cogitatio

POLITICS AND GOVERNANCE

Volume 11

Issue 2

2023

Hate Speech, Demonization, Polarization, and Political Social Responsibility

Edited by Luis M. Romero-Rodríguez, Pedro Cuesta-Valiño, and Bárbara Castillo-Abdul

Open Access Journal

ISSN: 2183-2463

Politics and Governance, 2023, Volume 11, Issue 2
Hate Speech, Demonization, Polarization, and Political Social Responsibility

Published by Cogitatio Press
Rua Fialho de Almeida 14, 2º Esq.,
1070-129 Lisbon
Portugal

Design by Typografia®
http://www.typografia.pt/en/

Cover image: © Mateusz Kropiwnicki from iStock

Available online at: www.cogitatiopress.com/politicsandgovernance

Table of Contents

Table of Contents

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 109–113
https://doi.org/10.17645/pag.v11i2.6663

Editorial

The Process of the Transfer of Hate Speech to Demonization and Social Polarization

Luis M. Romero-Rodríguez [1,2,*], Bárbara Castillo-Abdul [1], and Pedro Cuesta-Valiño [3]

[1] Department of Communication Sciences and Sociology, Rey Juan Carlos University, Spain
[2] ESAI Business School, Universidad de Especialidades Espíritu Santo, Ecuador
[3] Department of Economics and Business Administration, University of Alcalá, Spain

* Corresponding author (luis.romero@urjc.es)

Submitted: 2 January 2023 | Published: 17 May 2023

Abstract
We are living through a time of major political changes due to the rise of populist leaders and the resurgence of extreme ideological movements. The emergence of this phenomenon is due, to a large extent, to the ease with which these political actors can disseminate and spread their messages without any limits through social networks, leaving aside the former "fourth power" of the media as filterers and reinterpreters of information. Generally, the formula used by these leaders and movements is usually based on symbolic social division and polarization through hate speeches that allow demonizing their adversaries while antagonizing the issuers: a discursive "us" against "them" based on verbal violence to dehumanize an "exogroup." We want to discuss the importance of understanding the process of communicational transfer—which begins with hate speech and evolves into demonization and social polarization—as a strategic basis for creating an ideal scenario for the growth and strengthening of populist discourse, which is reductionist and simplifying in nature.

Keywords
hate discourse; political communication; political responsibility; social media

Issue
This editorial is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

Discussions around prejudice and rejection concerning an "exogroup" are extensive and long-standing in the social sciences (e.g., Brewer, 1999; Peherson et al., 2011). There is an intense connection between the use of language and the spread of prejudice toward "the other" (Maass et al., 1989). However, contrary to the paradigm of the intergroup linguistic bias theory hypothesis (Gorham, 2006; Whitley et al., 2016), in the transmission of prejudice through social networks, the most recent empirical evidence shows that descriptions of an exogroup cease to be vague or abstract to become specific, observable, and measurable when they are sheltered in official discourses (Crandall et al., 2018) or networks that allow one's identity to remain hidden (Arcila-Calderón et al., 2020, 2022; Fox et al., 2015).

Throughout history, there are many examples of how hate speech has been used from the discursive podium to demonize and astutely devalue different social groups and minorities. Verbal violence has served to dehumanize "otherness" and eliminate or diminish any empathy that may exist towards a group of "outsiders," symbolically stripping them of their humanity and rights and creating a scenario of forced polarization: an "us" versus "them" narrative that has even historically caused and justified genocide.

In this process, two phenomena implicitly occur: First, the demonization of the "other" becomes necessary to understand it as an enemy, to generate hatred towards

otherness, and place the other at the margins of the interests of the majority. The second phenomenon is the sacralization of the "we," unifying and uniting the heterogeneous group against the "symbolic enemy," giving a halo of heroism to the cause. This generates a "closing of ranks," a reification of "the other," and contributes to the feeling of exceptionality in attitudes and measures against the other. This has been the logic of the dehumanization process of almost all wars and the current basis of populist discourse, regardless of its ideology.

In this sense, populism, polarization, demonization, and hate speech are socio-political phenomena closely connected. They are based on conceptual simplifications that offer society a series of communicatively effective fallacies of causality, since simplistic dichotomies like bad/good, protagonist/antagonist, the people/enemies—without grey areas—have never ceased to be effective in the popular mobilization of feelings, especially the most effective ones of anger and fear.

Not surprisingly, recent research (e.g., Garzia & Ferreira Da Silva, 2022; Nai et al., 2022) has shown that affective-emotional polarization generates a stronger mobilization against a candidate (moved by rejection) rather than in their favor (moved by approval). This is a phenomenon called "negative voting," which is understood as a rational choice, a product of cognitive dissonance and/or retrospective evaluations. Voters may be reasoning less in favor of the best political alternative, instead seeking to avoid the one they deem to be worse. The more voters are flooded by predominantly negative information about some parties, candidates, and voters rather than others, the clearer it becomes that this information, which is readily accessible, "manipulates" their vote; when in doubt as to whom someone should vote for, they are triggered by negative information and moved by negative voting.

With the rise and popularization of the internet and social networks, since the beginning of this century there has been a process of disintermediation, or apomediation, in communication that makes it impossible to limit the spaces for the dissemination of this type of discourse, a role once played by the media as information filterers. This reticular and decentralized digital ecosystem has provided the perfect breeding ground for maximizing the visibility and scope of polarizing and demonizing arguments, which is why the emergence of populist movements, the growth of political disaffection and anti-politics, and the rise of intolerance towards those who do not think the same way are not trivial.

2. Hate Speech and Demonization: The Genesis of Social Polarization

Hate speech involves the promotion of messages that encourage the rejection, disparagement, humiliation, harassment, discrediting, and stigmatization of individuals or social groups based on very diverse attributes such as nationality, ideology, social class, race, creed, gender, or sexual orientation (Civila et al., 2020). In a context in which digital media and social networks facilitate the fast creation and dissemination of these messages, the relevance of hate speech derives, above all, from its role as a trigger for hate crimes. As several researchers suggest, there is a significant relationship between hate speech and physical violence (e.g., Muller & Schwarz, 2018).

Demonization, on the other hand, consists of the process by which the ideas and values of the orator (the source of the hate speech) are made sacred with dialectic and discursive resources, promoting a symbolic construction of reality based on the conceptual simplification of protagonist/antagonist where the "antagonist other" cannot be culturally accepted and is honestly inferior or inconsiderate. This "antagonist other" encapsulates hate paradigms, discriminations, and stereotypes that devalue it morally and perceptively, damaging its social identity and even its self-recognized identity (Goffman, 1963).

Through antagonization, demonization leads to an attitude of distrust against otherness among the public opinion. By turning "outsider" groups into morally inferior groups, criminalizing their opinions and all the while distracting and polarizing society, moral restrictions among the public are eliminated. On the other hand, the ideas of the orator (a sort of self-appointed representative of the majority) are exposed as correct and justified (Romero-Rodríguez & Römer-Pieretti, 2016).

The discursive process of demonization is composed in the following way (Romero-Rodríguez et al., 2015): The "demonizing" orator explores the interests of the audience, relating to them and gaining their empathy; they confront the interests of the "antagonist group," building on disagreement points between the latter and their target audience; they defend the interests of their audience and accuse the "antagonist group" of going against them; they use disqualifying adjectives, usually short, blunt, and easy to memorize, and repeat them frequently in their speech—repetition is the key to persuasion and "institutionalization" of realities; finally, they take their attacks to a radical extreme to mark out and stigmatize their opponent (Figure 1).

Continuous exposure to demonization generates polarization, which, from a sociological perspective, is a phenomenon that appears when individuals align their beliefs in extreme and conflicting positions while other individuals hold more moderate or neutral opinions (Isenberg, 1986; Sunstein, 2002).

Symbolic power, as it has a great scope of diffusion, increases the capacity of an individual or institution to construct realities (Searle, 1995; Watzlawick, 1976) and even imposes order on these constructed realities (Bourdieu, 1989). When an individual or institution vested with symbolic power uses rhetoric to construct a "should be unique," adequate or pertinent on pain of being illegitimate, it eliminates any possibility of alternative ways of thinking, crystallizes a stereotype typified as normal—even sacralized—while institutionalizing,

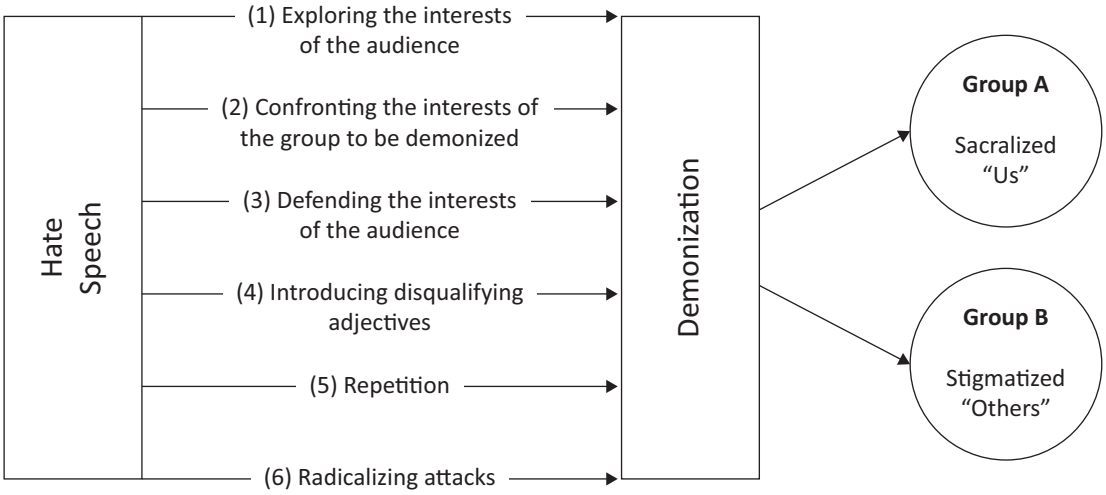


Figure 1. Process of communicational transfer of hate speech–demonization–polarization.

through discursive violence, the "other," disqualifying, trivializing, and rendering its social needs invisible. Society is divided into at least two opposing symbolic sectors, while the people's real problems are pushed into the background by the constant distraction allowed by the mediatized confrontation.

The conflict of polarization is rarely evaded, even if one does not want to enter into the discursive issues of political struggle (Prada Espinel & Romero-Rodríguez, 2018). One's positioning in a polarized scenario is compelled by social pressure—due to the so-called "spiral of silence" (Noelle-Neumann, 1993) or the "bandwagon" effect (Goidel & Shields, 1994)—to adapt to the group and reduce tensions. In this way, away from all reason and focused only on emotional management, it is in the fragmentation of the social fabric that common sense breaks down, and automatic solidarity appears, even with those leaders who operate against the most elementary human rights.

3. Political Social Responsibility: A Pending Matter for Stakeholders

Political actors, such as parties, candidates, citizen movements and activists, social media, among others, have a pending and unpostponable task on behalf of society: to lessen discursive confrontation, seek moderation, and return to politics as a means to solve people's problems.

No one can deny that polarization and demonization have become very useful and effective stratagems to distract society from the problems that affect them and mobilize the vote out of fear, anger, or extreme rejection. Therefore, the current situation of discursive confrontations in many Western countries such as the United States, the United Kingdom, Spain, Argentina, Brazil, Chile, Colombia, Peru, and Venezuela, among others, have very identifiable responsible parties who should be exhorted to moderation and ethics in the exercise of politics and social conciliation.

On the one hand, through legislative efforts, states must establish mandatory regulations to consider hate speech as a criminal offense while regulating and reducing as far as possible the demonizing and polarizing messages that political parties, electoral movements, activists, or other people with broad social reach in the media and social networks share. These efforts in creating laws that regulate hate speech must, of course, take care of the limits of freedom of expression, in the understanding that not every criticism against a group (ideological, economic, etc.) should be classified as a hate crime, so it is essential to be very accurate in these typifications to avoid that these laws can become a dictatorship of political correctness.

On the other hand, based on the premise that any information that society consumes has a political or economic interest behind it and that many political actors seek to divide society for electoral mobilization or mere emotional distraction, education should provide citizens with media and digital skills that will enable them to acquire critical and reflective analysis skills when consuming informative and opinionated content. Media education should be mandatory in primary and secondary schools since citizens are not usually taught how to consume media and messages (Pérez-Rodríguez et al., 2019; Romero-Rodríguez et al., 2021).

Finally, but no less importantly, political parties should establish programs for the ethical training of their bases and youth, as well as internal disciplinary measures against the use of hate speeches and demonizing or polarizing messages, both to their militants and their leaders. Understanding that the responsibility (political and legal) of political parties, as institutions for social organization, is to promote coexistence and not hatred is fundamental. Unfortunately, in many countries, hate speech, demonization, and polarization are becoming the modus operandi of campaigning, either because of the ease of construction of reductionist messages or because this type of content generates fear and anger

in audiences, which facilitates the negative vote or rejection vote.

In this thematic issue, we wanted to open a space for plural debate on the triangulation between hate speech–demonization–polarization and populism, so that researchers from around the world could show the path of these phenomena in various scenarios such as racism, islamophobia, russophobia, hate speech against LGTBI groups, the rejection of economic elites, the nationalist discourse against immigration, the discourses on Europeanization, and the social responsibility of journalism in the face of social confrontation.

Many spaces and social phenomena are currently continuously demonized by hate speeches, so any research intention from the social sciences can be reductionist and simplifying in itself. However, researchers must focus on building spaces for debate and reflection on social polarization to be agents of change toward harmony and citizen coexistence.

Conflict of Interests

The authors declare no conflict of interest.

References

Arcila-Calderón, C., Blanco-Herrero, D., & Apolo, M. B. V. (2020). Rejection and hate speech in Twitter: Content analysis of tweets about migrants and refugees in Spanish. *Revista Española de Investigaciones Sociológicas*, *172*, 21–40. https://doi.org/10.5477/cis/reis.172.21

Arcila-Calderón, C., Sánchez-Holgado, P., Quintana-Moreno, C., Amores, J. J., & Blanco-Herrero, D. (2022). Hate speech and social acceptance of migrants in Europe: Analysis of tweets with geolocation. *Comunicar*, *30*(71), 21–34. https://doi.org/10.3916/C71-2022-02

Bourdieu, P. (1989). Social space and symbolic power. *Sociological Theory*, *7*(1), 14–25. https://doi.org/10.2307/202060

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, *55*(3), 429–444. https://doi.org/10.1111/0022-4537.00126

Civila, S., Romero-Rodríguez, L. M., & Civila, A. (2020). The demonization of islam through social media: A case study of #StopIslam in Instagram. *Publications*, *8*(52). https://doi.org/10.3390/publications8040052

Crandall, C. S., Miller, J. M., & White, M. H. (2018). Changing norms following the 2016 US presidential election: The Trump effect on prejudice. *Social Psychological and Personality Science*, *9*(2), 186–192. https://doi.org/10.1177/1948550617750735

Fox, J., Cruz, C., & Lee, J. Y. (2015). Perpetuating online sexism offline: Anonymity, interactivity, and the effects of sexist hashtags on social media. *Computers in Human Behavior*, *52*, 436–442. https://doi.org/10.1016/j.chb.2015.06.024

Garzia, D., & Ferreira Da Silva, F. (2022). The electoral consequences of affective polarization? Negative voting in the 2020 US presidential election. *American Politics Research*, *50*(3), 303–311. https://doi.org/10.1177/1532673X221074633

Goffman, E. (1963). *Stigma*. Penguin.

Goidel, R. K., & Shields, T. G. (1994). The vanishing marginals, the bandwagon, and the mass media. *The Journal of Politics*, *56*(3), 802–810. https://doi.org/10.2307/2132194

Gorham, B. W. (2006). News media's relationship with stereotyping: The linguistic intergroup bias in response to crime news. *Journal of Communication*, *56*(2), 289–308. https://doi.org/10.1111/j.1460-2466.2006.00020.x

Isenberg, D. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality Social Psychology*, *50*(6), 1141–1151. https://doi.org/10.1037/0022-3514.50.6.1141

Maass, A., Salvi, D., Arcuri, L., & Semin, G. R. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social psychology*, *57*(6), 981–993. https://doi.org/10.1037/0022-3514.57.6.981

Muller, K., & Schwarz, C. (2018). *Fanning the flames of hate: Social media and hate crime*. SSRN. https://doi.org/10.2139/ssrn.3082972

Nai, A., Garzia, D., Aaldering, L., Ferreira Da Silva, F., & Gatermann, K. (Eds.). (2022). Negative politics: Leader personality, negative campaigning, and the oppositional dynamics of contemporary politics [Thematic issue]. *Politics and Governance*, *10*(4).

Noelle-Neumann, E. (1993). *The spiral of silence: Our social skin*. University of Chicago Press.

Peherson, S., Brown, R., & Zagefka, H. (2011). When does national identification lead to the rejection of immigrants? Crosssectional and longitudinal evidence for the role of essentialist in-group definitions. *British Journal of Social Psychology*, *48*(1), 61–76. https://doi.org/10.1348/014466608X288827

Pérez-Rodríguez, M. A., Delgado-Ponce, A., Marín-Mateos, P., & Romero-Rodríguez, L. M. (2019). Media competence in spanish secondary school students. Assessing instrumental and critical thinking skills in digital contexts. *Educational Sciences: Theory and Practice*, *19*(3), 33–48. https://doi.org/10.12738/estp. 2019.3.003

Prada Espinel, O., & Romero-Rodríguez, L. M. (2018). Polarización y demonización en la campaña presidencial de Colombia de 2018: Análisis del comportamiento comunicacional en Twitter de Gustavo Petro e Iván Duque [Polarization and demonizacion in the 2018 Colombia's presidential campaign: Analyzing the communicational behaviour of Gustava Petro and Iván Duque on Twitter]. *Revista Humanidades*, *9*(1), 1–26. https://doi.org/10.15517/h.v9i1.35343

Romero-Rodríguez, L. M., Aguaded, I., & Gadea, W.

(2015). From demonization to polarisation: An analysis of the Venezuelan governments and the political opposition's digital discourse. *Argos*, *32*(62), 97–117.

Romero-Rodríguez, L. M., & Römer-Pieretti, M. (2016). Proceso de demonización de la oposición política en los hitos discursivos de Hugo Chávez según la prensa digital [Process of demonization of the political opposition in Hugo Chávez discourses according to the digital press]. *Temas de Comunicación*, *32*, 95–124.

Romero-Rodríguez, L. M., Tejedor, S., & Pabón Montealegre, M. V. (2021). Populist attitudes and perceptions of public opinion and the media: A comparative correlational study between Spain and Colombia. *Revista Latina de Comunicación Social*, *79*, 43–66. https://doi.org/10.4185/RLCS-2021-1507

Searle, J. R. (1995). *The construction of social reality*. The Free Press.

Sunstein, C. (2002). The law of group polarization. *Journal of Political Philosophy*, *10*(2), 175–195.

Watzlawick, P. (1976). *How real is real: Confusion, disinformation, communication*. Random House.

Whitley, J. R., Bernard, E., & Kite, M. E. (2016). *Psychology of prejudice and discrimination*. Routledge.

About the Authors

Luis M. Romero-Rodríguez is an associate professor in the Department of Communication Sciences and Sociology at the Universidad Rey Juan Carlos (Madrid, Spain) and visiting professor at ESAI Business School, Universidad Espíritu Santo (Guayaquil, Ecuador). He holds a PhD in communication, a master's degree in social communication and law and international politics, and a bachelor's degree in social communication and law. He has received several awards and recognitions, such as the Napolitan Victory Awards and the Roblón 2020 as an emerging researcher in communication in Spain.

Bárbara Castillo-Abdul is a research fellow in the Department of Communication Sciences and Sociology at the Universidad Rey Juan Carlos (Madrid, Spain) and is a doctoral candidate in the Interuniversity Doctoral Program in Communication at the Universities of Huelva, Seville, Málaga, and Cádiz, and in the Doctoral Program in Social and Legal Sciences at the Universidad Rey Juan Carlos. She holds a master's degree in communication and audiovisual education and a bachelor's degree in social communication.

Pedro Cuesta-Valiño holds a PhD in economics and business (2002) and is a professor of marketing at the Universidad de Alcalá (Spain), where he is currently the chair of the Department of Economics and Management. He is editor-in-chief of the international journal of internet marketing and advertising, associate editor of humanities & social sciences communication, founder of the International University Network of Happiness, and chair of the research group Consumer Behaviour, Organizational, and Market Analytics.

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 114–126
https://doi.org/10.17645/pag.v11i2.6299

Article

Social Media and Otherness: The Case of #Islamterrorism on TikTok

Sabina Civila *, Mónica Bonilla-del-Rio, and Ignacio Aguaded

Department of Education, Huelva University, Spain

* Corresponding author (sabi.civila@dedu.uhu.es)

Submitted: 21 October 2022 | Accepted: 26 December 2022 | Published: 17 May 2023

Abstract
Social media and their participatory characteristics promote the construction of meanings that differ from those emitted by mainstream media outlets, becoming a tool that enables a reconfiguration of the dominant discourses. TikTok offers unique possibilities to confront the neoliberal imaginary and open a space for debate, incorporating political viewpoints and establishing itself as a new communication scenario. Regarding news about jihadism, many researchers have observed that those who practice Islam are classified as a monolithic entity, and this entire religious group is generalized as a threat to modern societies. The main objective of our research is thus to know the discourses used on TikTok to respond to the binomial Islam = terrorism spread by mainstream media and the affordances of this platform used to challenge this misconception. Using the snowball method, a multimodal analysis was conducted by identifying TikTok videos with the hashtags #yihadista, #yihad, and #islamterrorismo (in its English and Spanish versions) to explore the uses of the TikTok platform. The resulting selection criteria included: (a) content related to mainstream media discourses on jihadism, (b) discussion of a topic related to Islam and terrorism, and (c) where the content creator declares him/herself to be a Muslim. In addition, in-depth interviews were conducted to provide an enhanced understanding of how the media promote the need to generate a counter-narrative on TikTok. The results reveal that discourses from Muslims that combat Islam = terrorism discourses are constructed within the spiral of the dominant narrative, thus visualizing the negative discourses about Islam.

Keywords
Islam; media; prosumers; religion; social media; TikTok

Issue
This article is part of the issue “Hate Speech, Demonization, Polarization, and Political Social Responsibility” edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

This article analyzes the discourses of the Muslim community in relation to the messages transmitted in mainstream media about Islam and its association with the concept of terrorism. Specifically, we evaluate in depth their use of TikTok as a platform to respond to these discourses and how they take advantage of the affordances offered by this platform to question them and share their viewpoints. In this sense, we understand mainstream media as conventional news sources (newspaper, television, radio) that “are trapped within entrenched norms and values that predispose a certain perspective of the world” (Jean-Kenix, 2011, p. 2). In addition, mainstream media are considered credible, serious, and influential, so they could have a strong effect on viewers (Gondwe & Bhowmik, 2022). We thus explore the potential of TikTok as a medium that favors expression and allows bidirectional discourses, thereby opening a debate on religious, political, cultural, and social issues. This goes beyond approaches where the audience simply consumes information passively from the main agents (political leaders and mainstream media), turning them into prosumers who generate content and participate in conversations through TikTok. Using a qualitative approach, we carry out a multimodal content

analysis of TikTok videos shared by Muslims that specifically address discussions on topics related to Islam, terrorism, and jihadism. We also conducted semistructured interviews with subject-expert journalists, thus providing an enhanced understanding of the phenomenon.

