Criminal Policy and Research*, 29, 397-421. https://doi.org/10.1007/s10610-023-09542-0
- Vizcaíno-Verdú, A., Contreras-Pulido, P., & Guzmán-Franco, M. D. (2020). Construcción del concepto fanbullying: Revisión crítica del acoso en redes sociales (Building the fanbullying concept: Critical review of social media harassment). *Pixel-Bit. Revista de Medios y Educación*, (57), 211-230. https://doi.org/10.12795/pixelbit.2020.i57.09
- Von-Mettenheim, W. & Wiedmann, K. P. (2022). Hate speech and bad language: The ugly face of social influencers and its impact on brands. An abstract. In B. Jochims & J. Allen (Eds.), Optimistic marketing in challenging times: Serving ever-shifting customer needs (pp. 291-292). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-24687-6_116
- Walther, J. B. (2022). Social media and online hate. *Current Opinion in Psychology*, 45. 101298. https://doi.org/10.1016/j.copsyc.2021.12.010

ABOUT THE AUTHORS

ARANTXA VIZCAÍNO-VERDÚ, associate professor at the Universidad Internacional de la Rioja, Spain. She holds a Ph.D. in Communication from the Universidad de Huelva. She currently integrates several research groups and networks in Spain and Australia analyzing popular culture on social media platforms. Specializing in media and information literacy, her study focuses on user-generated content on the Internet, fandom, and transmedia storytelling from the perspectives of communication, advertising, education, and digital humanities.

(D) https://orcid.org/0000-0001-9399-2077

PALOMA CONTRERAS-PULIDO, journalist, social educator and Ph.D. in Education from the Universidad de Huelva. She is associate professor at the Universidad Internacional de la Rioja, specializing in educational technologies and didactics, and guest lecturer in masters courses on audiovisual communication and TV production. She currently supervises doctoral dissertations on media, technology, education, and social inclusion. Member of Alfamed and researcher in the COYSODI Group (UNIR), contributing to media literacy and social inclusion studies.

iD https://orcid.org/0000-0002-6206-7820