Currently, Islam is the first-most widely practiced religion in the world with approximately 1.6 billion followers (Statistics and Data, 2022), which means that one out of four people in the world is Muslim (Merino, 2020). The representation of this religion in the media can directly influence the public's perception of and attitudes toward Muslims as a group as well as Islam as a religion (Akbarzadeh & Smith, 2005). In this regard, previous studies have shown that the representation of Muslims in the media has an eminently negative focus. The Muslim community is often typecast as a terrorist group and as a threat to democratic systems and freedom of expression (Rahman, 2022).

Against this backdrop in which popular media culture has linked Islam to a negative representation in the West (Farooqui & Kaushik, 2022), social media have become a space that allows Muslims to counter such negative discourse regarding their religion and consolidate their identity and self-image (Deroo & Mohamud, 2022). Indeed, these platforms offer an opportunity for users to create their own content, which implies a greater diversity of media representations. Therefore, citizens have more information and viewpoints to better understand personal experiences and social events surrounding a given phenomenon. Given the narrative possibilities offered by social media, in this research we focus on TikTok, whose popularity has increased exponentially in recent years, becoming the most downloaded app worldwide (Sensor Tower, 2022).

We intend to understand how the Muslim community thrives on TikTok following two research questions:

> RQ1: What are the discourses used by the Muslim community to respond to the binomial Islam = terrorism on TikTok?
>
> RQ2: How are TikTok affordances used to challenge misconceptions spread by mainstream media?

Through this analysis, we contribute to understanding how the Muslim community uses TikTok to express themselves and to challenge misconceptions. According to Pearce et al. (2020), knowing how people express themselves on social media helps to design better social media communication experiences.

To address this approach the theoretical background includes discussion about the representation of Muslims in Western mainstream media, Islam on social media, and the TikTok affordances offered to challenge misconceptions. We then specified the methodology used and present the findings. The results enable us to identify the discourses used on TikTok to respond to the binomial Islam = terrorism. And, to understand the needs and concerns of the Muslim community about this phenomenon. Furthermore, it provides a wider view of how they thrive and represent themselves in social media spaces.

2. Theoretical Background

2.1. The Representation of Muslims in Western Mainstream Media

The relationship between the West and Islam has been categorized as conflictive by many researchers over the years (Canclini, 2004; Miles, 1989; Said, 1978). The expansion of the West results in the generation of an antagonistic dichotomy between the West and East, which maintains that the culture of Islam is not compatible with European values and is presented as a threat (Said, 1978). According to Miles (1989), there are two clear examples of this polarization: first, the definition of "others" as "blacks" and "savages," which results in situations of exclusion, since Westerners are considered white, logical, delicate, and virtuous (Donald, 1992); and secondly, culture, where the European representation of the Islamic world is associated with terrorism and barbarism. To achieve the aims of the current work, otherness is understood as a discursive process where the "other" is represented as a stranger (Iqbal, 2020), relating Islam to terrorism through the news frames of mainstream media.

According to Casteleiro-Ruiz-de-Azcarate (2015), the frames used to present terrorism-related news are full of cliches and simplifications, contributing to the identification of religion with terrorism and driving a state of ignorance about Islam and its nature (Villepin, 2003). Studies conducted during the last two decades have provided evidence that such news framing has focused on the representation of Islam as monolithic, sexist, homogeneous, fanatical, and terrorist (Civila et al., 2020; Ewart, 2012; Ibrahim, 2010; Korteweg, 2008). This results in demonization, stereotyping, and negative feelings toward the Muslim community (Akbarzadeh & Smith, 2005).

Along the same vein, in his book *Orientalism*, Edward Said argues that there is a particular discourse in Europe that promotes the differences between the familiar (Europe, the West, or "us") and the foreign (the East or the "others"; Said, 1978, p. 19). These discourses are promoted by the mainstream media, limiting Islam to a series of misconceptions that do not represent reality and that polarize societies (Poole, 2002). Brown and Levision (1987), as well as Lee and Pinker (2010), state that broadcasters use ambiguity, indirect discourses, and other strategies to inform, especially when the topic is about the "other," and even substantiate statements with personal experiences (Galasin'ska & Galasin'ski, 2003; Giglietto & Lee, 2017). However, most of the news is written from a non-Muslim perspective. Such coverage does not illustrate what Muslims consider about this phenomenon (Gabsi, 2015), but rather adopts a Western perspective.

When these negative discourses promoted by the media relate Islam and terrorism, they can be labeled as "Islamophobia" (Poole, 2002). The term has been discussed by many scholars but is too broad to provide a tight definition. According to Ross (2014), Islamophobia constitutes an understanding of Islam as a threat to Western societies and intolerance to those who practice this religion. Meanwhile, Acim (2019) defines Islamophobia as the representation of Muslims as "others.'' As well as the prohibition of Islam cultural elements in the West, such as the hijab in France or the construction of mosques in Switzerland.

Nowadays, social media offer a space for minorities to express themselves (Echchaibi, 2013), thus reflecting how Muslim people fight against the discourses issued by mainstream media, which are considered to be hegemonic. Following many studies demonstrating how Western media demonize Islam by relating it with terrorism (Ahmed & Matthes, 2017), we focus on determining how the Muslim community challenges such discourses on TikTok. Thus, we contribute to understanding how the Muslim community uses this platform to express themselves and challenge misconceptions spread by the media.

2.2. (De)Construction of Islam on Social Media

Social media have changed how we communicate by allowing people to create their own content. It is not only communication unidirectional, but users can participate in media debates and adopt positions in relation to their contexts, thus challenging dominant discourses (Jaramillo-Dent et al., 2022). The Internet and thereby social media offer the opportunity to represent the diverse identities that are present in the contemporary world as well as the autonomy to share beliefs and values with other users (Campbell & Evolvi, 2020). This helps minority groups express themselves, breaking misconceptions associated with their group and challenging hate speech (Gómez-García et al., 2021). According to Campbell and Evolvi (2020, p. 8), minority groups include those who are stigmatized by society, for example, Muslim people living in non-Muslim countries. Therefore, in the present article, Muslim people as represented in Western media are considered to belong to a minority group.

Minorities, in different contexts, are forced to challenge the discrimination they suffer from the dominant culture, which limits their opportunities to relate to and participate in some mainstream conversations. However, social media can be a tool to combat these threats and make their culture and identity visible (Abidin, 2021). We thus identify a generation of Muslim individuals in the West who are publicly committed to negotiating their religious values with modern societies and altering Western narratives by promoting the relationship between Islam and modernity (Echchaibi, 2013; Göle, 2002). Through this engagement, we observe profiles on blogs and social media whose mission is to challenge the representation of Islam in mainstream media, such as the Muslimah Media Watch or Yallah blog (Echchaibi, 2013; Evolvi, 2017). It is also remarkable the emergence of movements such as #Noenminombre (#Notinmyname in English) or #MuslimsareNOTterrorist following the waves of attacks that occurred in Europe during 2016–2018.

Studies related to "digital religion," a discipline that emerges from the combination of religion and Internet studies (Campbell, 2017), have examined how Islam is represented by Muslims on social media, claiming generally that they respond to dominant narratives and associations aimed at their collective (Echchaibi, 2013; Evolvi, 2017; Looy, 2015; Wheeler, 2014). This supports previous research finding that Muslim creators often create content to refute misconceptions about their religion (Mosemghvdlishvili & Jansz, 2012). Accordingly, some authors, such as El-Haj and Bonet (2011), question whether Muslim people represent themselves in social media exclusively in relation to controversies about their religion or, in contrast, for their own interests.

In this case, in line with Hoover and Echchaibi (2014), social media would act as a "third space," a term that arises in relation to evolving media discourses and that has been used in studies on digital religion. This third space refers to a hybrid space that offers new forms for the reconstruction, expression, and questioning of religion, as well as new possibilities for the construction of meanings: It describes an alternative space to the dominant ones. In this study, hybridization occurs between the discourses broadcasted by the mainstream media and those shared by TikTok users. This theory is thus applied to understand the discursive practices used on TikTok to challenge misconceptions about Islam, as previously applied to analyze aspects concerning Islam in Facebook groups (Al-Rawi, 2016; Illman & Sjö, 2015), memes (Aguilar et al., 2017), and other social media (Bahfen, 2018).

According to Mirra et al. (2018), young people today engage critically with the media and challenge misconceptions about Islam. Therefore, digital contexts can modify the perception of reality and generate conversations outside mainstream narratives. In this case, we focus on TikTok, the most downloaded app, with 188 million downloads during the first half of 2022 (Statista, 2022).

2.3. TikTok as a Tool to Challenge Mainstream Media Discourses

Since its inception, TikTok has been growing and has now become one of the most popular social media in the world, shaping engaged communities that respond to the app's affordances (Zhao & Wagner, 2022). In our research, we consider the affordances of TikTok as tools offered by this social media platform (audios, challenges, trends, hashtags, filters, virality, etc.) that allow users to

express themselves creatively (Kaye et al., 2021). TikTok confers power and opportunities for expression to users who acquire the role of the sender, which favors mass self-communication (Castells, 2013) and bi-directionality in the messages issued by conventional media. This social media platform is becoming a space for debate, incorporating a political viewpoint, and establishing itself as a new communication scenario (Cervi & Marín-Lladó, 2021). Users can thereby actively participate in social and political debates, taking advantage of TikTok's affordances and its particularities.

TikTok allows users to share short videos (although the option to create videos lasting up to ten minutes has been added in March 2022), which can be edited in the app itself adding filters, effects, or stickers. The innovations offered by this platform in terms of content distribution and discovery have established it as one of the most highly valued platforms. It includes recommendation algorithms aimed at the interests of the users themselves and its famous "For You" page which attracts and retains users (Zhao & Wagner, 2022). According to Vijay and Gekker (2021) and Literat et al. (2022), TikTok's affordances additionally favor content that responds to mainstream media and encourage users to interact with other videos (through options such as "green screen," "paste from," or video responses to comments from other accounts). Previous studies have shown that TikTok users employ the affordances offered by this platform to question media representations of "others" in the news, criticize the media framework, and expand narratives by offering new viewpoints (Literat et al., 2022). In this sense, TikTok becomes an environment for self-expression and sense-making of others and otherness (Schellewald, 2021).

This convergence between traditional and emerging media leads to a new media ecology in which communication is no longer the exclusive remit of classical actors (parties, political leaders, and the media); rather, social media break into the discussion, significantly influencing social mobilization and facilitating the democratization of dialog thanks to their interactivity and immediacy (Islas, 2015; Piñeiro-Otero & Martínez-Rolán, 2020). According to Şot (2022), users take advantage of TikTok's technological and practical features to create content that is tailored to their audience to generate an intimate and safe space for both them and their followers.

3. Method

3.1. Sample and Inclusion Criteria

The present analysis, conducted during May and September 2022, provides a case study with a qualitative design, mainly adopting two methods: (a) a multimodal content analysis that aims to explore the intersection between misconceptions of Islam in mainstream media and discourses on TikTok; and (b) interviews, a qualitative analysis technique with an exploratory and explanatory nature whose purpose is to understand the studied phenomenon in depth (Hernández-Sampieri et al., 2014).

For sample selection, a new TikTok profile was created, to avoid algorithmic contamination and obtain objective results. Once the account had been opened, we applied the snowball method, a qualitative data collection technique to reach difficult-to-access samples (Naderifar et al., 2017), and the sample was identified. To achieve this, the hashtags #yihadist, #yihad, and #islamterrorismo were searched for in the "explore" section of the app (in its English and Spanish versions), applying the filters "last 6 months" and ordering by "relevance." A total of 2,546 videos were then viewed and those that met the following criteria were selected: (a) content related to the discourses of mainstream media about jihadism, (b) discussion topics related to Islam and terrorism, and (c) where the content creator declared him/herself to be Muslim. The final sample comprised 62 videos. The files were downloaded by using the SnapTik app and collated in Excel using the video description and the numbers of likes, comments, and views. Finally, the data were analyzed using Atlas.ti 8, which allows qualitative analysis (Figure 1).

The data were examined utilizing a multimodal approach, which allows for the comprehension of both verbal and nonverbal data as well as the exploration of several representative modes (Dicks, 2019). TikTok's audiovisual, textual, interactive, and the way these modes are combined are all included in the multimodal analysis (Korhonen, 2010). Iterative analysis was used to categorize the data developing a codebook. In the first round, we analyzed the discourses spread by the Muslim community. Then, we related the affordances (audios, challenges, trends, hashtags, filters, virality, etc.) of TikTok with their discussion.

Moreover, to understand the content created by Muslim users on TikTok and how the media participates in the construction of the Islam = terrorism binomial in more depth, semistructured interviews were conducted during June and July 2022. Two main themes were addressed in the interviews: the use of terms and the stigmatization of Muslims. To obtain a diverse sample, seven journalists (who preferred to remain anonymous) were selected from different local, national, and international media outlets (Table 1). This number of journalists was selected because it was found that no new information was being collected thereafter, thus reaching what is known as thematic saturation (Guest et al., 2020). The following process was applied for data collection: (a) The interview was carried out according to previous research related to the object of study; (b) journalists with expertise in the subject were selected and contacted online to schedule an interview; (c) permission was requested to use the interviews for research purposes and their transcription; (d) the interviews were transcribed and summarized to identify the most important and relevant data obtained in terms of the object of study; (e) data analysis was performed by using Atlas.ti 8.

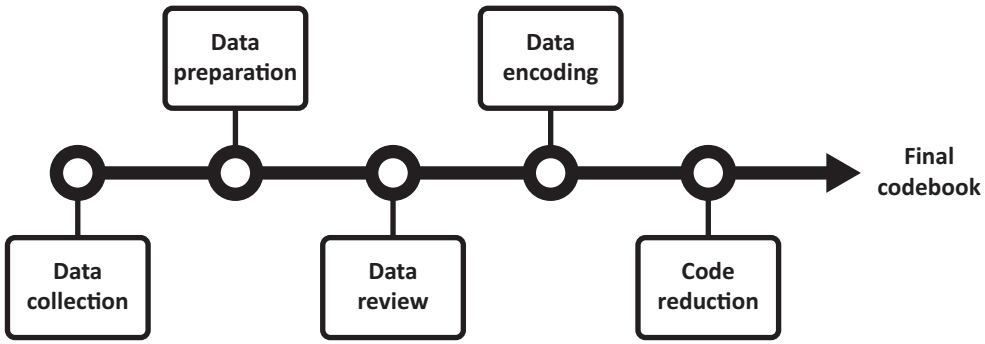


Figure 1. Data analysis procedure.

3.2. Codebook and Data Analysis

The data analysis was carried out using inductive–iterative reasoning, to identify the topics of interest for the current research. This approach enabled findings to be extracted from the data and allowed us to increase our knowledge about the subject (Vives-Varela & Hamui-Sutton, 2021). Three rounds of analysis were carried out, yielding a total of 30 codes. Finally, a total of seven of these were considered relevant to answer the research questions these codes are now presented in Table 2.

3.3. Ethical Considerations

Any reference to private persons or companies has been removed from the examples and images in this article, both to protect their identities and to avoid copyright infringement (Boyd & Crawford, 2012).

4. Results and Discussion

The results are described according to the research questions presented above. We also describe the conversations with journalists, which enable us to delve into the construction of this meaning and the understanding of the discourses created by Muslim users on TikTok.

4.1. Discourses Used by the Muslim Community to Respond to the Binomial Islam = Terrorism

The most relevant discourses propagated by the Muslim community about the Islam = terrorism binomial include

Table 1. Interviewees' characteristics.

Name	Nationality	Type of media outlet
EN1	Argentinean	International
EN2	Spanish	National
EN3	Moroccan	National
EN4	Spanish	International
EN5	Italian	International
EN6	Spanish	Local
EN7	Spanish	National

Table 2. Codebook for content analysis.

Codebook	Definition
DIS	This code helps us to classify the discourses used to challenge the idea that Islam = terrorism.
OTHER	This code helps us to identify them versus us discourses (otherness).
OPIN	This code describes the opinion of users regarding the news published by mainstream media.
STOP	This code allows us to identify discourses against islamophobia.
AFFOR	This code helps us to analyze the affordances used to challenge misconceptions spread by mainstream media.
STRA	This code allows us to classify content creation strategies to respond to the binomial Islam = terrorism.
LANG	This code describes the use of multimodal elements such as captions, hashtags, icons, gifs, etc.

the explanation of their religion, its terminology, and the dismantling of dominant discourses. Thus, we identify two main issues: the simplistic presentation of Islam and the representation of Muslims in the Western media. Following Echchaibi (2013), it is thus observed how the possibilities offered by social media platforms promote the desire of modern participants to subvert the mainstream discourses and become actively involved in political discourses. The formation of unconventional discourses related to religion on TikTok by this collective shows, according to Hoover and Echchaibi (2014), that digital platforms act as a third space that allows the negotiation of the values of Islam in contemporary societies and a deconstruction of the established social imaginaries regarding this religion. The third space becomes a place where not only can religious practices be exercised but also alternatives to the hegemonic narratives can be constructed. This locates the present study in the fourth wave of digital religion, which emphasizes the political and social aspects of research on religions (Campbell & Evolvi, 2020).

As stated by Jaramillo-Dent et al. (2022) in their research on the creative practices of migrant TikTokers, content creation enables minorities to express themselves and construct alternative meanings to the prevailing conversations. In the case of TikTok, its affordances offer a space to develop involvement and engagement in politics, helping users to mobilize (Cervi & Marín-Lladó, 2021; Galpin, 2022) and generating new forms of expression that enable the social empowerment of the Muslim community. Moreover, presenting such discourses on TikTok amplifies a narrative that lacks visibility in other spaces (Echchaibi, 2013). The present analysis revealed four main types of discourses that are used to challenge social-political conceptions about Islam: (a) explanatory, (b) controversial, (c) demystifying, and (d) performative.

Firstly, we identified content aimed at explaining and deepening knowledge of what Islam is and clarifying different concepts that sometimes are used indifferently (Ahmed & Matthes, 2017). These concepts are Arab, Muslim, Islamic, Islamist, and jihad(ist). Owing to the relationship between jihad and terrorism, there is an ongoing desire to explain the meaning of the word "jihad," whose translation would actually be "effort." However, two strands can be detected. The first defends that the word "jihad" indeed means "effort," despite the appropriation of the term by the ISIS terrorist group. On the other hand, other TikTokers accept this appropriation and, when explaining concepts, argue that jihad is indeed related to terrorism. This disparity of opinions indicates that, contrary to its presentation in mainstream media (Ahmed & Matthes, 2017), the Muslim community is not a monolithic entity. However, according to Campbell and Evolvi (2020), such disparate contributions on social media can lead to a misrepresentation of identities rather than to social change, despite the appearance of a collective purpose, in this case, to convey that Islam is a religion of peace.

The journalists' responses reveal the need to take advantage of this third space to raise awareness of these different terms, which could break the misconceptions about Islam (Pearce et al., 2020). According to the interviews conducted, content creators thus need to explain them and improve their understanding by the majority. According to Romero-Rodriguez et al. (2021), knowledge of those considered to be different avoids falling into the traps of stereotyping and polarization. From the interviews:

> IN2: In a society of immediate consumption, the traditional media tends to draw a broad picture and tell things in general terms, without differentiating between terms.

> IN4: When writing, in order to avoid repetition, we sometimes use terms as synonyms whereas, actually, they have their own nuances and differences.

Secondly, the discourse classified as controversial refers to those publications aimed at questioning mainstream media. This content is focused on expressing opposition to news frames, highlighting how mainstream media use terrorist attacks to drive the concept that Islam is a threat to Western values. Moreover, as part of this type of discourse, TikTokers aim to raise awareness among the majority population that the mainstream media are politically biased sources (Jean-Kenix, 2011; Poole, 2002). These discourses agree with Ahmed and Matthes (2017), who support that western media demonize Islam by relating it to terrorism. Despite numerous studies (Casteleiro-Ruiz-de-Azcarate, 2015; Miles, 1989; Said, 1978) have demonstrated that news frames applied to Islam are stigmatizing, our data evidence how this is currently undergoing a process of change and improving the representation of Islam in the media. From the interviews:

> IN3: I am aware, both due to the changes made in the media I work and in which other colleagues work, that work is being done to not link Islam with a political and military ideology in a generalized way. More and more specialized journalists in newsrooms and international organizations are reviewing news of this nature.

> IN4: The change is growing, and although a Muslim is rarely the protagonist of a positive news story, the activism of the collective on social media is promoting change in the traditional media.

These responses related to representation in the media reveal how the pressure exerted by this minority group on social media is shaping the hierarchies of power and diminishing the weight of the media in public opinion as observed by Echchaibi (2013) and Evolvi (2017). Platforms such as TikTok facilitate the dissemination of

minority opinions and challenge hegemonic discourses (Literat et al., 2022), allowing them to make their culture and identity visible (Abidin, 2021).

Furthermore, the demystifying discourse, which refers to content focused on emphasizing the positive aspects of Islam, promotes not only visibility but recognition. We observed how the creator demystifies the religion and encourages its recognition by relating Islam to love and peace, thus seeking common ground with "them." However, according to Gray (2013), this need not necessarily promote real recognition, as it may be a neoliberal technique to celebrate diversity rather than a recognition of the structural problems associated with Islam.

In the videos focused on demystifying the symbols of Islam, it is mainly argued that they are not terrorists, and that the Muslim community is the main victim of the terrorism that is classified as jihadist. This is evidence that they're committed to negotiating their religious values with modern societies and altering Western narratives by promoting the relationship between Islam and modernity (Göle, 2002). Interaction with the dominant culture and common ground is sought to reduce polarized discourse, which, in accordance with Vizcaíno-Verdú and Aguaded (2022), enhances opportunities to build new social relationships and reinforce recognition. On the other hand, according to the analysis conducted by Hamid et al. (2022), for messages to have an impact they must transmit new narratives, construct discourses on what we have in common, and not focus on differences (them versus us), to enable a feeling of identification in the audience and promote real recognition. The results obtained from the interviews show the pressure they suffer to create demystifying discourses and promote action to make their voice heard. From the interviews:

> IN5: In investigative reporting on jihadist terrorism, they wheel out Muslim people from the neighborhood to make themselves look good, but the main message is still "beware that a neighbor may be a potential terrorist."
>
> IN6: The voice of the Muslim collective only appears in the media when there is a terrorist attack.
>
> IN7: The media generates ignorance about Islam since we often work under the dictates of political groups.

Finally, we identify the performative discourse which refers to those narratives constructed through mimesis helping users express themselves creatively and connect with their collective (Kaye et al., 2021). This content is created using audio and videos that are constantly copied on the platform with the addition of new meanings (Literat et al., 2022; Vijay & Gekker, 2021). The use of these trends reveals an intention to get visible and challenge polarization by using the vernacular language spoken on the Internet.

4.2. TikTok Affordances Used to Challenge Misconceptions Spread by Media

After analyzing the discourses, we correlated the codes focused on the discourse with those focused on the affordances in the sample units. We identified that the most relevant affordances used by these content creators on TikTok are audio, hashtags, text, embedded text, and visual elements. In this section, we delve into the use of TikTok affordances to challenge misconceptions spread by media. This will allow us to better understand how they deploy the discourses to respond to the binomial Islam = terrorism.

Our analysis revealed that the use of original audio in combination with visual affordances such as embedded texts and images—that visually strengthen the message—enables us to explain and argue the different concepts (see Figure 2): Arab, Muslim, Islamic, Islamist, and jihad(ist). Furthermore, the use of the hashtags #aprendeconTikTok (learningwithTikTok), #curiosidadesenTikTok (curiositiesonTikTok), or #educación (education) is also observed, reflecting the explanatory aim of such posts. Previous research has shown that activist political practices on social media are more effective when followed by an explanation and intellectual argument, thus strengthening the message and enhancing the critical level of society (Raley, 2009).

On the other hand, the embedded text and captions are the most used affordance to express their annoyance with the media's treatment of Islam. However, visual elements such as embedded image carousels are also identified, being used to highlight the positive aspects of Islam and reinforce the textual argument (Figure 3). This evidence that TikTok users employ the creative options offered by this platform to question media representations of "others" and expand alternative narratives (Literat et al., 2022). These affordances are accompanied by hashtags such as #islam or #muslims, as well as music that is trending on the platform and thus extends the reach of the transmitted message and the number of views (Zhao & Wagner, 2022). This reflects new approaches for campaigning in the face of otherness and power structures, supporting the theory proposed by Bossetta et al. (2018) that political participation is improved and increased using social media. According to Echchaibi (2013), the propagation of these discourses allows the topic of Islam to be located at the center of modern society, promoting its acceptance and recognition in the society to which it belongs. However, recent research conducted within the MAGIC project (Muslim women and communities Against Gender Islamophobia in soCiety; see Hamid et al., 2022) indicates that, despite the importance of challenging mainstream media on social media to change the dominant discourse, one should go further and contact the media to raise awareness of the message transmitted.

These affordances are also used to demystify the discourse about Islam. TikTok offers a space for debate

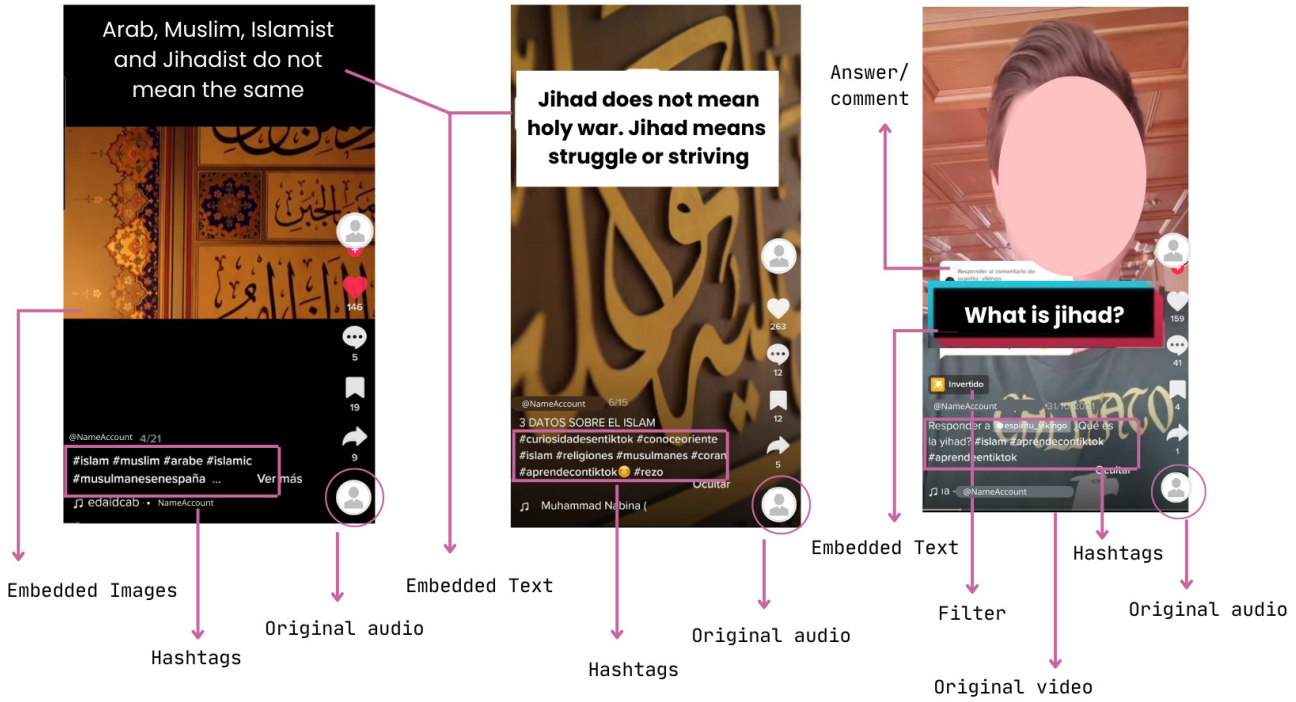


Figure 2. Use of visual and oral affordances to explain religion.

and to combat these threats (Abidin, 2021) helping to challenge the misconception that Muslims are terrorists. These affordances facilitate action and offer an opportunity to justify their religious values to be understood by others. The hashtags most used to demystify Islam are #allah, #hijabi, #muslim, and #islam, followed by positive ones such as #love and #peace, thus relating Islam with peace (Figure 4). At this point in the analysis, it can thus be seen how TikTok offers a space to deconstruct misconceptions and promote acceptance and recognition, whereas the use of this platform by this collective does not promote total change. Gray's (2013) theory, i.e., that social media increase the visibility but not necessarily the actual recognition of minority groups, holds. This drives us to support the idea that marginalized people are between risk and opportunity on social media (Pearce et al., 2020).

Our analysis evidence how TikTok facilitates the creation of content based on popular trends to reach a wider audience (Kaye et al., 2021), and offers a space for

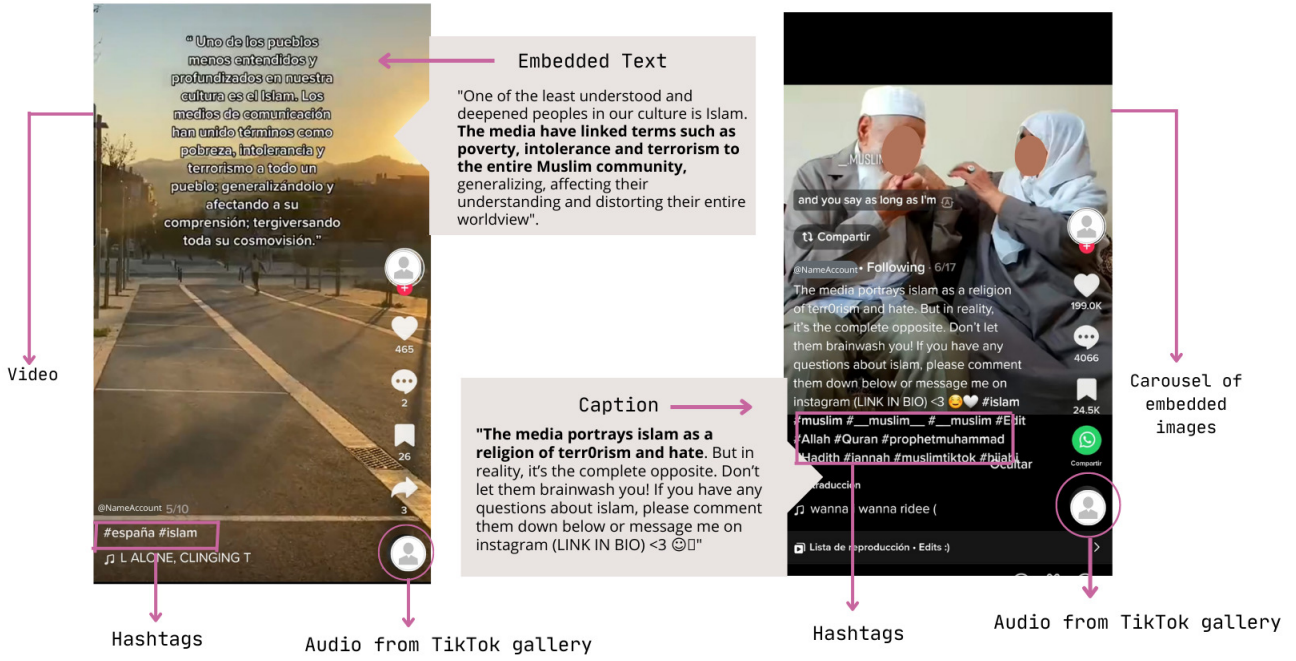


Figure 3. TikTok affordances used to challenge hegemonic discourse.

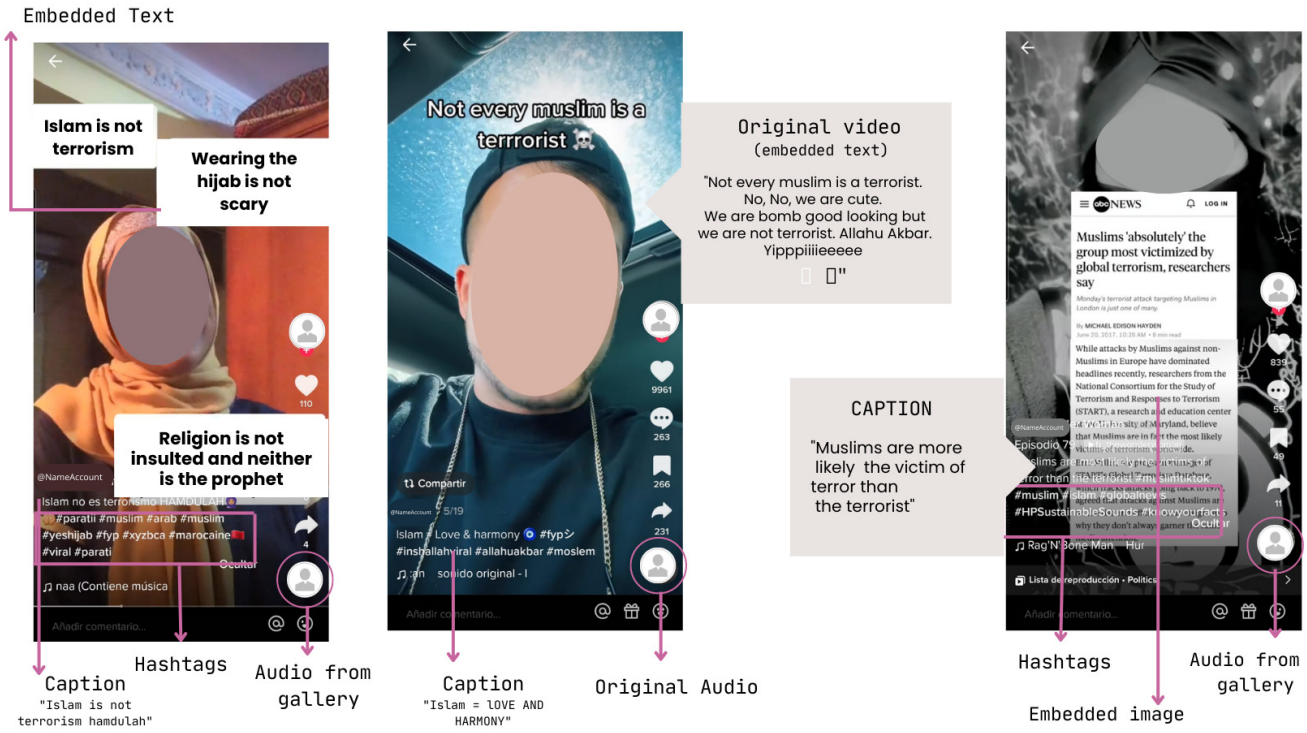
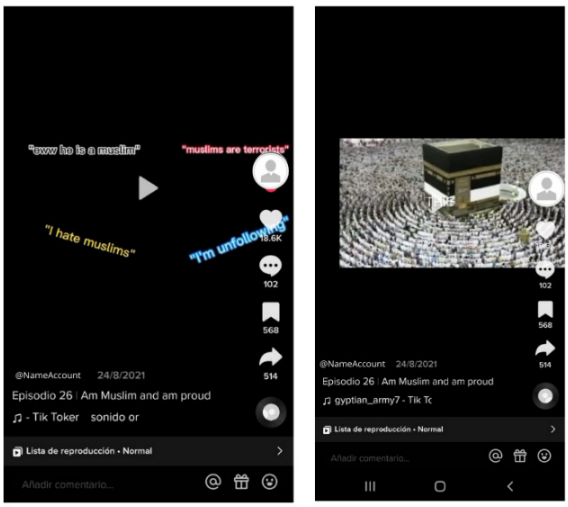


Figure 4. Uses of TikTok to demystify the discourse about Islam.

sense-making of "other" and "otherness" (Schellewald, 2021). It notes the use of hashtags such as #viral, #followme, #goviral, #FyP, and #Foryou, to extend the reach and thereby improve the visibility of the post (Klug et al., 2021). Moreover, the discourses to increase visibility are created from trending audios and videos, lacking original audio and generally presenting visual arguments. In the case of the two most used trends, what the West thinks of Islam is presented using embedded text, while the corresponding replies, which are considered to be "the reality," are presented with images that reflect the positive values of Islam. The arguments with embedded text are accompanied by a sad face emoji, reflecting the Muslim community as a victim. The use



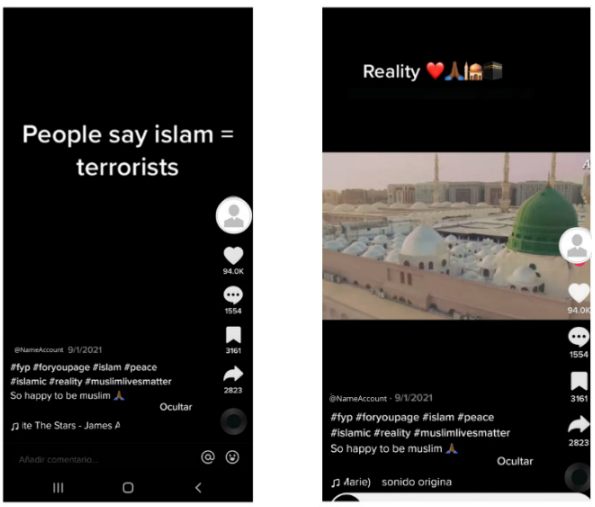


Figure 5. Uses of TikTok trends to increase the visibility of religion.

of humorous strategies with hashtags such as #humor or #joke to challenge oppressive structures is also identified. It is found that the struggle of otherness is not handled in a positive way, emphasizing the differences between "us" and "them" rather than the empowering aspects of differences, which is fundamental to promote recognition (Romero-Rodriguez et al., 2021).

5. Conclusions

The analysis of TikTok content created by the Muslim community about the binomial Islam = terrorism spread by mainstream media reveals this platform as an environment to challenge prevailing misconception narratives. In this sense, it appears that Muslim minorities need to motivate discourses in a way that penetrates societies, thus promoting social change and challenging misconceptions. TikTok allows this community to seek recognition as well as to generate discourses that make their culture visible (Abidin, 2021). However, the way these creators use TikTok highlights the misconceptions about Islam spread by the mainstream media. This does not promote the construction of an alternative narrative as evidenced by Hamid et al. (2022).

Regarding our first research question, the discourses used to address the dichotomy Islam = terrorism made negative misconceptions about Islam and Muslims more visible. Instead of constructing new discourses, the arguments used to challenge Western mainstream narratives and take part in political debate are based on negative narratives. Thus, the reconfiguration of dominant discourses on TikTok is constructed within the spiral of the prevailing narrative. However, according to the journalists interviewed, we found that messages posted by this collective destabilize Western narratives, promoting the relationship between Islam and modernity (Göle, 2002). As supported by El-Haj and Bonet (2011), these discourses deployed on TikTok are a result of the pressure and need to constantly challenge power structures. They intend to be recognized by others and to value what is considered contrary.

In response to the second research question, the results show that TikTok affordances encourage political participation (Brown et al., 2022) and provide tools to foster social empowerment (Vizcaíno-Verdú & Aguaded, 2022), providing a space to challenge the discursive soapbox and the holistic construction of the counter-narrative. However, the use of affordances does not necessarily contribute to diminishing the difference between "us" and "them." This means that textual and visual elements are used to highlight the differences between what "we are" versus what "they are," while other affordances such as "duo" or "green screen" do not encourage the challenging of opinions, according to Zulli and Zulli (2020). Following Civila and Jaramillo-Dent (2022), the spread of such content, beyond empowering the Muslim collective, provides visibility and recognition of discriminatory practices, which reinforces Gray's (2013) theory of recognition. As a result, it seems that the affordances of the platform have an impact on the construction of otherness, while TikTok potentially fails to provide a space for these groups to confront established mechanisms of domination.

The study limitations were presented in terms of sample and platform. Although the sample provides answer to our research questions, we acknowledge that according to the TikTok algorithm, our location and search term could limit the results shown by this platform (Jaramillo-Dent et al., 2022). Furthermore, our study is focused on one platform, which could be amplified in further research. Accordingly, we suggest that further research approaches interview the users to have a better understanding of how they represent themselves in social media spaces.

Acknowledgments

This study was supported by the R+D+I Project (2019–2021), entitled Youtubers and Instagrammers: Media Competence in Emerging Prosumers, under code RTI2018–093303-B-I00, financed by the Spanish Ministry of Science, Innovation, and Universities, the European Regional Development Fund (ERDF), and the R+D+I project (2020–2022) entitled Instagramers and Youtubers for the Transmedia Empowerment of the Andalusian Citizenry. Media literacy of the Instatubers, with code P18-RT-756, was financed by the Government of Andalusia in the 2018 call for tenders (Andalusian Plan for Research, Development and Innovation, 2020) and the European Regional Development Fund (ERDF).

Conflict of Interests

The authors declare no conflict of interests.

References

Abidin, C. (2021). Mapping internet celebrity on TikTok: Exploring attention economies and visibility labours. *Cultural Science Journal*, *12*(1), 77–103. https://doi.org/10.5334/csci.140

Acim, R. (2019). Islamophobia, racism, and the vilification of the Muslim diaspora. *Islamophobia Studies Journal*, *5*(1), 26–44. https://www.jstor.org/stable/10.13169/islastudj.5.1.0026

Aguilar, G. K., Campbell, H. A., Stanley, M., & Taylor, E. (2017). Communicating mixed messages about religion through internet memes. *Information, Communication & Society*, *20*(10), 1498–1520. https://doi.org/10.1080/1369118X.2016.1229004

Ahmed, S., & Matthes, J. (2017). Media representation of Muslims and Islam from 2000 to 2015: A meta-analysis. *International Communication Gazette*, *79*, 219–244. https://doi.org/10.1177/1748048516656305

Akbarzadeh, S., & Smith, B. (2005). *The representation of*

Islam and Muslims in the media (The Age and Herald newspapers). Monash University.

Al-Rawi, A. (2016). Facebook as a virtual mosque: The online protest against innocence of Muslims. *Culture and Religion*, *17*(1), 19–34. https://doi.org/10.1080/14755610.2016.1159591

Bahfen, N. (2018). The individual and the ummah: The use of social media by Muslim minority communities in Australia and the United States. *Journal of Muslim Minority Affairs*, *38*(1), 119–131. https://doi.org/10.1080/13602004.2018.1434939

Bossetta, M., Dutceac-Segesten, A., & Trenz, H. J. (2018). Political participation on Facebook during Brexit: Does user engagement on media pages stimulate engagement with campaigns? *Journal of Language and Politics*, *17*(2), 173–194.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Brown, P., & Levision, S. C. (1987). *Politeness: Some universals in language usage.* Cambridge University Press.

Brown, Y., Pini, B., & Pavlidis, A. (2022). Affective design and memetic qualities: Generating affect and political engagement through bushfire TikToks. *Journal of Sociology*. Advance online publication. https://doi.org/10.1177/14407833221110267

Campbell, H. A. (2017). Surveying theoretical approaches within digital religion studies. *New Media + Society*, *19*(1), 15–24. https://doi.org/10.1177/1461444816649912

Campbell, H. A., & Evolvi, G. (2020). Contextualizing current digital religion research on emerging technologies. *Human Behavior and Emerging Technologies*, *2*, 5–17. https://doi.org/10.1002/hbe2.149

Canclini, N. (2004). *Diferentes, desiguales y desconectados: Mapas de la interculturalidad* [Different, unequal and disconnected: Maps of interculturality]. Gedisa.

Casteleiro-Ruiz-de-Azcarate, J. (2015). *Islam, terrorismo y medios de comunicación* [Islam, terrorism and media]. Instituto español de estudios estratégicos.

Castells, M. (2013). Prefacio: Autocomunicación de masas y movimientos sociales en la era de Internet [Preface: Mass self-communication and social movements in the internet age]. *Anuario Del Conflicto Social*, *1*(1), 11–19. https://bit.ly/3fIcME5

Cervi, L., & Marín-Lladó, C. (2021). What are political parties doing on TikTok? The Spanish case. *Profesional de la información*, *30*(4). https://doi.org/10.3145/epi.2021.jul.03

Civila, S., & Jaramillo-Dent, D. (2022). #Mixedcouples on TikTok: Performative hybridization and identity in the face of discrimination. *Social Media + Society*, *8*(3). https://doi.org/10.1177/20563051221122464

Civila, S., Romero-Rodríguez, L. M., & Civila, A. (2020). The demonization of Islam through social media: A case study of #Stopislam in Instagram. *Publications*, *8*(4). https://doi.org/10.3390/publications8040052

Deroo, M. R., & Mohamud, I. (2022). "I would rather be informed than misinformed": Critical conversations supporting transnational religious identity across time and space. *English Teaching: Practice & Critique*, *21*(3), 254–266. https://doi.org/10.1108/ETPC-08-2021-0105

Dicks, B. (2019). Multimodal analysis. In P. Atkinson, S. Delamont, A. Cernat, J. Sakshaug, & R. Williams (Eds.), *SAGE research methods foundations*. SAGE. https://doi.org/10.4135/9781526421036831970

Donald, J. (1992). *Cultural Identity*. SAGE.

Echchaibi, N. (2013). Muslimah media watch: Media activism and Muslim choreographies of social change. *Journalism*, *14*(7), 852–867. https://doi.org/10.1177/1464884913478360

El-Haj, T. R. A., & Bonet, S. W. (2011). Education, citizenship, and the politics of belonging: Youth from Muslim transnational communities and the "war on terror." *Review of Research in Education*, *35*(1), 29–59. https://bit.ly/3EvAuxo

Evolvi, G. (2017). Hybrid Muslim identities in digital space: The Italian blog Yalla. *Social Compass*, *64*(2), 220–232. https://doi.org/10.1177/0037768617697911

Ewart, J. (2012). Framing an alleged terrorist: How four Australian news media organizations framed the Dr. Mohamed Haneef case. *Journal of Media and Religion*, *11*(2), 91–106. https://doi.org/10.1080/15348423.2012.688667

Farooqui, J. F., & Kaushik, A. (2022). Growing up as a Muslim youth in an age of Islamophobia: A systematic review of literature. *Contemporary Islam*, *16*(5), 1–24. https://doi.org/10.1007/s11562-022-00482-w

Gabsi, Z. (2015). Response to Islamophobia in the Arabic Islamic discourse: A critical discourse Analysis. *Intellectual Discourse*, *23*(2), 229–254. https://journals.iium.edu.my/intdiscourse/index.php/id/article/view/629

Galasin'ska, A., & Galasin'ski, D. (2003). Discursive strategies for coping with sensitive topics of the Other. *Journal of Ethnic and Migration Studies*, *29*, 849–863. https://doi.org/10.1080/1369183032000149604

Galpin, C. (2022). At the digital margins? A theoretical examination of social media engagement using intersectional feminism. *Politics and Governance*, *10*(1), 161–171. https://doi.org/10.17645/pag.v10i1.4801

Giglietto, F., & Lee, Y. (2017). A hashtag worth a thousand words: Discursive strategies around #JeNe SuisPasCharlie after the 2015 Charlie Hebdo Shooting. *Social Psychological and Personality Science*, *6*(7), 814–822. https://doi.org/10.1177/1948550615590199

Göle, N. (2002). Islam in public: New visibilities and new imaginaries. *Public Culture*, *14*(1), 173–190. https://muse.jhu.edu/article/26271

Gómez-García, S., Paz-Rebollo, M., & Cabeza-San-Deogracias, J. (2021). Newsgames against hate speech in the refugee crisis. *Comunicar*, *29*(63), 123–133. https://doi.org/10.3916/C67-2021-10

Gondwe, G., & Bhowmik, S. (2022). Visual representation of the 2020 Black Lives Matter protests: Comparing US mainstream media images to citizens' social media postings. *Online Journal of Communication and Media Technologies*, *12*(4), Article e202238. https://doi.org/10.30935/ojcmt/12494

Gray, H. (2013). Subject(ed) recognition. *American Quarterly*, *65*(4), 771–798. https://doi.org/10.1353/aq.2013.0058

Guest, G., Namey, E., & Chen, M. (2020). A simple method to assess and report thematic saturation in qualitative research. *Plos One*, *15*(5), Article e0232076. https://doi.org/10.1371/journal.pone.0232076

Hamid, R., Olea, A., & Rodríguez-Alarcón, L. (2022). *Speak up for diversity.* European Institute of the Mediterranean. https://bit.ly/3rRCdWd

Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, M. P. (2014). *Metodología de la investigación* [Research methodology] (6th ed.). McGraw Hill Education. https://bit.ly/2JLPtUM

Hoover, S. M., & Echchaibi, N. (2014). *Media theory and the "third spaces" of digital religion*. University of Colorado.

Ibrahim, D. (2010). The framing of Islam on network news following the September 11th attacks. *International Communication Gazette*, *72(*1), 111–125. https://doi.org/10.1177/1748048509350342

Illman, R., & Sjö, S. (2015). Facebook as a site for inter-religious encounters: A case study from Finland. *Journal of Contemporary Religion*, *30*(3), 383–398. https://doi.org/10.1080/13537903.2015.1081341

Iqbal, Z. (2020). *Islamophobia: History, context and deconstruction.* SAGE. https://bit.ly/3T7iOMG

Islas, O. (2015). La ecología de los medios: Metadisciplina compleja y sistémica [Media ecology: Complex and systemic metadiscipline]. *Palabra clave*, *18*(4), 1057–1083. https://doi.org/10.5294/pacla.2015.18.4.5

Jaramillo-Dent, D., Contreras-Pulido, P., & Pérez-Rodríguez, A. (2022). Immigrant influencers on TikTok: Diverse microcelebrity profiles and algorithmic (in)visibility. *Media and Communication*, *10*(1), 208–221. https://doi.org/10.17645/mac.v10i1.4743

Jean-Kenix, L. (2011). *Alternative and mainstream media: The converging spectrum*. Bloomsbury.

Kaye, D. B. V., Chen, X., & Zeng, J. (2021). The co-evolution of two chinese mobile short video apps: Parallel platformization of douyin and TikTok. *Mobile Media and Communication*, *9*(2), 229–253. https://doi.org/10.1177/2050157920952120

Klug, D., Qin, Y., Evans, M., & Kaufman, G. (2021). Trick and please. A mixed-method study on user assumptions about the TikTok algorithm. In *WebSci '21: 13th ACM web science conference 2021* (pp. 84–92). Association for Computing Machinery. https://doi.org/10.1145/3447535.3462512

Korhonen, V. (2010). Dialogic literacy: A sociocultural literacy learning approach. In M. Lloyd & S. Talja (Eds.), *Practising information literacy: Bringing theories of learning, practice and information literacy together* (pp. 1–16). Centre for Information Studies.

Korteweg, A. C. (2008). The Sharia debate in Ontario gender, Islam, and representations of Muslim women's agency. *Gender & Society*, *22*(4), 434–454. http://www.jstor.org/stable/27821662

Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review*, *117,* 785–807. https://doi.org/10.1037/a0019688

Literat, I., Boxman-Shabtai, L., & Kligler-Vilenchik, N. (2022). Protesting the protest paradigm: TikTok as a space for media criticism. *The International Journal of Press/Politics*. Advance online publication. https://doi.org/10.1177/19401612221117481

Looy, J. (2015). Online games characters, avatars, and identity. In *The international encyclopedia of digital communication and society* (pp. 1–11). Wiley-Blackwell. https://doi.org/10.1002/9781118767771.wbiedcs106

Merino, Á. (2020). *El islam en el mundo* [Islam in the World]. EOM. https://bit.ly/3ynbQv5

Miles, R. (1989). *Racism*. Open University Press.

Mirra, N., Morrell, E., & Filipiak, D. (2018). From digital consumption to digital invention: Toward a new critical theory and practice of multiliteracies. *Theory into Practice*, *57*(1), 12–19. https://doi.org/10.1080/00405841.2017.1390336

Mosemghvdlishvili, L., & Jansz, J. (2012). Framing and praising Allah on YouTube: Exploring user-created videos about Islam and the motivations for producing them. *New Media & Society*, *15*(4), 482–500. https://doi.org/10.1177/1461444812457326

Naderifar, M., Goli, H., & Ghaljaie, F. (2017). Snowball sampling: A purposeful method of sampling in qualitative research. *Strides in Development of Medical Education*, *14*(3). https://doi.org/10.5812/sdme.67670

Pearce, K., Gonzales, A., & Foucault, B. (2020). Introduction: Marginality and social media. *Social Media + Society*. https://doi.org/10.1177/2056305120915588

Piñeiro-Otero, T., & Martínez-Rolán, X. (2020). Para comprender la política digital. Principios y acciones [To understand digital politics. Principles and actions]. *Vivat Academia*, *152*, 19–48. https://doi.org/10.15178/va.2020.152.19-48

Poole, E. (2002). *Reporting Islam: Media representations of British Muslims.* I.B. Tauris. https://bit.ly/3MlEank

Rahman, K. A. (2022). Shifting the dynamics in popular culture on Islamophobic narratives. *Pacific Journalism Review*, *28*(1/2), 19–28. https://doi.org/

10.24135/pjr.v28i1and2.1271
Raley, R. (2009). *Tactical media*. University of Minnesota Press.
Romero-Rodriguez, L. M., Civila, S., & Aguaded, I. (2021). Otherness as a form of intersubjective social exclusion: Conceptual discussion from the current communicative scenario. *Journal of Information, Communication and Ethics in Society*, *19*(1), 20–37. https://doi.org/10.1108/JICES-11-2019-0130
Ross, W. (2014). *Social studies curriculum: Purposes, problems, and possibilities*. State University of New York. https://bit.ly/3CLgmGi
Said, E. (1978). *Orientalism*. Penguin.
Schellewald, A. (2021). Communicative forms on TikTok: Perspectives from digital ethnography. *International Journal of Communication*, *15*(2021), 1437–1457. https://ijoc.org/index.php/ijoc/article/view/16414/3389
Sensor Tower. (2022). *Q1 2022: Store intelligence data digest*. https://bit.ly/3VdloTi
Şot, İ. (2022). Fostering intimacy on TikTok: A platform that "listens" and "creates a safe space." *Media, Culture & Society*, *44*(8), 1405–1601. https://doi.org/10.1177/01634437221104709
Statista. (2022). *TikTok: Datos estadísticos* [TikTok: Statistical data]. https://es.statista.com/temas/7541/tiktok
Statistics and Data. (2022). Most popular religions in the world. https://statisticsanddata.org/data/most-popular-religions-in-the-world-1950-2022
Vijay, D., & Gekker, A. (2021). Playing politics: How Sabarimala played out on TikTok. *American Behavioral Scientist*, *65*(5), 712–734. https://doi.org/10.1177/0002764221989769
Villepin, D. (2003). Islam y occidente [Islam and West]. *Revista política exterior*, *2003*(95). https://bit.ly/3COnYb8
Vives-Varela, T., & Hamui-Sutton, L. (2021). La codificación y categorización en la teoría fundamentada, un método para el análisis de los datos cualitativos [Coding and categorization in grounded theory, a method for qualitative data analysis]. *Investigación en educación médica*, *10*(40), 97–104. https://doi.org/10.22201/fm.20075057e.2021.40.21367
Vizcaíno-Verdú, A., & Aguaded, I. (2022). #ThisIsMeChallenge and music for empowerment of marginalized groups on TikTok. *Media and Communication*, *10*(1), 157–172. https://doi.org/10.17645/mac.v10i1.4715
Wheeler, K. R. (2014). Remixing images of Islam. The creation of new Muslim women subjectivities on YouTube. *Journal of Religions on the Internet*, *6*, 144–163. https://doi.org/10.11588/rel.2014.0.17364
Zhao, H., & Wagner, C. (2022). How TikTok leads users to flow experience: Investigating the effects of technology affordances with user experience level and video length as moderators. *Internet Research*. Advance online publication. https://doi.org/10.1108/INTR-08-2021-0595
Zulli, D., & Zulli, D. (2020). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media and Society*, 24(8), 1872–1890. https://doi.org/10.1177/1461444820983603

About the Authors

Sabina Civila is a FPI predoctoral researcher and a PhD candidate in the Interuniversity Communication Program (Málaga, Huelva, Sevilla y Cádiz) in the educommunication line. She has a master's in strategic communication and communication innovation (Malaga, Huelva, Seville, and Cádiz), and graduated in advertising and public relations from the University of Cádiz. Sabina is a researcher at the "Ágora" Research Group (HUM-648). Currently, she focuses her research on social networks, educommunication, media literacy, and Islamophobia.

Mónica Bonilla-del-Rio is a PhD candidate in the Interuniversity Communication Program at the University of Huelva in the media literacy and educommunication line. Monica graduated in teaching preschool education at the University of Cantabria (Spain) and has a master's in emotional, social, and creativity education in the University of Cantabria, and in communication and audiovisual education in the Internacional University of Andalucia and University of Huelva. Nowadays, she is a PhD fellow at the University of Huelva and is developing her dissertation about the digital and social inclusion of young people with disability.

Ignacio Aguaded is a full professor of education and communication at the Universidad de Huelva. He is Editor-in-chief of *Comunicar* and the president of Grupo Comunicar, a long-standing media literacy collective in Spain. He is also head of the "Ágora" investigation team that forms part of the Andalusia Investigation Plan (HUM-648), is the director of the international master of communication and education (UNIA, UHU), and the coordinator of the UHU Interuniversity doctoral program in communication (US, UMA, UCA, UHU).

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 127–137
https://doi.org/10.17645/pag.v11i2.6302

Article

Demonising Migrants in Contexts of Extremism: Analysis of Hate Speech in UK and Spain

Darren Lilleker [1] and Marta Pérez-Escolar [2,*]

[1] Department of Media and Communication, Bournemouth University, UK
[2] Department of Communication, Universidad Loyola Andalucía, Spain

* Corresponding author (martaperez@uloyola.es)

Submitted: 21 October 2022 | Accepted: 12 January 2023 | Published: 17 May 2023

Abstract
Hate speech has been seen as a problem within democratic societies that has been exacerbated by social media. While platforms claim to moderate content, this proves impossible. Studying popular platforms in the UK and Spain and examining content within community pages dedicated to right-wing parties, we use framing analysis to identify the predominant frames in user comments that contained hate speech against migrants. Our research demonstrates a frequent use of arguments that encourage xenophobic and discriminatory attitudes to flourish. Specifically, we find that immigrants are commonly framed as potential criminals, people who steal resources and erode norms of the dominant culture and traditions. The fact that these frames are commonly used is worrying and indicates xenophobic attitudes exist within both societies under study. However, it is difficult to imagine regulatory systems that would prevent these attitudes from being expressed. Rather, we argue that it is more important to focus on correcting the conditions that cause such attitudes to take hold within a society.

Keywords
extremism; far-right parties; hate speech; immigration; social media; Spain; UK

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

Research on user comments has flourished across different disciplines over the last few years (Schindler & Domahidi, 2021), including studies specifically concerned with the increased use of hate speech online (Ernst et al., 2017), commenters' motivations, and how they relate to the increasingly polarized political environment (see Almoqbel et al., 2019). Citizens' comments can be used as cues by other readers regarding the stance they should take on an issue (Springer et al., 2015) and can counter more accepted or socially acceptable viewpoints (Weber et al., 2017), particularly when they are viewed as the "authentic voice" of the ordinary person. Hence, and as per previous studies of online user comments, we view these comments as instrumentalist interactions (Lilleker & Bonacci, 2017; Zurutuza-Muñoz & Lilleker, 2018). Comments, from this perspective, are made by users to express their views and gain reactions from others.

Comments within spaces on social media platforms can be homogenous, in which case this space acts as an echo chamber (Auxier & Vitak, 2019), or heterogeneous and antagonistic. However, more extreme views are most likely shared in spaces populated by like-minded individuals with a specific ideological perspective. Research suggests that these types of spaces seem to be breeding grounds where polarisation, hate, incivility, and conspiracy theories become prevalent in online discourse (Bolsover, 2020). While it is unclear the extent

to which content from within these spaces infects the wider discourse on any social media platform, there are concerns that this might be the case (Bennett & Livingston, 2018). Evidence shows that holding strong affectively polarised and extremist positions is mostly caused by obtaining news from social media (Nguyen & Vu, 2019). In particular, when there is a strong link between identity and a political position, it is more likely for discourse to lead to the use of othering and hate speech (Wasilewski, 2019). The discourses within these spaces often represent a rejection of political correctness and reflect a raw emotional response towards a target deemed as the "other" (Hamed, 2020).

Using framing analysis, this article explores, comparatively, user comments regarding migrants on social media platforms that are popular both in the UK and Spain. The main purpose of this research is to identify the predominant frames used by users whose comments contained hate speech against migrants on party pages. These spaces are largely unmoderated, with groups only being closed sporadically and when identified as engaging in illegal activity. Hence, they may appear as spaces where the true views of users can be expressed and so provide research insights into currents of thinking within societies.

2. Hate Speech and Othering

Hate speech is a term that is used widely, sometimes weaponised to silence certain viewpoints (Gelber & McNamara, 2016). However, despite attempts to prevent its spread (see Jougleux, 2022), definitions of hate speech remain loose and open to interpretation. Hate speech is used to encompass language or discourse that expresses strong dislike and discrimination, encourages violence or any kind of attack, or diminishes a person, group, or institution. This broad definition includes expressions used to attack or threaten others or their rights (e.g., other commenters, journalists, politicians, or specific races, ethnicities, etc.) but to attack ideas that are damaging to their image as a member of a group.

Wardle and Derakhshan (2017) set out a broad taxonomy to help researchers frame hate speech alongside the post-truth communication phenomenon. They suggested three types of "information disorder," which describe the extent the content is intended to cause harm (Wardle & Derakhshan, 2017, p. 20). Firstly, they classify disinformation as content that is false and deliberately created to harm a person, social group, organization, or country (e.g., by creating a false connection using misleading or manipulated content). Secondly, they classify misinformation as false but not created deliberately to cause harm (e.g., satirical content). Thirdly, they classify malinformation as content based on reality but used to inflict harm on a person, organization, or country (e.g., leaks, harassment, or stereotyping). Based on this taxonomy, hate speech can fall into any category as it exaggerates threats intending to violate norms and attack a group due to a specific identity factor such as race, gender, religion, ethnicity, or nationality, among others (Emcke, 2019; Rossini, 2020). Rossini (2020, p. 6) argues in this vein that hate speech is a subtype of intolerance because the term proves "too narrow in scope to address intolerant expressions that occur in relatively public digital spaces." Regarding online hate speech, Rossini (2020) emphasizes the need to tackle complex or abusive forms of online discourse. The most overt hate speech is simple to identify; that being said, though false stories or loaded articles that question the fitness of women to govern (Sheckels et al., 2012) or suggest children growing up with same-sex parents could be mentally or even sexually abused (Strand & Svensson, 2019) are clearly forms of hate speech, they are not always classified as such and remain prevalent despite platform moderation. Hence, many types of hate speech are explicit and relatively unequivocal as they explicitly cause harm to victims, for example, dehumanizing people by comparing them to animals or vermin (Williams, 2021). Yet, hate speech can also be implicit, as in arguments that reinforce negative stereotypes and can lead society to develop negative impressions of groups (Rieger et al., 2021). In fact, a widely used form of hate speech, which is not in itself directly threatening, is the use of negative stereotypes, low-level insults, micro-aggressions towards individuals, groups, or institutions, and/or discriminatory and/or negative content based on characteristics, such as race, nationality, religion, gender, physical attributes, ideas, ideologies, etc. Such arguments can move the Overton window, the range of acceptable viewpoints within a society, to encompass xenophobic or misogynistic perspectives (Lilleker & Ozgul, 2021). These forms of discourse should be classified as hate speech as they spread the impression that the targeted other is inferior, less than human or evil, and so represents an existential threat that is harmful to society. However, the challenge in identifying all forms of hate speech is that not all hate speech involves a direct attack or threat and so the hateful nature of the speech is only visible if one takes into account the full context of an argument.

Hence, to understand how hate speech academics should work with a broader definition that includes the context encompassing terms that become weaponised for political purposes. The negative connotation that has become attached to the term "woke" exemplifies an example of implicit hate speech. The term woke was originally popularised by Black Americans in the early 20th century to highlight the importance of being empowered to recognise and overthrow racism and oppression in society. The word gained prominence again with the rise of the Black Lives Matter movement, campaigns in support of legalizing same-sex marriage, and the #Metoo movement in the 2010s. However, woke has recently been adopted among social conservatives as a pejorative term employed to dismiss the arguments of those classified as "woke." The "woke" are thus classified as extremists who promote progressive social justice

policies which run counter to traditional values. Hence, woke went from a social justice concept to a pejorative term weaponised by the alt-right in an attempt to discredit or silence progressive voices (Cammaerts, 2022). The context here is crucial and relates to how a frame is constructed with clear negative connotations. Framing refers to how a shared understanding and meaning is constructed within a wider social context within which a group can be positioned, for example, the moral versus immoral frame is found to be common within the use of hate speech (Armstrong & Wronski, 2019). Those using hate speech are defenders of the values of a society and so are involved in a moral crusade against those who undermine those values; those framed as immoral can be migrants with different cultural backgrounds or the "woke" who call out anti-immigration campaigners as racist. But frames are complex constructions that may not involve overtly pejorative terms. Consequently, hate speech can only be understood by analysing how terms are used in context; a term may not be in itself derogatory but can be used in a way to define members of an outgroup and connote their exclusion from the mainstream. As Lynch (2022) notes, due to the framing of an argument, terms can develop a fixed meaning which can signal the othering and exclusion of individuals of certain backgrounds or political opinions, the process of othering can shape attitudes which underpin affective polarisation and the dehumanisation of groups.

Framing research dates back to sociology (Goffman, 1974; Sádaba, 2006) and psychology (Tversky & Kahneman, 1981), and has become a reliable method in communication research (Cappella & Jamieson, 1996), media studies (Herrero-Diz & Pérez-Escolar, 2022), and political studies (Fenoll & Rodríguez-Ballesteros, 2017). For Entman (1993, p. 52), framing means to "select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation for the item described." The frame emphasizes or gives salience to certain characteristics of a subject and forms the central idea that structures and organizes the information (Gamson & Modigliani, 1989). In this sense, frames act as cognitive filters that determine how the receiver will interpret and understand the complex world they inhabit (Lippmann, 1922). Hence, to understand the framing strategy within a text it is important to analyse the discourse, the terminology used, the linkages made between different items, the stereotypes employed, etc. These are all common heuristic devices used when framing an argument during social interactions to imply negative and positive associations, as strategically these aid audiences understand the wider context surrounding the specific event and its implications (Lakoff & Johnson, 2001). Thus, the frame is designed to determine the perspective that represents the shared understanding of a situation or problem (Takeshita, 1997). In the field of communication, authors have divided frames into two general categories: media frames and audience frames (Lin, 2016; Scheufele, 1999). However, in the current digital ecosystem, it is difficult to clearly distinguish between media and audience frames, as they can combine both elite messages and representations of how the individual or a collective perceives, organizes, and interprets (cognitive processes) a topic, event, or situation (Goffman, 1974; Lin, 2016). As Qin (2015, p. 169) suggests, "a major challenge is whether the frames in social media are media or audience frames in nature, given that social media are a mixture of institutional accounts and individual accounts." Therefore, in this study, it has been deemed appropriate to adopt the perspective of audience frames because individual users and communities actively produce content (user-generated content) on the analysed social platforms.

There is significant evidence that hate speech is a feature of online discourse (Rossini, 2020), and this has led to its prioritisation in discussions regarding the regulation of online environments with little impact on its prevalence. Rieger et al. (2021) suggest five reasons why hate speech is particularly problematic online. The first one is continuity: As Gagliardone et al. (2015) note, moderators can remove hateful content but this may have already been reposted to the same or different platforms. This highlights the second reason: Hate speech is spreadable. Due to its contentious and emotive nature, hate speech gains visibility quickly within an ecosystem designed for sharing and spreading content (Jardine, 2019). Thirdly, and most relevant to certain platforms, users tend to be more aggressive or extreme when they feel they are anonymous. Anonymity makes users feel less accountable, which empowers them to "be more outrageous, obnoxious, or hateful" (Brown, 2018, p. 298) either to gain rewards from other users or because they feel they are free to say what they really think. Mondal et al. (2017) found anonymity fuels hate in online media environments, particularly concerning race or sexual orientation, and that, in turn, this can lead to further and more extreme expressions. Fourthly, using hate speech is seen, instantly, as victimless as the actual target tends to be an invisible other unlikely to be present in the space where hate speech is expressed (Rieger et al., 2021). This leads to a disinhibition that is often absent in real-life contexts. Finally, hate speech can be "memeified" and mixed with satire or humour (Rieger et al., 2021), making it even more spreadable. Memeification is a common practice among members of alt-right movements (Rieger et al., 2021) with sharing denoting membership of a user community with specific knowledge which delineates them from "clueless outsiders" (Tuters & Hagen, 2020).

Attempts to monitor and exclude such forms of language have proven largely ineffectual (Jougleux, 2022). Some terms can be detected automatically (e.g., swear words, obscenities) however it is a simple process for commenters to use acronyms, symbols, or substitute words to offend or affront others and go undetected by automated moderators. Artificial intelligence can only

detect certain words and phrases or styles of argumentation which are likely to correlate with the use of hate speech. They can detect demeaning and negative language, words in capital letters (indicating yelling), mockeries, and insults. However, the full range of language forms encompassed by definitions of hate speech cannot be detected and often the context of an argument matters. Hence the diverse ways language online clearly violates social norms of conversation and politeness and is consistent with racist, xenophobic, misogynistic, and discriminatory language remains highly visible and prevalent (see Suhay et al., 2017). The detection and understanding of hate speech is important. Firstly, revealing such strains of argumentation sheds light on currents of thought within a society (Lilleker & Bonacci, 2017; Zurutuza-Muñoz & Lilleker, 2018), if not their strength. Secondly, in understanding what attitudes exist, strategies can be developed to counter those attitudes. Blunt methods such as censorship are limited in their ability but also come under attack for placing limits on freedom of speech despite it being recognised that caution is needed when allowing hate speech under the aegis of free speech. In the paradox of tolerance, Popper (1996) highlighted the inconsistencies of extending freedom of speech to extremism as their discourse by nature threatens core democratic values of pluralism and inclusivity. As Popper (1996) noted, if extremists gained power they would immediately silence dissenting voices. Thus, those who cry foul when their freedom of speech is restricted are likely the same individuals who would advocate its extinction. However, while moderation is valid, its imperfection as a tool means alternative approaches are required. Radical behaviours and extremist ideological polarization have become a global concern, substantially affecting social relationships and impacting public understanding of the world (Pérez-Escolar & Noguera-Vivo, 2022). Hence, it is far more important to find ways to prevent people from using hate speech, to stop its spread, and to ensure prejudiced and discriminatory attitudes do not become normalised within society.

3. The Context of Hate Speech: Migration and the Far-Right in the UK and Spain

Hate speech that targets immigrants tends to be the preserve of far-right political parties and groupings that adopt an exclusionary populist discourse (Jessoula et al., 2021). Hence, we analyse discourse within user comments on the pages dedicated to the UK Independence Party (UKIP) and VOX. Both parties have been at the forefront of increasing the visibility of immigration as an issue, causing more mainstream and centrist parties to adopt tougher policies.

3.1. UK

In the UK, Enoch Powell, then a prominent Conservative, made immigration a political issue in 1968 when warning of "rivers of blood" flowing through the streets in the coming conflict between the white British and growing Black populations (Atkins, 2018). Powell's speech gave succour to a growing far-right that brought violence to the streets of many UK cities leading to a backlash from immigrant communities (Renton, 2018). These tensions calmed, but residual xenophobia remained despite declarations that the UK was a truly multicultural nation (Wetherell, 2008). The bombings of the London tube by Islamist extremists in 2007, the increase of migration from Eastern European countries after the accession of Poland to the EU in 2004, and that of Bulgaria and Romania in 2014, giving them access to freedom of movement, and the migration of refugees across the English Channel, which emerged as an issue in 2018, kept immigration on the agenda from the 2010s onwards. As early as 2005, Conservatives had an election pledge to "control immigration" in response to migration from Poland. Pressure from UKIP, who claimed 30 million Bulgarians and Romanians would have open access, aligned immigration with EU membership, partially leading to the close vote to leave the union in 2016. In the lead-up to the referendum, UKIP was found to have a large and very active following on social media (Lilleker & Jackson, 2017) and their controversial discourse allowed them to gain significant attention from mainstream media (Murphy & Devine, 2020). The referendum was widely seen as a largely successful attempt by the Conservative government to head off the electoral challenge from UKIP (Smith, 2018) although the 2015 election victory meant the promise of a referendum had to be honoured. The referendum result saw an initial backlash against any person perceived to be of a non-British heritage, but focus was recently placed upon refugees attempting to make their way to the UK via dangerous crossings of the English Channel by small boats. The numbers are reported to have increased from 299 in 2018 to 28,526 in 2021 according to official figures reported by Sky News (Scott, 2022). UKIP's prominence declined after the referendum and its charismatic leader Nigel Farage left to form the Brexit party. However, it built a following amongst working-class white males with an angry disposition towards liberal values by positioning itself as a defender of Britishness. UKIP's anti-immigration discourse, focusing on both refugees and economic migrants also infected political discourse with a number of news outlets as well as leading ministers being called out for fuelling anti-immigrant sentiment (Creighton & Jamal, 2022).

3.2. Spain

Immigration has not featured among the top concerns of the Spanish public but has been placed onto the agenda by far-right party VOX repeatedly making statements and messages vilifying immigrants, specifically those from Muslim countries. This political strategy has incited extremist attitudes in some circumstances. On November 4, 2019, in Sevilla, Rocío Monasterio,

VOX's leader, visited a centre for unaccompanied minors in La Macarena to condemn young immigrants as provoking "insecurity," generating "huge problems" for the neighbourhood and creating an "unsustainable" situation. Although some traditional media and fact-checking platforms, such as *Maldita.es* ("Los bulos que ha usado VOX," 2019) and *Newtral* (González, 2019), denounced Monasterio's falsehoods, her hate speech fuelled xenophobic attacks in some neighbourhoods where migrant populations resided. The Office of the Prosecutor is yet to declare whether Monasterio's speech constituted an incitement to racial hatred.

In early 2021, VOX placed posters inside the Puerta del Sol station (Madrid), claiming: "An unaccompanied minor, [earns] 4,700 euros per month. Your grandmother's pension, [is] 426 euros per month." Even though the claim is untrue and could incite hatred, the Spanish justice system dictated it was not a hate crime as it was political propaganda. For these reasons, organizations such as UNICEF or International Amnesty (Equipo de Migración y Refugio de Amnistía Internacional, 2021), have alerted that unaccompanied minors have been targets of disinformation in Spain, enabling hate speech and discrimination. Even the term "MENA" (after *menor extranjero no acompañado*, or unaccompanied foreign minors in English) has become a pejorative term due to its framing by VOX as a synonym for criminality (Rubio Hancock, 2019).

Through their strategy, VOX has become the third force in the Spanish Congress (Martín Plaza, 2019) and gained popularity on social networks (Aladro Vico & Requeijo Rey, 2020). VOX's followership on Instagram overtook Podemos and has four times the numbers of Ciudadanos, PP, and PSOE. Although currently experiencing an internal crisis (López Agudín, 2022), VOX's popularity among particular societal groups is growing, mainly white males of the millennial generation, indicating they attract voters disenchanted with the political establishment (Morillo, 2022). Immigration, income, and conservative traditions are the party's central issues (Aragó, 2019) and it gains support in cities and towns with higher rates of immigration from outside the EU. Indeed, VOX has become ingrained in areas that have not developed their own political identity, e.g., Murcia, Almería, and Castilla-La Mancha (Aragó, 2019), though it struggles to gain supporters where there is a consolidated political identity, such as Catalonia, the Basque Country, Navarra, or Galicia. VOX also lure voters away from the PP as well as Ciudadanos (Lerín Ibarra, 2022). In addition, VOX's voters are usually Catholics and frequent churchgoers who strongly identify with the Spanish national state; they also have lower education levels but a higher income (Lerín Ibarra, 2022). Therefore, VOX appears to gain similar supporters to former US President Trump.

4. Methodology

We conducted a qualitative analysis involving the close reading of comments of posts that attracted the most intense discussion (multiple comments by different users) on pages set up to be supportive of UKIP and VOX on social media platforms (Facebook and Reddit in the case of the UK; Forocoches and Reddit for the Spanish context). These pages are public: In the case of Reddit they are accessible to anyone visiting the site; for Facebook groups you can view the posts independently or if you are a member of a group. While we recognise that a minority of social media users follow any political organisations—few of them follow parties and fewer follow more extremist parties—their discourse is argued to gain high traction due to the shareability of their content (Bennett & Livingston, 2018). Hence, focusing on the content of the pages of these parties provides insights into the more extreme content that might be available online. Given the debates around immigration in both countries and high-profile cases where politicians have been accused of utilising pejorative language, it is useful to detect how supporters of the most extreme parties construct arguments that may filter into online discourse. Hence our qualitative analysis of discourse is designed to identify the predominant frames in user comments that contained hate speech against migrants on these party's pages on popular platforms. All posts referencing migrants, immigrants, or immigration in English or Spanish were collected from the pages of the parties on both platforms from January 2020 to December 2022. Posts were selected purposefully (based on them receiving more than 100 comments) and then a random sample of 50 per party and per platform, a total of 200 posts and comment threads, was selected for discourse analysis, which determined the most common terms and narratives used referring to immigrants (Laver et al., 2003). Although there were intuitive expectations, we did not develop hypotheses regarding the frames that would be discovered. Instead, we chose to assess the patterns that existed within the discourse of these users without developing predetermined categories. The reason for this strategy is that language is an essential element that allows us to conceptualize everyday life. This implies, therefore, that commenters innately introduce their convictions and perceptions in the construction of even the most straightforward arguments (Lind & Salo, 2006). Hence, the analysis is a reflection of how these users interpret their reality and construct shared frames, in particular how their framing meant them "identifying a type of object or experience by emphasising certain properties, delocalising others, and obscuring others'' (Lakoff & Johnson, 2001, p. 205). Due to the selection of the overall sample of posts by topic and then random selection of posts due to user comments, we cannot make claims about their representativeness of the attitudes of the supporters of these parties. But, our analysis allows us to understand the dynamics of discourse, which is triggered by the topic of immigration in a way that cannot be reliably determined using automated text analysis programmes (Angus et al., 2013). The qualitative approach enables us to develop an analytical

narrative regarding references to immigrants as well as offer examples of the discourse produced by the more highly engaged followers of these parties. We set out the results in a thematic narrative identifying the prevalent common frames prior to drawing the data together within our conclusion. To comply with research ethics policies, no names are used to attribute quotes despite the pages being publicly viewable.

5. Results

The pages we draw our data from are vibrant communities. UKIP's Facebook page has 484,000 followers and VOX 491,000. Across the posts selected, there were over 800 unique usernames who left comments, suggesting many followers monitor the pages, or the posts that are made visible through the Facebook algorithm, and given the nature of the comments many of them respond when the posts trigger an emotional reaction. UKIP's Facebook page has declined in support since Brexit and the departure of its leader and main publicist Nigel Farage. The page is almost entirely dedicated to commentary on the UK border, the relationship between the UK and the EU, and the nature and composition of UK society. Tellingly one of the highest-rated posts is a screenshot of a list of ethnic groups arranged alphabetically with the accusation that "white" being last is an example of "woke" culture. There is also a degree of transphobia and misogyny expressed by some commenters as part of a general anti-woke campaign which references the evils of cancel culture that those designated as woke are imposing on society. These terms are used widely as criticisms of events that are described without context. Reddit discussions are less easy to locate, with the UKIP community page having been dormant for some years. However, references to the party, which often combines them with more fringe far-right groupings such as the British Democrats and English Democrats, suggest the far-right represents the only voices of truth and common sense. The frames that are found within the selected posts across both sites and nations are set out below.

5.1. Framing Immigrants: Leaches and Criminals Who Do Not Belong

The least tendentious frame used when immigration relates to them being a drain on resources. Posts will frame the housing of migrants as an injustice and comments will then be supportive of that frame. Initially, comments can use a measured town and highlight reasonable concerns regarding the strains on public services. But there is a clear juxtaposition between the undeserving outsiders and the deserving British or Spanish. The latter group is framed as increasingly marginalized—economically or socially—because "the system" is biased in favour of outsiders who should not have rights. Hence criticisms couched within valid concerns incite other users to inflame the argument using pejorative terminology. Immigrants are described as taking jobs away from British workers or taking advantage of Spanish public services, such as health assistance or education. Anecdotal evidence is used widely to reinforce these points. Stories are told of "a friend who lost his job" or "a friend who knows someone" who did not receive urgent hospital care because doctors were treating migrants instead. On the VOX Reddit page users complained that Spanish families did not receive aid to buy school textbooks, while the children of immigrants, who "spurn Spanish education and not even go to classes" (*desprecian la educación española y no van ni a clase*), receive free books. The comments relating to this frame are mixed, some promote hostility by framing immigrants as competitors for resources who are advantaged by the system, but some comments do include pejorative terms. "Leeches" is a common term, suggesting immigrants "suck" resources out of society but give nothing back. Some claim immigrants actively seek to steal resources. Hate speech is used in both VOX communities displaying anger and indignation when claiming Spanish citizens lack decent housing because the state supposedly gives away houses to illegal immigrants; or for not having work because immigrants steal jobs. VOX communities argue that immigrants arrive to claim benefits or take money from the system to send to their home countries are often referenced as facts with minimal evidence. There is also evidence of prejudice against immigrants of colour. Some commenters question the housing of Ukrainian refugees as further stretching resources, but this policy is defended on the grounds their stay is temporary as well as claiming Ukrainian refugees arrive with values closer to those of the host nation as this argument suggests:

> A second-generation African who fails at school and cannot get a job is more likely to join the ranks of BLM and other radical leftist organisations in the belief that he has been the victim of discrimination, than will a second-generation Ukrainian who will be more likely to attribute his failure to either his level of ability or the amount of effort he has made to succeed.

The frame, however, emphasizes the notion that only those who "belong" in each nation should be entitled to societal benefits and, therefore, cutting immigrants' rights to access public services is necessary. The notion of immigrants deliberately arriving to "steal" benefits they have not earned is made explicit at points, as in the argument that should the state "cut their benefits...they won't come."

The second prevalent frame involves arguing that immigrants do not possess the same values as the British or Spanish and so they will not follow recognised norms and laws of society. On one level their lack of loyalty to the nation positions them as willing to exploit the benefits of being in the nation, at another it frames all immigrants as potential criminals. The claim is most

explicitly made in comments on VOX's pages. Consistent with the party line, users claim illegal migrants "invade" Spain to commit serious crimes; therefore, Spain must "expatriate them to guarantee that Spanish people live safely." In the UK and Spain criticism is levelled at the European international cooperation, users claim it is a mechanism that encourages illegal immigration mafias which encourage massive illegal migratory flows and act as recruiters for jihadist groups. On Reddit communities, conspiracy theories and disinformation usually feed hate speech. There are claims that illegal immigrants who arrived in dinghies, *pateras* in Spanish, were guilty of spreading coronavirus all over the country. Forocoches users offer anecdotes describing crimes committed by MENA. These are presented as evidence that the Spanish Criminal Code must be revised urgently so there is harsher sentencing. In Ceuta and Melilla, it is argued that it is imperative to "build a wall" to prevent people from Morocco from invading Spanish territories. On UKIP pages, any stories of people of colour committing crimes are deployed as evidence of the dangers of migration independent of the actual backgrounds of those accused and/or convicted. Any evidence is used to promote an anti-immigration and xenophobic narrative, with comments on shared news items adhering to an exclusionary line. As one user exclaimed regarding a story about two shop owners with Asian names who were found guilty of sexually assaulting an underage boy in their storeroom: "Dirty fucking scum!!! Then people wonder why we don't want them here."

There is clear evidence that isolated incidents involving individuals that can be identified as having non-British origins are used to tarnish the reputations of all from similar backgrounds. Within this frame, there are also claims that certain areas of cities, in particular London, have become ghettos for immigrant communities where "it is no longer safe for white people to go." Claims that parks can no longer be used "unless you're a n*****r" offer the impression that there are deep racialised divides within cities with communities co-existing antagonistically as opposed to integrating. This theme, which focuses on a lack of shared norms and values, is most prominent on Reddit and Forocoches. Alongside evidence of criminality, there are comments that immigrants violate a range of norms, and claims that they are uncivilised and likely to abuse women and children. This frame also incites the strongest forms of hate speech. There is frequent use of terms such as "dirty," "plague-ridden," or "scum," there are frequent examples of dehumanisation, comparing them to rats, plague rats, vermin, etc.

The third common frame explores the notion of the clash of civilisations, is largely Islamophobic but suggests most migrants of colour are from a homogenous, alien culture which links to a longstanding conspiracy theory that coloured races are attempting to replace white European populations. Some stories seem anodyne, such as the post reporting the closure of a British public house and conversion to a Mosque. But comments quickly reference the conspiratorial metanarrative, for example: "Typical, they want to destroy every aspect of white British culture."

Similarly focusing on the clash of cultures, many users comment negatively about the cruelty of preparation of halal meat, cultural norms, around arranged marriages, and gender segregation practised in Mosques presenting these as evidence Muslims pose an existential threat to European society. Hence a key feature is an Islamophobic discourse and the expression of strong Islamophobic attitudes. It is common to find pejorative expressions on Forocoches and Reddit such as *moros de mierda* ("shitty Moors") or in the UK to suggest all Muslims are potential rapists or terrorists. The notion of a new "Moorish" invasion is discussed in Reddit communities, referencing the Reconquest of Spain in 1492. Comments such as "We Spaniards did not reconquer Spain for nothing" (*Los españoles no reconquistamos España para nada*) highlight some feel a new "reconquest" is needed, giving strength to the motto "Make Spain great again!" (*Hagamos España grande otra vez!*), echoing Trump's and VOX's electoral propaganda.

The Islamophobic narrative links to a wider rejection of multiculturalism. Forocoches users criticize all other ethnic groups, their language, customs, and religion. Consistent with VOX campaign themes they promote the homogeneity of the nation and like many populist far-right groups claim all foreign influence to be a threat (Carter & Pérez, 2015; Mudde, 2000). These influences range from Brussels and the European Parliament as well as the more nebulous forces of multiculturalism and globalism. The discourse indicates an aspiration that all citizens must share the same national origin and ethnic features favouring ethnic Hispanicism. These statements of preference encompass expressions like "Spanish people first" (*los españoles primero*) but do not necessarily involve offensive language or insults toward migrants; they can however incite tensions between communities.

An interesting example of how ethnicity and culture become prominent was found across the UKIP pages during the September 2022 Conservative leadership campaign with the final choice being between British Asian Rishi Sunak and White British female Liz Truss. Beyond the comment that Britain was "no longer a place for white men," concerns were raised over what Sunak might legalise. The comment "will I be allowed loads of wives" was tongue in cheek but indicative of the notion Sunak might impose non-British norms on UK society. Despite being of Indian Hindu descent, some asked whether Sunak would open the door to imposing Sharia law in Britain, when challenged on that commenters responded "they're all the same, that lot" suggesting any person of colour has alien values. Similarly, arguments suggested that people of colour stick together and support one another, so disadvantaging the white British. This frame exacerbates perceptions of racial differences while at the same time offering no differentiation of people of colour

independent of their background. In this vein, it was expected that such figures would be softer on immigration, although many commenters referenced the introduction of harsh environmental policies by former Home Secretary Priti Patel, including a plan to remove illegal immigrants and house them in Rwanda. But racial solidarity was even referenced in this case, as her policies "only involved Africans…bet she opens the door if they're Pakis"; despite the fact that Patel is Ugandan-Indian and not Pakistani. The heritage of Patel and Sunak is correctly labelled by one commenter, claiming they were part of a "Hindu mafia taking over our country." The outsider narrative remains strong and users consistently questioned the motives of any person of colour, promoting a culture of suspicion of those of non-British racial heritage by classifying them as un-British.

6. Conclusion

The study of hate speech towards immigrants in extremist online communities has provided a valuable framework for understanding the ideological paradigm of these groups. Results show that hate speech towards migrants in the UK and Spain share keyframes. Immigrants are framed as a drain on resources, illegal, potential, or actual criminals, and discourse is coloured by the worship of traditional social norms and national customs—conventionalism. The strident language used represents an aggressiveness towards otherness which leads to homogenous thinking (Rogers & Kincaid, 1981), the need to transfer strength to the nation and the communities or leaders who are like-minded and defend traditional homeland values, and a closed or dogmatic mentality (Rokeach, 1960). These attitudes and beliefs may be expressed in such bold forms in ghettoes within the social media environment, but likely reflect a wider current of thinking within both societies. These ideas can also be transferred out of these communities and spread more widely across social media platforms and in some cases become tropes in some mainstream media outlets. Evidence suggests that exposure to negative stories relating to immigration can lead to the development of more extremist attitudes as users view increasingly extreme content (Mihelj & Jiménez-Martínez, 2021). Research suggests that the attitudes that lead to the use of hate speech can become more widespread normalising the attitudes and potentially the use of hateful frames when referencing any individual perceived as being of another race, ethnicity, or culture.

It may seem surprising that extreme language can be found on social media platforms. However, as noted, even where automated systems prevent such expressions they can be circumnavigated. Moderation remains valuable and proved crucial in fighting disinformation during the Covid-19 pandemic (Pérez-Escolar & Herrero-Diz, 2022); however, it is a blunt instrument that is limited both by the constraints of automated text detection as well as by the implementation policies of platform owners. Therefore, legislation on social media platforms can only have a limited impact. Censorship, such as the bans imposed on Trump and other individuals, can only push such attitudes to the fringes, although the less visible such arguments are the less they can be normalised. Avoiding normalisation is crucial as it can lead to widespread societal rejection of any representations of groups which are routinely framed as the other. Censorship can also, to some extent, ensure the claims of far-right voices do not feature in mainstream media reporting. If their voices do not appear on mainstream channels they are less likely to be seen, shared, or promoted by bots, and so remain ghettoed. But, to prevent hate speech requires an eradication of both the attitudes and conditions which provide succour to such arguments. Firstly, this means correcting xenophobic and other discriminatory attitudes within society. Our data suggest these attitudes exist and that there is a strong emotional attachment to the frames constructed around certain migrant groups, hence these societies would appear to contain a serious threat to multiculturalism and inclusion. Secondly, however, it is important to develop strategies that will correct the societal inequities that fuel anger and which can be channelled and given a voice by rightist extremists. Without these corrections it is hard to see a way by which the attitudes that are expressed in hate speech can be eradicated.

Conflict of Interests

The authors declare no conflict of interests.

References

Aladro Vico, E., & Requeijo Rey, P. (2020). Vox political party discourse, strategies and interactions on its official Instagram account during the A-28 elections. Radical right and social networks. *Revista Latina de Comunicación Social*, *2020*(77), 203–229. https://www.doi.org/10.4185/RLCS-2020-1455

Almoqbel, M. Y., Wohn, D. Y., Hayes, R. A., & Cha, M. (2019). Understanding Facebook news post comment reading and reacting behavior through political extremism and cultural orientation. *Computers in Human Behavior*, *100*, 118–126.

Angus, D., Rintel, S., & Wiles, J. (2013). Making sense of big text: A visual-first approach for analysing text data using Leximancer and Discursis. *International Journal of Social Research Methodology*, *16*(3), 261-267.

Aragó, L. (2019, December 4). Los distintos perfiles del votante de Vox [The different profiles of the Vox voter]. *La Vanguardia*. https://cutt.ly/1V6M9MX

Armstrong, G. M., & Wronski, J. (2019). Framing hate: Moral foundations, party cues, and (in)tolerance of offensive speech. *Journal of Social and Political Psychology*, *7*(2), 695–725.

Atkins, J. (2018). "Strangers in their own country": Epideictic rhetoric and communal definition in Enoch Pow-

ell's "Rivers of Blood" speech. *The Political Quarterly*, *89*(3), 362–369.

Auxier, B. E., & Vitak, J. (2019). Factors motivating customization and echo chamber creation within digital news environments. *Social Media + Society*, *5*(2), 1–13.

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139.

Bolsover, G. (2020). *Black Lives Matter discourse on US social media during Covid: Polarised positions enacted in a new event*. ArXiv. https://arxiv.org/abs/2009.03619

Brown, A. (2018). What is so special about online (as compared to offline) hate speech? *Ethnicities*, *18*(3), 297–326.

Cammaerts, B. (2022). The abnormalisation of social justice: The "anti-woke culture war"discourse in the UK. *Discourse & Society*, *33*(6), 730–743.

Cappella, J. N., & Jamieson, K. H. (1996). News frames, political cynicism, and media cynicism. *Annals of American Academy of Political and Social Science*, *546*(1), 71–84.

Carter, N. M., & Pérez, E. O. (2015). Race and nation: How racial hierarchy shapes national attachments. *Political Psychology*, *37*(4), 497–513. https://doi.org/10.1111/pops.12270

Creighton, M. J., & Jamal, A. A. (2022). An overstated welcome: Brexit and intentionally masked anti-immigrant sentiment in the UK. *Journal of Ethnic and Migration Studies*, *48*(5), 1051–1071.

Emcke, C. (2019). *Against hate*. Polity.

Entman, R. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, *43*(4), 51–58.

Equipo de Migración y Refugio de Amnistía Internacional. (2021, May 22). Problemas a los que se enfrentan los MENAS. Desmontando bulos [Problems faced by unaccompanied minors. Debunking hoaxes]. *Amnistía Internacional*. https://cutt.ly/tV6R8W8

Ernst, N., Engesser, S., Büchel, F., Blassnig, S., & Esser, F. (2017). Extreme parties and populism: An analysis of Facebook and Twitter across six countries. *Information, Communication & Society*, *20*(9), 1347–1364.

Fenoll, V., & Rodríguez-Ballesteros, P. (2017). Análisis automatizado de en-cuadres mediáticos. Cobertura en prensa del debate 7D 2015: El debate decisivo [Automated analysis of media in-frames. Press coverage of the 7D 2015 debate: The decisive debate]. *El profesional de la información*, *26*(4), 630–640.

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. UNESCO. https://cutt.ly/NVhWnuQ

Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, *95*(1), 1–37.

Gelber, K., & McNamara, L. (2016). Evidencing the harms of hate speech. *Social Identities*, *22*(3), 324–341.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harper & Row.

González, S. (2019, November 6). Los datos oficiales niegan la alta delincuencia de los menores extranjeros no acompañados en Andalucía [Official data denies high delinquency of unaccompanied foreign minors in Andalusia]. *Newtral*. https://cutt.ly/LV5JrHw

Hamed, D. M. (2020). Ideologically-based polarisation and racism in discourse. *Journal of Law and Social Sciences*, *4*(1), 1–13.

Herrero-Diz, P., & Pérez-Escolar, M. (2022). Análisis de los bulos sobre Covid-19 desmentidos por Maldita y Colombiacheck: Efectos de la infodemia sobre el comportamiento de la sociedad [Analysis of Covid-19 hoaxes disproved by Maldita and Colombiacheck: Effects of infodemics on societal behavior]. *Palabra Clave*, *25*(1), Article e2517. https://doi.org/10.5294/pacla.2022.25.1.7

Jardine, E. (2019). Online content moderation and the dark web: Policy responses to radicalizing hate speech and malicious content on the darknet. *First Monday*, *24*(12). https://doi.org/10.5210/fm.v24i12.10266

Jessoula, M., Natili, M., & Pavolini, E. (2021). "Exclusionary welfarism": A new programmatic agenda for populist right-wing parties? *Contemporary Politics*, *28*(4), 447–468.

Jougleux, P. (2022). *Facebook and the (EU) law*. Springer.

Lakoff, G., & Johnson, M. (2001). *Metáforas de la vida cotidiana* [Everyday life metaphors]. Cátedra.

Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from poli-tical texts using words as data. *American Political Science Review*, *97*(2), 311–331.

Lerín Ibarra, D. (2022). *La emergencia política de la derecha radical en España. El caso de Vox* [The political emergence of the radical right in Spain. The case of Vox] [Doctoral dissertation, Universidad Complutense de Madrid]. E-Prints Complutense. https://eprints.ucm.es/id/eprint/73525

Lilleker, D. G., & Bonacci, D. (2017). The structure of political e-expression: What the Brexit campaign can teach us about political talk on Facebook. *International Journal of Digital Television*, *8*(3), 335–350.

Lilleker, D. G., & Jackson, D. (2017). The social media campaign: Mobilisation and persuasion. In D. Wring, R. Mortimore, & S. Atkinson (Eds.), *Political communication in Britain* (pp. 293–313). Palgrave Macmillan.

Lilleker, D. G., & Ozgul, B. A. (2021). *The psychology of democracy*. Routledge.

Lin, C. S. L. (2016). Convergence in election campaigns: The frame contest between Facebook and mass media. *Convergence: The International Journal of Research Into New Media Technologies*, *22*(2), 199–214.

Lind, R. A., & Salo, C. (2006). The framing of feminists and feminism in news and public affairs programs in U.S.

electronic media. *Journal of Communication*, *54*(1), 211–228.

Lippmann, W. (1922). *Public opinion*. Free Press.

López Agudín, F. (2022, September 27). La crisis anunciada a imparable de Vox [The announced and unstoppable crisis of Vox]. *Público*. https://cutt.ly/6V8EnJQ

Los bulos que ha usado VOX para hablar de inmigración: Manadas, ayudas al alquiler, Mena y "paguitas" de 664 euros [The hoaxes VOX has used to talk about immigration: Herds, rent subsidies, unaccompanied minors and "pays" of 664 euros]. (2019, November 11). *Maldita.es*. https://cutt.ly/CV5JaOu

Lynch, K. L. (2022). Rereading black like me: Speech matters, context matters. In J. P. Messina (Ed.), *New directions in the ethics and politics of speech* (pp. 91–109). Routledge.

Martín Plaza, A. (2019, November 11). Vox se dispara hasta colocarse como tercera fuerza con 52 diputados, más del doble que en abril [Vox soars to become the third force with 52 deputies, more than double the number in April]. *RTVE*. https://cutt.ly/CV8Q0sk

Mihelj, S., & Jiménez-Martínez, C. (2021). Digital nationalism: Understanding the role of digital media in the rise of "new" nationalism. *Nations and Nationalism*, *27*(2), 331–346.

Mondal, M., Silva, L. A., & Benevenuto, F. (2017). A measurement study of hate speech in social media. In *HT '17: Proceedings of the 28th ACM conference on hypertext and social media* (pp. 85–94). Publication History. https://doi.org/10.1145/3078714.3078723

Morillo, I. (2022, January 1). Vox consolida su crecimiento en el voto joven, punto débil del bipartidismo [Vox consolidates its growth in the young vote, a weak point of the bipartisan system]. *El Confidencial*. https://cutt.ly/YV4IIe9

Mudde, C. (2000). *Ideology of the extreme right*. Manchester University Press.

Murphy, J., & Devine, D. (2020). Does media coverage drive public support for UKIP or does public support for UKIP drive media coverage? *British Journal of Political Science*, *50*(3), 893–910.

Nguyen, A., & Vu, H. T. (2019). Testing popular news discourse on the "echo chamber" effect: Does political polarisation occur among those relying on social media as their primary politics news source? *First Monday*, *24*(6). https://doi.org/10.5210/fm.v24i6.9632

Pérez-Escolar, M., & Herrero-Diz, P. (2022). The role of wikipedia in the dissemination of new knowledge: Analysis of the entry "desinformación" as a changeable concept. *Communication & Society*, *35*(2), 257–268. https://doi.org/10.15581/003.35.2.257-268

Pérez-Escolar, M., & Noguera-Vivo, J. M. (2022). *Hate speech and polarization in participatory society*. Routledge. https://cutt.ly/CCGhHBE

Popper, K. (1996). *The open society and its enemies*. Routledge.

Qin, J. (2015). Hero on Twitter, traitor on news: How social media and legacy news frame Snowden. *The International Journal of Press/Politics*, *20*(2), 166–184.

Renton, D. (2018). *Never again: Rock against racism and the anti-nazi league 1976–1982*. Routledge.

Rieger, D., Kümpel, A. S., Wich, M., Kiening, T., & Groh, G. (2021). Assessing the extent and types of hate speech in fringe communities: A case study of alt-right communities on 8chan, 4chan, and Reddit. *Social Media + Society*, *7*(4), 1–14. https://doi.org/10.1177/20563051211052906

Rogers, E., & Kincaid, L. (1981). *Communication networks. Towards a new paradigm for research*. Free Press.

Rokeach, M. (1960). *The open and closed mind: Investigations into the nature of belief systems and personality systems*. Basic Books.

Rossini, P. (2020). Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk. *Communication Research*, *49*(3), 1–27. https://doi.org/10.1177/0093650220921314

Rubio Hancock, J. (2019, November 14). How the term "MENA" hurts unaccompanied migrant minors in Spain. *El País*. https://acortar.link/HAQR9K

Sádaba, T. (2006). *Framing: Una teoría para los medios de comunicación* [A theory of media]. Ulzama Digital.

Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of Communication*, *49*(1), 103–122.

Schindler, M., & Domahidi, E. (2021). The growing field of interdisciplinary research on user comments: A computational scoping review. *New Media & Society*, *23*(8), 2474–2492.

Scott, J. (2022, August 28). More than 25,000 migrants have crossed channel this year, latest figures reveal. *Sky News*. https://cutt.ly/CV5JaOu

Sheckels, T. F., Gutgold, N. D., & Carlin, D. B. (2012). *Gender and the American presidency: Nine presidential women and the barriers they faced*. Lexington Books.

Smith, J. (2018). Gambling on Europe: David Cameron and the 2016 referendum. *British Politics*, *13*(1), 1–16.

Springer, N., Engelmann, I., & Pfaffinger, C. (2015). User comments: Motives and inhibitors to write and read. *Information, Communication & Society*, *18*(7), 798–815.

Strand, C., & Svensson, J. (2019). Towards a situated understanding of vulnerability: An analysis of Ugandan LGBT+ exposure to hate crimes in digital spaces. *Journal of Homosexuality*, *40*(4), 77–95. https://doi.org/10.1080/23743670.2019.1665565

Suhay, E., Brandt, M. J., & Proulx, T. (2017). Lay belief in biopolitics and political prejudice. *Social Psychological and Personality Science*, *8*(2), 173–182.

Takeshita, T. (1997). Exploring the media's roles in defining reality. From issue-agenda setting to attribute-agenda setting. In M. E. McCombs, D. L. Shaw, & D. H. Weaver (Eds.), *Communication and democracy: Exploring the intellectual frontiers in agenda-setting theory* (pp. 15–27). Lawrence Erlbaum.

Tuters, M., & Hagen, S. (2020). (((They))) rule: Memetic antagonism and nebulous othering on 4chan. *New Media & Society*, *22*(12), 2218–2237. https://doi.org/10.1177/1461444819888746

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Wardle, C., & Derakhshan, H. (2017). *Information disorder. Toward an interdisciplinary framework for research and policymaking*. Council of Europe. https://bit.ly/2OsHEHY

Wasilewski, K. (2019). Hate speech and identity politics. An intercultural communication perspective. *Przegląd Europejski*, *3*(53), 175–187.

Weber, P., Prochazka, F., & Schweiger, W. (2017). Why user comments affect the perceived quality of journalistic content. *Journal of Media Psychology*, *31*(1), 24–34.

Wetherell, M. (2008). Speaking to power: Tony Blair, complex multicultures and fragile white English identities. *Critical Social Policy*, *28*(3), 299–319.

Williams, M. (2021). *The science of hate: How prejudice becomes hate and what we can do to stop it*. Faber & Faber.

Zurutuza-Muñoz, C., & Lilleker, D. (2018). Writing graffiti on the Facebook wall: Understanding the online discourse of citizens to politicians during the 2016 Spanish election. *Communication & Society*, *31*(3), 27–43.

About the Authors

Darren Lilleker is professor of political communication at Bournemouth University. His expertise is political campaigning and public engagement in politics, in particular the positive and negative impact of digital technologies on political engagement and participation. He has published widely on political communication and its impacts. Past works include *Political Communication and Cognition* (Palgrave, 2014) and *The Psychology of Democracy* (Routledge, 2022). He is co-editor of the *Journal of Visual Political Communication*.

Marta Pérez-Escolar is assistant professor in the Department of Communication at Loyola University (Seville, Spain). Her research emerges from the synergy of civic and political participation, public opinion, disinformation, hate speech, and ideology. She has recently published *Hate Speech and Polarization in Participatory Society* (Routledge, 2022) and obtained a José Castillejo International Mobility Stays for Young Post-Doctoral Researchers in 2022, where she spent six months in a research stay at Bournemouth University (UK).

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 138–146
https://doi.org/10.17645/pag.v11i2.6306

Article

The Side Effect of Political Standing: Corporate Activism and Its Impact on Stock Returns

Teresa Pintado Blanco [1], Carlota López-Aza [1], Joaquín Sánchez [1], and Pablo Gutiérrez-Rodríguez [2,*]

[1] Department of Marketing, Universidad Complutense de Madrid, Spain
[2] Department of Business Administration, Universidad de León, Spain

* Corresponding author (pablo.gutierrez@unileon.es)

Submitted: 21 October 2022 | Accepted: 12 January 2023 | Published: 17 May 2023

Abstract
Taking sides on controversial political issues such as gun control, abortion, immigration, or diversity is increasingly common among large companies. What remains unclear, however, is whether this type of strategy—known as "corporate activism"—has positive or negative effects on these companies. The use of the concept of corporate activism on different variables affecting the companies is relatively recent. This article analyses the effect of corporate activism on the stock market performance of US companies through the analysis of the sample collected. Although there are some recent articles published on this topic, none of them measures the risk associated with the use of this type of strategy. For the development of the research, the well-known Fama–French modelling framework is applied to estimate the differences between companies that participate in corporate activism initiatives versus those that remain outside this strategy. The findings complement previous research showing that companies that use corporate activism have lower market risk than companies that do not engage in this type of strategy. These results can be useful in identifying the advantages and disadvantages of corporate activism initiatives and, in addition, they can also help companies to evaluate the use of corporate activism as a strategic tool and as a driver of social change.

Keywords
corporate activism; corporate social advocacy; corporate social responsibility; firm value

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

Controversial political issues have been on the rise dramatically in recent years. For example, after the US Supreme Court struck down abortion rights in the US last June, many US companies (Apple, Microsoft, Meta, Yelp, Netflix, Uber, Warner Bros, Levi Strauss, Bumble, etc.; Duffy & Korn, 2022) took a stand, offering to provide support and financial assistance for their employees if they needed this procedure. Such immediate actions would have been unthinkable a few years ago, as companies did not meddle in controversial issues. However, currently, companies are increasingly called upon to speak out on polarising issues such as immigration, gun control, climate change, and LGBTQ rights (Kotler & Sarkar, 2018). These are issues of interest to millennials, who see that the current political system does not meet their demands, and therefore disengage from the political process (LaCombe & Juelich, 2019) by shifting their demands to private companies. According to the 2021 Edelman Trust Barometer (Edelman, 2021), 86% of consumers expect brands to take actions beyond their core business instead of governments. This strategy is known as corporate activism and consists of a "company's willingness to take a stand on social, political, economic, and environmental issues to create societal change by influencing

the attitudes and behaviours of actors in its institutional environment" (Eilert & Nappier Cherup, 2020, p. 3).

Adopting corporate activism initiatives involves a risk for the company because taking the wrong stand on a controversial issue, that is, a stand not aligned with consumers' attitudes, values, and opinions, produces a public backlash and a lack of consumer brand identification (Mukherjee & Althuizen, 2020). These partisan actions cause uncertainty for shareholders, too, as these actions move investors away from profit maximisation and change strategic priorities (Bhagwat et al., 2020). Thus, shareholders still are not clear on how to perceive and value the effects of corporate activism on financial results (Villagra, Monfort, & Méndez-Suárez, 2021).

In fact, even though corporate activism is attracting growing attention within the marketing research community and professional discourse (Villagra et al., 2022), the gap between this strategy and corporate finance results is not clear. In previous research, authors have viewed corporate activism as part of strategic issue management and have discussed its importance to companies (Dodd & Supa, 2014).

Some authors have pointed out the possibility that there may be an impact on financial results if a company takes a stand on a controversial issue (Korschun et al., 2019; Mukherjee & Althuizen, 2020). Other authors have gone a step further and reported a link between corporate activism and financial results, finding abnormal returns in different contexts (Bhagwat et al., 2020; Dodd, 2018; Dodd & Supa, 2014, 2015; Villagra, Monfort, & Méndez-Suárez, 2021; Weinzimmer & Esken, 2016). Previous studies do not agree on the results of this relationship and the aforementioned articles have found positive investor reactions to corporate activism actions (Haq et al., 2022; Weinzimmer & Esken, 2016), negative investor reactions (Bedendo & Siming, 2021; Bhagwat et al., 2020; Glambosky & Peterburgsky, 2022; Villagra, Monfort, & Méndez-Suárez, 2021), and non-significant investor reactions (Villagra, Monfort, & Méndez-Suárez, 2021).

This research empirically demonstrates the positive or negative financial effect of corporate activism on corporations. Using the Fama–French modelling framework from the field of finance, it is demonstrated that organisations that have implemented corporate activism actions are associated with changes in the stock market compared to companies not using an activist strategy.

The article will be divided as follows: Section 2 offers a more detailed review of the principles of corporate activism. Section 3 offers a summary of the existing studies, indicates gaps in the existing literature, and explains this study's contribution. Sections 4 and 5 provide the empirical evidence for the association between corporate activism and firm performance through the Fama–French methodology using a sample of the stock market performance of 96 US companies, and the final section includes conclusions, limitations, and future lines of research.

2. Principles of Corporate Activism

The phenomenon of corporate activism has been studied from different perspectives such as communication and public relations, and it is related to several concepts such as corporate social advocacy (CSA), corporate political activity (CPA), corporate social responsibility (CSR), and so on (Bhagwat et al., 2020).

The previous literature shows us that corporate activism is an evolution of CSR (Kotler & Sarkar, 2018) and is differentiated from CSR by the partisan nature of the causes supported (whether progressive or conservative) and by its focus on company values rather than the consequences for sales or other variables (Vredenburg et al., 2020). Consequently, corporate activism has much higher levels of risk than CSR actions (Bhagwat et al., 2020) and as a result, controversial issues have an impact on financial outcomes (Dodd & Supa, 2014).

These authors (Dodd, 2018; Dodd & Supa, 2014) define CSA as controversial corporate socio-political initiatives beyond CSR actions in which firms or their CEOs take a stand aligned with their values, whether intentionally or not. This concept arises due to changes in society's expectations whereby instead of demanding these social changes from governments, they demand them from companies and their leaders (Dodd, 2018; Hoppner & Vadakkepatt, 2019). This concept has been increasingly developed over the past few years and it is often included within public relations, bringing together different key areas: CSR, problem management, and strategic issue management (Dodd & Supa, 2014). Thus, CSA is linked to strategic business planning and social responsibility practices, which affect companies' stakeholders (Heath & Palenchar, 2008).

Another theoretical approach related to corporate activism is CPA. CPA is defined as public support for individuals, groups, ideals, or values that seek to persuade others to do the same (Wettstein & Baur, 2016) and is therefore similar to the concepts described above: CSR and CSA. However, CPA has a more radical purpose, abandoning consensus in communication and transcending the economic commitments of the organisation (Ciszek & Logan, 2018; Ferguson, 2018), because it is focused on actions that are seen as a driver of social change (Ciszek & Logan, 2018). Like the previous concepts, it has an important relationship with public relations and can be considered a form of activism. Furthermore, political activity acts as a buffer mechanism for companies developing CPA strategies, preventing them from entering into agreements with external social activists who try to influence their business policies (Hadani et al., 2019).

For their part, business leaders are increasingly involved in corporate activism by taking stances on matters of current social or political debate, with the primary aims of visibly weighing in on the issue and influencing opinions in the espoused direction (Hambrick & Wowak, 2021, p. 34). This is known as the CEO activism approach and it can be beneficial to companies (Chatterji

& Toffel, 2016; Krebel Chang, 2017). Research on CEO activism is growing rapidly (Branicki et al., 2021). For example, Korschun et al. (2019) analysed the relationship between CEO activism actions and consumer perception, finding a positive relationship between whether consumers' values matched those supported by the company. Other research classified different types of CEO according to their degree of morality and the level of corporate self-interest related to the issue and studied if CEO activism is a genuine ethical practice or not (Branicki et al., 2021). Additionally, Hambrick and Wowak (2021) presented a CEO activism model that explores whether a CEO's position is motivated solely by their values or if it is also moderated by employees' and customers' values. Lastly, Bedendo and Siming (2021) explored the relationship between CEO activism and the evolution of companies' shares.

In summary, previous empirical literature has focused on the relationship between corporate activism and purchase intention (Corcoran et al., 2016; Krebel Chang, 2017; Overton et al., 2021; Robinson et al., 2012), changes in consumer attitudes (Atanga et al., 2022; Parcha & Kingsley Westerman, 2020), the impact on brand equity (Korschun et al., 2019; Vredenburg et al., 2020), reputation (Den-Hond et al., 2014), or both (Villagra, Clemente-Mediavilla, et al., 2021), as well as understanding the antecedents of the concept (institutional and corporate credibility and authenticity) from the consumer's point of view (Villagra et al., 2022). Moreover, the relationship between corporate activism and financial performance is receiving research attention (Bhagwat et al., 2020; Dodd, 2018; Dodd & Supa, 2014; Glambosky & Peterburgsky, 2022; Villagra, Monfort, & Méndez-Suárez, 2021; Weinzimmer & Esken, 2016), which is closely related to the principal objective of this research.

3. The Relationship Between Corporate Activism and Firm Performance

There are currently not many articles that focus on analysing how corporate activism can affect the financial performance of companies (Bhagwat et al., 2020; Villagra, Monfort, & Méndez-Suárez, 2021). Some studies have focused on theoretical approaches that have had no subsequent empirical application (Eilert & Nappier Cherup, 2020; Vredenburg et al., 2020) and others have not provided conclusive results (Parcha & Kingsley Westerman, 2020).

One of the ways to analyse this type of relationship is to take consumer behaviour as a reference. Thus, to analyse how corporate stances on social or political issues affected the financial performance of companies, Dodd and Supa (2014) conducted an experimental study based on consumer purchase intentions, as they have a subsequent impact on sales and, therefore, on the economic performance of companies. To do so, they identified six companies (Starbucks, Chick-fil-A, Walmart, Whole Foods, Hobby Lobby, and Nike) that had publicly taken various stances on three controversial social issues: gay marriage, healthcare reform, and emergency contraception. Study participants were randomly exposed to these companies' messages and asked how these stances affected their purchase intentions, and whether their attitudes were congruent with what the companies had said. They then ranked the participants according to whether their attitudes were congruent or incongruent with the organisations' messages on these three topics and analysed whether this had significantly affected their purchase intentions. The results showed that participants had a higher purchase intention when they were exposed to messages that matched their own attitudes towards the social issues advocated by the organisations; therefore, a company's economic outcomes could be better in this situation, and worse if the messages did not match consumer attitudes.

Shortly thereafter, these same authors (Dodd & Supa, 2015) developed an article along the same lines, in which they again studied the impact of companies' social stances on consumers' purchase intention and, therefore, on their economic performance. In this case, the researchers identified two organisations that had taken public stances on same-sex marriage (Starbucks in favour, and Chick-fil-A against), and found that participants' purchase intention was more favourable toward companies that had a stance akin to their own, which generated better financial results. Subsequently, Dodd (2018) put forward research based on a theoretical conceptualisation of the role of corporate involvement in controversial socio-political issues within democratic societies. His work argues that the erosion of traditional institutions has led to companies playing an increasingly relevant role in decision-making on issues affecting society; therefore, companies are becoming increasingly politicised and this, in turn, affects their communications, which become of public interest.

On the other hand, Villagra et al. (2022) used a sample of 1,521 consumers to propose a theoretical model on the antecedents of corporate activism; these authors showed that institutional and corporate credibility and authenticity act as antecedents of this phenomenon, thus conditioning and explaining which circumstances contribute to the use of corporate activism. They also found that the higher the perceived credibility and authenticity of companies, the more positively their corporate activism actions will be regarded. In addition, Villagra, Clemente-Mediavilla, et al. (2021) found that consumers' political ideology could act as a moderating variable of the effects of corporate activism. On the one hand, consumers with a more conservative ideology do not consider that the activist actions of companies will have positive consequences on their reputation and brand value. However, consumers with a more liberal ideology do consider that activism can have positive effects on both aspects so that companies should be involved in this type of activist actions that transcend their traditional functions.

Furthermore, Weinzimmer and Esken (2016) studied how taking a particular stance on a sensitive social issue could affect companies' financial performance. According to these authors, this type of action by companies can confuse employees and can also impact corporations' image and influence the consumers' purchasing behaviour. In their work, Weinzimmer and Esken (2016) evaluated the advantages and disadvantages of adopting corporate positions on sensitive social issues, finding that sometimes what is important is not the position taken by the leader but the way in which the leader defends it. They also argued that before a company takes a stand on a social issue, it must use a strategic approach and understand the legal implications of its behaviour. However, such a stance on a controversial issue, and the way a leader defends it, can be significant and, therefore, the financial outcome could be positive.

Other authors have also studied how CEO activism can affect the value of corporate stocks. Thus, Bedendo and Siming (2021) analysed the resignation of a group of business leaders as advisors to President Trump. In this case, shareholders feared that their company would have less political influence, so they reacted negatively. On the other hand, the results of this work revealed that CEOs' public stances were driven to a greater extent by their personal ideology, while their companies' involvement was of lesser importance in such public stances.

Similarly, Bhagwat et al. (2020) confirmed that investors react unfavourably to companies that engage in controversial socio-political issues, as they move away from profit-oriented objectives towards risky activities with uncertain outcomes. The authors identified two moderators in this relationship: on the one hand, the deviation of these controversial actions from stakeholder values and brand image; on the other hand, the characteristics of the implementation of these actions, which can affect investor and customer responses. In their study, Bhagwat et al. (2020) analysed 293 controversial socio-political actions, in 149 companies from 39 different sectors, and their effects on several variables: the value for their audiences, the company's brand image, the size of the company, the actions or statements related to such socio-political issues, whether there was any specific statement from the CEO, or whether any communication on these issues was disseminated. These authors found that investors react unfavourably to companies that engage in controversial socio-political issues and that these investor reactions are worse if: (a) the controversial actions deviate from stakeholders' political values, (b) they take the form of actions (rather than statements), (c) they are announced by the CEO (rather than someone else in the company), (d) they do not explicitly communicate any business interests, and (e) it is an action by a single company (rather than a coalition with other companies).

Also, Glambosky and Peterburgsky (2022) studied how investors reacted to companies taking a stand against the Russia–Ukraine conflict by leaving the Russian country, which is a political corporate activism action. These authors found that markets react negatively to the company's announcement of divestment from Russia, but these activist companies recover their losses over the following two weeks. Furthermore, these authors showed that early activist companies lose more stock price declines than company followers, so the key is the timing of the announcement of its corporate activism action, according to these authors.

Moreover, Villagra, Monfort, and Méndez-Suárez (2021) analysed the impact of corporate activism on company value; these authors used Facebook and the "Stop Hate for Profit" campaign as a reference; on June 17, 2020, this campaign, launched by six organisations, accused Facebook of passivity in the face of violent or racist content on its platform, and invited advertisers to withdraw their ads until the company changed its stance. Many of them discontinued their advertising, and Facebook's shares fell. Villagra, Monfort, and Méndez-Suárez (2021) used a descriptive study in which they analysed the reaction of 33 companies to this particular event, and their results showed that there is a significant negative effect on the stock market for the company subject to the boycott; however, there is no effect on the sponsoring companies.

Another article that explores this question in more depth is Haq et al.'s (2022), who used the event study to analyse the reaction of investors to corporate activism of a socio-political nature, focusing specifically on racial discrimination; to do so, they considered 197 statements issued by companies after the assassination of George Floyd, an event that caused great social stir on the issue. Their results showed that investor reaction is more positive for companies that engage in more intense activism. In such a case, investors perceive that the public will reward to a greater extent companies that take a radical stance in advocating socio-political issues that involve broad social change (in this case, in relation to racial discrimination); therefore, sometimes issuing a statement can be beneficial to company performance, at least in the short term.

Although the influence of corporate activism on firm performance is theoretically appealing, its empirical application poses some challenges. In the academic literature, especially in the field of finance, this type of problem has been successfully addressed using the methodology proposed by Fama and French (1993). The model uses three factors to evaluate the return of certain economic assets, such as stocks: (a) market risk premium, which captures the additional return of an asset relative to another risk-free asset; (b) the size of the organisation; and (c) the value derived from stocks with a high book-to-market ratio, versus the value derived from stocks with a low book-to-market ratio. Among these three factors, the first one is the most useful for assessing the risk of corporate activism, while the other two allow us to isolate some effects that may affect the corporation's stock value. For example, Chan et al. (2001)

state that, especially in the more developed economies, corporations have very important intangible assets, such as their brand image, patents, or the know-how of their employees, and criticise the fact that they are rarely taken into account in their strategic evaluation. To show the relationship between these intangible assets and a corporation's performance, they use the Fama–French model using R&D investment. Similarly, Madden et al. (2006) apply the same model to assess the impact of brand equity, as defined by Interbrand, on corporate performance. Their results empirically show how strong or better-valued brands have a lower associated risk that provides them with competitive advantages such as lower volatility, better resistance to crises, or less uncertainty regarding their performance in the market. Therefore, the measurement of the risk associated with intangible assets is not new in the literature, but its application to corporate activism is a novel contribution. As mentioned above, due to corporate activism's own dynamics, this phenomenon can be seen as a source of risk for the corporations that put it into practice.

In short, and having analysed the previous research, the direction of the relationship between corporate activism actions and corporate finance results is not clear and previous authors have not taken into account the company risk associated with the use of corporate activism as a strategic tool. So, the following hypothesis is proposed:

> H1: Companies that engage in corporate activism have a lower investment risk than companies that do not.

4. Methodology

4.1. Data

After reviewing the literature on corporate activism and its possible effects on the financial value of companies that incorporate it into their strategies, is it possible that this type of initiative affects the risk of these companies? In our study, we have adopted a methodology followed in similar comparative studies (see Madden et al., 2006) based on the three-factor model of Fama and French (1993) which, in essence, tries to compare a portfolio of companies with some distinctive feature in the way they operate on the market, versus the portfolio of companies that make up the rest of the market.

To test the hypothesis, the entire portfolio (not a sample) of US companies listed on the New York Stock Exchange (NYSE) that had adopted activist initiatives from 2014 to June 2022 was selected and compared with the portfolio comprised of the rest of the NYSE companies. The FACTIVA database, newspaper libraries, and other historical news archives were used to locate the portfolio of activist companies. Stock data for each company was obtained from the Thompson Reuters Datastream database, and the final number of activist companies included in this portfolio was 96 (4% of the NYSE).

4.2. Modelling Approach

The Fama–French three-factor model (Fama & French, 1993) is an extension of the capital asset pricing model, which incorporates additional components to more accurately assess the risk associated with a stock (Black et al., 1972). Specifically, the Fama–French model postulates the relationship between the expected return $E(R_{it})$ and its associated risk, which it measures using three risk factors: (a) the market return R, (b) the difference between the return of large companies and the return of small companies (SMB = small minus big), and (c) the difference between the return of companies with a high book-to-market ratio and the return of companies with a low book-to-market ratio (HML). If R_{ft} is the risk-free interest rate, then:

$$E\left(R_{it}\right) = R_{ft} + \beta_{iM}\left[R_t - R_{ft}\right] + \beta_{iSMB}SMB_t + \beta_{iHML}HML_t$$

This specification of the Fama–French model estimates whether the difference between a safe asset is higher or lower than the expected return of other types of investments. To operationalise this estimate, the usual practice is to calculate this difference using the following expression (Madden et al., 2006):

$$R_{it} - R_{ft} = \alpha_{it} + \beta_{iM}\left[R_t - R_{ft}\right] + \beta_{iSMB}SMB_t + \beta_{iHML}HML_t + \varepsilon_{it}$$

If $\beta_{iM} > 1$, this means that the asset receiving the investment shows a higher risk than expected, and in the opposite direction if it is less than 1. The rest of the indicators (β_{iSMB} and β_{iHML}) provide additional measures of an asset's risk; values close to zero indicate its associated risk coincides with the expected risk of other benchmark assets in the market. To estimate the model more robustly, a system of two simultaneous equations was designed. The first used the portfolio of activist companies as the dependent variable, and the second used the general index of the reference market. This specification allows for an understanding of the possible correlation between the residuals of the different equations, and estimates more consistent parameters (Zellner & Huang, 1962).

5. Results

The estimation of the system of equations shows a high fit (r^2 = 0.95), and all parameters are significant and with the expected sign. Of particular interest is the value obtained for the parameter β_{iM}. The estimated coefficient corresponding to the activist companies is 0.851, which is significantly different from the estimated coefficient for the companies that make up the reference market (F = 142.35; p < 0.000). Table 1 shows the results of the estimation of the system of equations.

Table 1. Seemingly unrelated regression estimation results.

	Dependent variable	
	Activist corporations	Reference market
M	0.851***	0.998***
	(0.006)	(0.004)
SMB	0.105***	0.056***
	(0.011)	(0.008)
HML	0.249***	0.288***
	(0.007)	(0.005)
(Intercept)	0.000	0.000
	(0.000)	(0.000)
Observations	*n* = 1,756	*n* = 1,756
R^2	0.941	0.967
SSR	0.016	0.008
Correlation of residuals	0.439	
System R^2	0.953	

Notes: *** $p < 0.01$; standard errors in parenthesis.

Therefore, according to these results, companies that engage in activist initiatives present a lower risk than companies that make up the reference market. These results are in line with the findings from previous research by Haq et al. (2022) and Weinzimmer and Esken (2016).

6. Conclusions

6.1. Theoretical Conclusions

Corporate activism is a business strategy that has recently become widespread due to the demands of consumers today (Atanga et al., 2022). The theoretical arguments of this article are based on the concepts of CSA, CSR, and CPA studied from the point of view of communication and public relations. On one hand, research on corporate activism is recent but the literature review has shown that if corporate activism is used as a business strategy (aligning company values, activist messages or actions and stakeholders' values) it has effects on companies' results. On the other hand, social change is a fundamental aim of corporate activism (Vredenburg et al., 2020) so that in addition to its impact on stock market performance, corporate activism opens a way to examine the social impact of this strategy in relation to investors and other stakeholders (customers, employees, media, lobbies, government, etc.). Thus, does this corporate activism promote stakeholders from taking real action for social change?

This research has shown that companies that use corporate activism have lower market risk than companies that do not engage in this type of strategy. These results allow us to evaluate the possible advantages of using corporate activism as a business strategy and draw comparisons to those that do not use it. But, to what extent might corporate activism that targets one clear segment while alienating other segments be more effective than other business strategies? This work shows that corporate activism may not be associated with the investor risk predicted by some authors related to customer boycotts or employee strikes. Along this same line of argument, this research has demonstrated that organisations that have implemented corporate activism actions are associated with change in the stock market compared to companies not using an activist strategy.

6.2. Managerial Implications

One of the main managerial implications should be that managers should be aware of their customers' values in order to make the decision on how to engage in corporate activism. This would lead to sales growth because of the impact that corporate activism can have on companies in a lasting way as customers continue to remember companies long after the implementation of this type of initiative (Bhagwat et al., 2020). Likewise, companies should think about how they should communicate their activism actions and emphasise to their investors that it is a strategic activity with much lower risk than they think, in fact much lower than companies that do not engage in corporate activism as a business strategy. In addition, other implications that can be deduced from this research have to do with companies that do not currently engage in corporate activism. This research shows that these companies are missing the opportunity to align themselves with the needs and expectations of their customers and this is not the time to remain neutral on controversial issues. As for the investors of non-activist companies, a good internal public relations and communications strategy is necessary for them to understand the strategic importance of good use of corporate activism.

6.3. Limitations and Future Research

In line with the above, further research should be aimed at replicating the results of the analysis described here on other stock exchanges with companies from around the world in order to verify whether these results can be generalised, as only 96 US companies on the NYSE were analysed in this study. Furthermore, another limitation of this study is that variables such as cultural differences, political tradition, and investor ideology have not been considered; these variables could affect these results and future results if the research is expanded to other regions or countries. Complementarily, it would be also pertinent to analyse the authenticity or type of corporate activism initiatives to know whether those factors affect how investors respond to corporate activism in this study. Lastly, this study only analyses the short-term financial consequences of single corporate activism events, but it does not examine the potential long-term effects of a corporate activism strategy on investors or other stakeholders.

Acknowledgments

This work has been partially supported by the Government of Madrid (Comunidad de Madrid, Spain) under the multiannual agreement with Universidad Complutense de Madrid, in the line Excellence Programme for University Teaching Staff, in the context of the V PRICIT (Regional Programme for Research and Technological Innovation).

Conflict of Interests

The authors declare no conflict of interests.

References

Atanga, B. A., Xue, X., & Mattila, A. S. (2022). The impact of corporate sociopolitical activism (CSA) on brand attitude. *International Journal of Hospitality Management*, *107*, Article 103290. https://doi.org/10.1016/j.ijhm.2022.103290

Bedendo, M., & Siming, L. (2021). To advocate or not to advocate: Determinants and financial consequences of CEO activism. *British Journal of Management*, *32*(4), 1062–1081. https://doi.org/10.1111/1467-8551.12459

Bhagwat, Y., Warren, N. L., Beck, J. T., & Watson, G. F., IV. (2020). Corporate sociopolitical activism and firm value. *Journal of Marketing*, *84*(5), 1–21. https://doi.org/10.1177/0022242920937000

Black, F., Jensen, M. C., & Scholes, M. (1972). *The capital asset pricing model: Some empirical tests*. SSRN. https://ssrn.com/abstract=908569

Branicki, L., Brammer, S., Pullen, A., & Rhodes, C. (2021). The morality of "new" CEO activism. *Journal of Business Ethics*, *170*(2), 269–285. https://doi.org/10.1007/s10551-020-04656-5

Chan, L. K., Lakonishok, J., & Sougiannis, T. (2001). The stock market valuation of research and development expenditures. *The Journal of Finance*, *56*(6), 2431–2456.

Chatterji, A. K., & Toffel, M. W. (2016). *Do CEO activists make a difference? Evidence from a field experiment* (Harvard Business School Working Paper No. 16–100). https://dash.harvard.edu/bitstream/handle/1/26274858/16-100.pdf;jsessionid=2B6EA394CF72DA87D822ED78F04A04E5?sequence=1

Ciszek, E., & Logan, N. (2018). Challenging the dialogic promise: How Ben & Jerry's support for Black Lives Matter fosters dissensus on social media. *Journal of Public Relations Research*, *30*(3), 115–127. https://doi.org/10.1080/1062726x.2018.1498342

Corcoran, M., Newman, K., & Devasagayam, P. R. (2016). Consumer perception of corporate activism: Strategic implication for marketing. *International Journal of Academic Research in Business and Social Sciences*, *6*(10), 52–61. https://doi.org/10.6007/IJARBSS/v6-i10/2331

Den-Hond, F., Rehbein, K. A., De-Bakker, F. G. A., & Kooijmans-Van-Lankveld, H. (2014). Playing on two chessboards: Reputation effects between corporate social responsibility (CSR) and corporate political activity (CPA). *Journal of Management Studies*, *51*(5), 790–813. https://doi.org/10.1111/joms.12063

Dodd, M. (2018). Globalization, pluralization, and erosion: The impact of shifting societal expectations for advocacy and public good. *The Journal of Public Interest Communications*, *2*(2). https://doi.org/10.32473/jpic.v2.i2.p221

Dodd, M. D., & Supa, D. (2015). Testing the viability of corporate social advocacy as a predictor of purchase intention. *Communication Research Reports*, *32*(4), 287–293.

Dodd, M. D., & Supa, D. W. (2014). Conceptualizing and measuring corporate social advocacy: Examining the impact on corporate financial performance. *Public Relations Journal*, *8*(3), 2–23.

Duffy, C., & Korn, J. (2022, June 27). These US companies will cover travel costs for employees who need an abortion. *CNN Business*. https://edition.cnn.com/2022/06/24/tech/companies-abortion-reaction/index.html

Edelman. (2021). *21st annual Edelman trust barometer*. https://www.edelman.com/trust/2021-trust-barometer

Eilert, M., & Nappier Cherup, A. (2020). The activist company: Examining a company's pursuit of societal change through corporate activism using an institutional theoretical lens. *Journal of Public Policy & Marketing*, *39*(4), 461–476.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Ferguson, M. A. (2018). Building theory in public rela-

tions: Interorganizational relationships as a public relations paradigm. *Journal of Public Relations Research*, *30*(4), 164–178. https://doi.org/10.1080/1062726x.2018.1514810

Glambosky, M., & Peterburgsky, S. (2022). Corporate activism during the 2022 Russian invasion of Ukraine. *Economics Letters*, *217*, Article 110650. https://doi.org/10.1016/j.econlet.2022.110650

Hadani, M., Doh, J. P., & Schneider, M. (2019). Social movements and corporate political activity: Managerial responses to socially oriented shareholder activism. *Journal of Business Research*, *95*, 156–170.

Hambrick, D. C., & Wowak, A. (2021). CEO sociopolitical activism: A stakeholder alignment model. *Academy of Management Review*, *46*(1), 33–59. https://doi.org/10.5465/AMR.2018.0084

Haq, H., Shin, J., & Winnebeck, D. (2022). *Owning the issue pays off: Investor reaction to intensity of corporate socio-political activism*. SSRN. https://ssrn.com/abstract=4223037

Heath, R. L., & Palenchar, M. J. (2008). *Strategic issues management: Organizations and public policy challenges*. SAGE.

Hoppner, J. J., & Vadakkepatt, G. G. (2019). Examining moral authority in the marketplace: A conceptualization and framework. *Journal of Business Research*, *95*, 417–427. https://doi.org/10.1016/j.jbusres.2018.07.045

Korschun, D., Rafieian, H., Aggarwal, A., & Swain, S. D. (2019). *Taking a stand: Consumer responses when companies get (or don't get) political*. SSRN. http://dx.doi.org/10.2139/ssrn.2806476

Kotler, P., & Sarkar, C. (2018). *Brand activism: From purpose to action*. Idea Bite Press.

Krebel Chang, K. (2017). Starbucks for president: The disappearing line between government and business as agents for social change. *Revista Empresa y Humanismo*, *XX*(1), 35–64. https://doi.org/10.15581/015.xx.1.35-64

LaCombe, S. J., & Juelich, C. (2019). Salient ballot measures and the millennial vote. *Politics and Governance*, *7*(2), 198–212.

Madden, T. J., Fehle, F., & Fournier, S. (2006). Brands matter: An empirical demonstration of the creation of shareholder value through branding. *Journal of the Academy of Marketing Science*, *34*, 224–235. https://doi.org/10.1177/0092070305283356

Mukherjee, S., & Althuizen, N. (2020). Brand activism: Does courting controversy help or hurt a brand? *International Journal of Research in Marketing*, *37*(4), 772–788.

Overton, H., Kim, J. K., Zhang, N., & Huang, S. (2021). Examining consumer attitudes toward CSR and CSA messages. *Public Relations Review*, *47*, Article 102095. https://doi.org/10.1016/j.pubrev.2021.102095

Parcha, J. M., & Kingsley Westerman, C. Y. (2020). How corporate social advocacy affects attitude change toward controversial social issues. *Management Communication Quarterly*, *34*(3), 350–383. https://doi.org/10.1177/0893318920912196

Robinson, S. R., Irmak, C., & Jayachandran, S. (2012). Choice of cause in cause-related marketing. *Journal of Marketing*, *76*(4), 126–139.

Villagra, N., Clemente-Mediavilla, J., López-Aza, C., & Sánchez-Herrera, J. (2021). When polarization hits corporations: The moderating effect of political ideology on corporate activism. *Profesional de la Información*, *30*(6). https://doi.org/10.3145/epi.2021.nov.02

Villagra, N., Monfort, A., & Méndez-Suárez, M. (2021). Value impact of corporate activism: Facebook and the stop hate for profit campaign. *Journal of Business Research*, *137*, 319–326.

Villagra, N., Sánchez-Herrera, J., Clemente, J., & Pintado, T. (2022). Deconstructing corporate activism: A consumer approach. *Journal of Management & Organization*. Advance online publication. https://doi.org/10.1017/jmo.2022.70

Vredenburg, J., Kapitan, S., Spry, A., & Kemper, J. A. (2020). Brands taking a stand: Authentic brand activism or woke washing? *Journal of Public Policy & Marketing*, *39*(4), 444–460. https://doi.org/10.1177/0743915620947359

Weinzimmer, L. G., & Esken, C. A. (2016). Risky business: Taking a stand on social issues. *Business Horizons*, *59*(3), 331–337.

Wettstein, F., & Baur, D. (2016). "Why should we care about marriage equality?": Political advocacy as a part of corporate responsibility. *Journal of Business Ethics*, *138*(2), 199–213. https://doi.org/10.1007/s10551-015-2631-3

Zellner, A., & Huang, D. S. (1962). Further properties of efficient estimators for seemingly unrelated regression equations. *International Economic Review*, *3*, 300–313.

About the Authors

Teresa Pintado Blanco holds a PhD in communication and is professor of the Department of Marketing at Universidad Complutense de Madrid. She has participated in marketing and communication research projects, and she has several publications in academic journals and different books. She is member of the editorial board of several journals. Her professional experience has been developed in multinational advertising agencies, as well as marketing and business management departments. ORCID: https://orcid.org/0000-0003-2545-3770

Carlota López-Aza holds a PhD in business economics and is an associate professor in the Department of Marketing at Universidad Complutense de Madrid. She is in charge of the Master's in Business Administration and Management (MBA) at the UCM's Faculty of Economics and Business. Her scientific production is linked to corporate activism, communication, and corporate social responsibility. She has several publications in scientific journals, book chapters, and a book. Also, she has collaborated in several teaching innovation projects. ORCID: https://orcid.org/0000-0002-7723-066X

Joaquín Sánchez holds a PhD in communication and is a full professor at the Department of Marketing at Universidad Complutense de Madrid. He has participated in more than 20 communication and marketing books. He has published his research in prestigious international journals, such as: *Journal of Advertising Research*, *Journal of Marketing Channels*, and *Journal of Retailing and Consumer Services*. He is member of the editorial board of several international journals. ORCID: https://orcid.org/0000-0003-0737-1210

Pablo Gutiérrez-Rodríguez holds a PhD in business economics and has completed several research stays in different countries and universities such as Penn University (US), Stockholm University (Sweden), Bournemouth University (UK), and University of San Francisco (US). He is co-author of more than 40 books and book chapters. He has also participated in more than 100 articles in scientific journals on topics related to international marketing, public marketing, retailing, and communication. He is a member of the editorial board and guest editor of several journals. ORCID: https://orcid.org/0000-0001-5407-4265

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 147–159
https://doi.org/10.17645/pag.v11i2.6330

Article

Dilemmas Between Freedom of Speech and Hate Speech: Russophobia on Facebook and Instagram in the Spanish Media

David Caldevilla-Domínguez [1], Almudena Barrientos-Báez [1], and Graciela Padilla-Castillo [2,*]

[1] Department of Communication Theories and Communication Analysis, Complutense University of Madrid, Spain
[2] Department of Journalism and New Media, Complutense University of Madrid, Spain

* Corresponding author (gracielp@ucm.es)

Submitted: 25 October 2022 | Accepted: 7 December 2022 | Published: 17 May 2023

Abstract
On March 11, 2022, Russia opened a criminal case against Meta, the parent firm that owns Facebook, Instagram, and WhatsApp. It alleged that Facebook had modified its community standards, broadening its concept of freedom of speech to allow alleged hate speech against Russian citizens, amid the conflict in Ukraine. Reuters (2022, para. 1) refers to a "temporary change in the company's hate speech policy," according to confidential Facebook documents. The Office of the High Commissioner for Human Rights called the change "worrying" ("Rusia y Ucrania," 2022, para. 11). In this context, this article addresses two objectives: (a) to explore and comment on the state of the art on freedom of expression in social networks and its deontological limitations to prevent hatred against nationalities (EU legislation, scientific research, Twitter, and Instagram deontological limitations); and (b) to study the emergence of possible cases of Russophobia, in a mediatized form, through the news of Spanish media and the comments they generated on their Facebook and Instagram sites. A triangular methodology is used: analytical and longitudinal commentary on EU definitions and standards on hate speech; quantitative analysis of news items in Spanish media on Russophobia, on Facebook and Instagram, published between January 1, 2022, and October 20, 2022; and mixed analysis of the engagement of these news items, thanks to the Fanpage Karma tool. The media coverage of Russophobia is scarce, with an average of one news item per media and, exceptionally, with two news items in very few cases. It is also striking that in such a long period, only six hashtags are used.

Keywords
deontology; Facebook; freedom of speech; hate speech; Instagram; media; Meta; Russia; Russophobia; social media

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

On March 11, 2022, Russia opened a criminal case against Meta, the parent company that owns Facebook, Instagram, and WhatsApp. It alleged that Facebook had modified its community standards, broadening its concept of freedom of speech to allow alleged hate calls against the Russian citizenry, in the heat of the war against Ukraine. The Investigative Committee of the Russian Federation described Meta as an "extremist organization" ("Rusia abre una causa," 2022). And the Tverskoy court in Moscow, through a communiqué on Telegram, banned the activity of the US multinational in the federal territory ("La Justicia rusa," 2022).

After learning the news, Meta did not want to make statements about the accusation, nor about its inclusion in the Rosfinmonitoring or Federal Service for Financial Monitoring list as "organizations and individuals about which there is information about their involvement in extremist activities or terrorism" ("Russia's financial monitoring agency," 2022, para. 4). However, two weeks after the opening of the criminal case, a Meta

spokesperson unofficially stated that Facebook had only temporarily relaxed its rules on political speech, allowing some words and comments that previously would have been censored ("Rusia abre una causa," 2022).

The United Nations' Office of the High Commissioner for Human Rights (OHCHR) described the change as "worrying" ("Rusia y Ucrania," 2022, para. 11) even though "freedom of expression is a fundamental human right, enshrined in Article 19 of the Universal Declaration of Human Rights" (OHCHR, 2022, para. 1). The main reason for this concern is not an attack on Russian leaders, but attacks on those born in that country, even if they reside elsewhere in the world. It is Russophobia or irrational fear or hatred of Russia or Russian citizens. Russian identity is put in the bull's eye (Correyero-Ruiz & Sánchez Martínez, 2022; Soto Ivars, 2022) and this cancellation generalizes and harms the same Russians who protest against the war in Ukraine.

In Spain, this Russophobia materialized against some Russian companies, products, and citizens (Gómez Díaz, 2022). According to the Instituto Nacional de Estadística (2022), on January 1, 2021, there were officially 79,485 Russian citizens living in Spain (52,829 women and 26,656 men). There are no subsequent official data, as they are published only at the beginning of each new year. Catalonia, Valencia, and Andalusia are the autonomous communities with the most Russians per inhabitant, in decreasing order: 25,093, 22,264, and 14,172, respectively (Instituto Nacional de Estadística, 2022). The president of the Casa de Rusia in Alicante, Aleksandr Chepurnoy, denounced Russophobia in schools, in an intervention in the TV show Noticias 8 Mediterráneo (Osuna, 2022). Ximo Puig, president of the Generalitat Valenciana, defended any blockade of Putin without falling into Russophobia (García, 2022). In the cultural aspect, the Russian Museum of Malaga remains open but does not renew its exhibitions (RNE, 2022) and the Teatro Real de Madrid canceled the performances of the Bolshoi Ballet, scheduled for May 2022, due to "the war conflict unleashed by Russia in Ukraine, which is causing a serious crisis in the world and a painful humanitarian emergency" (Teatro Real, 2022, para. 1).

2. Methodology

In this context, this research work addresses two main objectives: (a) to explore and comment on the state of the art on freedom of expression in social networks and its deontological limitations to prevent hatred against nationalities (EU legislation, scientific research, Twitter, and Instagram deontological limitations); and (b) to study the emergence of possible cases of Russophobia, in a mediatized form, through the news of Spanish newspapers and the comments they generated on their Facebook and Instagram sites. This second objective is related to the first one, as we want to show a tangible and real case study, whose successes and errors serve as an exemplary study that can be extrapolated from Spain to other countries in Europe and the world. A triangular methodology is employed: analytical and longitudinal commentary on the definitions and community standards on hate speech; quantitative analysis of news in Spanish media, on Russophobia on Facebook and Instagram, published between January 1, 2022, and October 20, 2022; and mixed analysis of the engagement of such news, thanks to the Fanpage Karma tool.

For the first step (a longitudinal, analytical, and synthetic literature review), we studied the concepts of freedom of expression, hate speech, and deontological limitations in social networks. The term "longitudinal" is understood as the chronological bibliographic analysis, from the appearance of the concept or of the first research, in reverse chronological order, up to the present day. The term "analytical" comes from decomposition, which moves from the whole to the parts, seeking a detailed analysis of all the above-mentioned terms. Coming from the philosophical current of the same name, it emphasizes the importance of precision in language. The term "synthetic" refers to reasoning analysis that seeks the essential, reintegrating all the investigated parts, abstracting what is important (found in the analytical part), and updating it.

The literature review has been carried out through first-level bibliographic sources, and with articles of maximum impact and prestige, from publications indexed by the Journal Citation Reports and Scimago Journal & Country Rank. As a literature review, it seeks to describe the qualities of the context in which this research is framed to support the validity of the sample and period chosen. The aim is not to prove to what degree a certain quality is found in this study context, but to discover as many qualities and orientations as possible, based on the most recent and prestigious studies (Rodríguez-Vidales & Padilla-Castillo, 2018). Therefore, it pursues in-depth understanding, rather than the accuracy of purely quantitative research, as the context of social networks is rapidly changing in terms of data (Bernárdez-Rodal et al., 2021; Caldevilla-Domínguez et al., 2021a, 2021b; Díaz-Altozano et al., 2021; Padilla & Rodríguez, 2022). Precisely because of the longitudinal, analytical, and synthetic review, any literature review on social networks can quickly become outdated if it focuses on descriptive, quantitative, or situational aspects. For this reason, we opt for research that does not focus solely on quantitative aspects, but rather on the search for the profound reasons that give rise to these numbers or situations.

For the second step, which addresses the objective of studying the emergence of possible cases of Russophobia—in a mediatized form, through the news of Spanish media and the comments they generated on their Facebook and Instagram sites—the Fanpage Karma tool is employed. It was founded by Nicolas Graf von Kanitz and Stephan Eyl in 2012, in Berlin, Germany (Fanpage Karma, 2022), and is one of the most comprehensive platforms for social network analysis. It offers four solutions for social media work and analysis:

analytics, engagement, publishing, and discovery. This research has only employed discovery, which allows studying trends, topics, and influencers, on Twitter and Instagram, with unlimited analysis of profiles (Fanpage Karma, 2022). It offers top posts, the posts with the most interaction on the two mentioned social networks, top influencers with the most active profiles, and a precise search by region, language, category, keyword, and hashtag, in eligible and customizable periods, from the last hour to the last year.

The validity of Fanpage Karma rests on three pillars relevant to the objectives of this work: It collects only public information, meaning that there is no concern about confidential data or privacy violation (Gmiterek, 2021; Trunfio & Della Lucia, 2019); it has a quantitative nature by using data, although it offers common characteristics that allow to compare, establish origins, and consequences (Martínez-Sánchez et al., 2021); and only analyses business or professional social media profiles (Barrios-Rubio & Pedrero-Esteban, 2021; Manca et al., 2022), which ensures that the publications obtained come from official media accounts. Specifically, only accounts that Fanpage Karma categorizes as "media" or "news" were used. It also allows filtering by language (for this work, the Spanish language was chosen), by country (the search was only in Spain), and by date (for the sample to be valid, the period of the first 10 months of the year, between January 1, 2022, and October 20, 2022, was chosen to collect comments before and during the conflict).

3. Freedom of Expression in Social Networks and Its Deontological Limitations

Social media has grown exponentially in the number of users, during and after the international socio-health confinement (Abuín-Penas & Abuín-Penas, 2022; Demuner Flores, 2021), as well as in influence (Martínez-Fresneda Osorio & Sánchez Rodríguez, 2022). Studies conducted during 2021–2022 point to an unprecedented increase in the use of social networks and in the number of users who use them. Major reports from IAB Spain (2022), Statista (2022), The Social Media Family (2022), and We Are Social (2022) highlight Instagram and TikTok as the fastest-growing networks. According to the data, society would tend to be informed about a topic, for much longer, on social networks before traditional media (Calderón-Garrido et al., 2019; Gong et al., 2022; Hernandez-Cruz, 2021; Padilla-Castillo & Rodríguez-Hernández, 2022; Ryzhova, 2022; Segura-Mariño et al., 2020).

These social networks should not evade the legislation under which information is disseminated in traditional media. Because information, whatever its platform, has an ethical nature and its extreme complexity can carry perverse stereotypes, including those towards certain nationalities (S. Chen et al., 2019; Y. Chen et al., 2021; Matamoros-Fernández & Farkas, 2021; Rodríguez-Pérez et al., 2022; Tsuriel et al., 2021). Hate crime in Europe has been ethically contemplated for several decades (Padilla-Castillo, 2020). According to the original classical Greek definition, legislation will be understood as part of ethics, following the treatises of Socrates and Descartes. Deontology, on the other hand, is a part of ethics, which deals with the rules for the practice of a profession or with the rules laid down autonomously, when it includes the community rules of social networks. The International Covenant on Civil and Political Rights (United Nations, 1966), in its Article 2, dictates the importance of respecting and guaranteeing to all individuals the recognized rights, without distinction as to race, color, gender, language, religion, political or other opinions, national or social origin, property, birth, or another status. Its Article 18 refers to the freedom of thought, Article 19 refers to the freedom of opinion, and Article 26 protects non-discrimination (United Nations, 1966). And the internet has become a kind of fertile breeding ground for maliciously altered information (Almendros & Rojano, 2022; Quintana-Pujalte & Pannunzio, 2021).

In 2000, the Charter of Fundamental Rights of the European Union was ratified. In Articles 1, 10, 21, and 47 it was established that the EU guarantees the right to human dignity, freedom of conscience, religion, thought non-discrimination, and the protection of an impartial judge (Charter of Fundamental Rights of the European Union, 2000). The next milestone after the European Constitution was the 2003 constitutional treaty, which was signed in Rome by the heads of government of the member countries and ratified by the European Parliament in 2005. Part II of the European constitution deals with the fundamental rights of the EU and names: dignity, freedoms, equality, solidarity, citizenship, and justice (Treaty Establishing a Constitution for Europe, 2004).

Following EU legislation, it is essential to comment on Directive 2012/29/EU of the European Parliament and of the Council of October 25, 2012, establishing minimum standards on the rights, support, and protection of victims of crime, that replaced Council Framework Decision 2001/220/JHA. It specifies that hate crime encompasses any violence directed at another person, because of their "property, birth, disability, age, gender, gender expression, gender identity, sexual orientation, residence status, or health" (Directive 2012/29/EU of 25 October 2012, 2012, p. 2, para. 9). It is also important to mention the Office for Democratic Institutions and Human Rights of the Organization for Security and Co-Operation in Europe (OSCE) and the Helsinki Final Act (1975), as the outcome of the third phase of the Conference on Security and Co-Operation in Europe. Today, the OSCE brings together 57 states from Europe, North America, and Central Asia and "provides support, assistance and expertise to participating States and civil society to promote democracy, rule of law, human rights and tolerance and non-discrimination" (OSCE, n.d., para. 1).

In 2022, the European Commission opened a public consultation, claiming that "recent developments in

some EU countries affecting the ownership, management, or operation of certain media point to increasing interference in this sector" (European Commission, 2022, p. 2, para. 1). This consultation, open between January 10 and March 25, 2022, wanted to "ensure a better functioning of the EU media market by increasing legal certainty and removing obstacles to the internal market" (European Commission, 2022, para. 2). The Commission's proposal for regulation, published on September 16, 2022, is in the process of receiving comments as this article is being published. Until then, we know that one of its central thrusts is to protect freedom of expression, where the role of the media as actors in democratic systems based on the rule of law is developed (European Commission, 2022).

Reviewing the deontological or community standards of the two studied social networks, created and modified according to the use and abuse of users (Segado-Boj et al., 2020), we see that Twitter says in its safety chapter: "Violence: You may not threaten violence against an individual or a group of people. We also prohibit the glorification of violence"; "Terrorism/violent extremism: you may not threaten or encourage terrorism or violent extremism"; and "Hateful conduct: You may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories." (Twitter, 2022, paras. 2–6). However, numerous studies focus on criticizing the social network, which would allow hateful tweets (Eriksson, 2018; Jacobs et al., 2020; Wang & Ding, 2022), shielding behind the fact that "the purpose of Twitter is to be in the service of public conversation" (Twitter, 2022, para. 1).

For its part, Instagram offers two fundamental lines for this research: "There is no place on Instagram for people who support or praise terrorism, organised crime or hate groups" (Instagram, 2022, para. 11) and "We remove any content that includes credible threats or hate speech, content directed at individuals for the purpose of humiliating or embarrassing them, personal information used to blackmail or harass someone, and repeated unwanted messages" (Instagram, 2022 para. 14). However, like Twitter, it receives criticism for not censoring some posts in the interest of being a "a diverse and positive community" (Instagram, 2022 para. 14). Numerous articles have analyzed how Instagram has focused too much on censoring supposedly erotic photographs (for example, female nipples, breastfeeding, breast cancer surgery), rather than pursuing quasi-pornographic, violent, or covert advertising posts (Chung & Wihbey, 2022; Ging & Garvey, 2018; Kermani & Hooman, 2022; Savolainen et al., 2022).

4. Results

Following the methodology, the second objective of the research is to study the emergence of possible cases of Russophobia, in a mediatized form, through the news of Spanish media and the comments they generated on their Facebook and Instagram sites. In accordance with the triangular methodology, after analytically commenting on the definitions and community standards on hate speech, we proceed to the quantitative analysis of news about Russophobia in the Spanish media, which was published on their official Facebook and Instagram accounts, in the 10 months covered by the study. The quantitative analysis, under the spirit of the research, is completed with a mixed analysis of engagement, thanks to the Fanpage Karma tool (likes, comments, sharing, and reactions).

The first search focuses on "media," which are official Spanish and Spanish-language media accounts. Table 1 collects the name of the account, the number of "likes," the number of public comments, the number of times the post was shared, and the number of overall reactions (the sum of "like," "love," "haha," "sad," "angry," and "wow," in the case of Facebook; and "like" in the case of Instagram).

Table 1. Spanish media posts about Russophobia on Facebook and Instagram.

Order of priority	Account	Number of "likes"	Number of comments	Number of times shared	Number of total reactions
1	*El País*	2008	1000	186	3194
2	*20minutos.es*	2372	431	296	3099
3	*ABC.es*	920	413	33	1366
4	*La Vanguardia*	735	501	65	1301
5	*El País*	941	224	76	1241
6	Telecinco	446	362	24	832
7	*La Voz de Galicia*	320	480	16	816
8	*Público*	315	420	30	765
9	*20minutos.es*	468	246	15	729
10	*El Mundo*	353	267	34	654
11	*Ultima Hora Mallorca*	225	209	16	450
12	*Público*	124	259	20	403
13	Antena 3 Noticias	241	132	8	381

Table 1. (Cont.) Spanish media posts about Russophobia on Facebook and Instagram.

Order of priority	Account	Number of "likes"	Number of comments	Number of times shared	Number of total reactions
14	Cuatro	209	144	11	364
15	*ABC.es*	134	162	15	311
16	Antena 3 Noticias	129	155	7	291
17	*La Razón*	206	43	0	249
18	*elDiario.es*	119	72	23	214
19	*El Mundo*	78	119	9	206
20	RTVE	102	69	3	174
21	*El Independiente*	132	15	27	174
22	*Levante-EMV*	147	18	7	172
23	*elDiario.es*	100	50	8	158
24	*Público*	98	28	21	147
25	*infoLibre*	56	61	26	143
26	*Crónica Global*	38	74	5	117
27	*20minutos.es*	36	77	0	113
28	*El Confidencial*	50	50	13	113
29	*Público*	24	59	6	89
30	*elDiario.es*	57	16	9	82
31	*La Nueva España Digital*—LNE.es	59	15	2	76
32	NAIZ	53	11	9	73
33	*Gigantes del Basket*	30	22	1	53
34	laSexta Noticias	15	27	0	42
35	Espejo Público	6	21	1	28
36	*Diario de Avisos*	17	5	3	25
37	*El Español*	11	14	0	25
38	*Esquire Spain*	13	3	4	20
39	*La Opinión de Murcia*	9	6	2	17
40	*Larazon.es*	4	12	0	16
41	Más Vale Tarde	6	8	0	14
42	*El Periódico de Aragón*	3	5	2	10
43	*La Voz de Galicia*, Edición Lugo	2	4	0	6
44	*Diario Córdoba*	3	2	0	5
45	RTVE Noticias	4	1	0	5
46	*Informacion.es*	3	1	0	4
47	*El Periódico Extremadura*	4	0	0	4
48	*Larazon.es*	0	2	2	4
49	*El Día*	3	0	0	3
50	*Diario de Sevilla*	2	1	0	3
51	*Diario de Navarra*	1	2	0	3
52	*El Periódico Extremadura*	1	2	0	3
53	*eldiariocv.es*	2	0	0	2
54	*Leonoticias Diariodigital*	2	0	0	2
55	*El Periódico de Aragón*	2	0	0	2
56	*The Objective*	1	0	1	2
57	*The Objective*	1	1	0	2
58	*EC Cultura*	1	0	1	2
59	*Larazon.es*	2	0	0	2
60	*Crónica Global*	0	2	0	2
61	*eldiariocv.es*	1	0	0	1
62	*Leonoticias Diariodigital*	1	0	0	1
63	*Diario SUR*	0	0	1	1
64	*La Opinión de Zamora*	1	0	0	1
65	*Huelva Información*	0	0	0	0
66	*Economía Digital*	0	0	0	0
67	*El País Cat*	0	0	0	0

Table 1. (Cont.) Spanish media posts about Russophobia on Facebook and Instagram.

Order of priority	Account	Number of "likes"	Number of comments	Number of times shared	Number of total reactions
68	*Crónica Global*	0	0	0	0
69	La 8 Mediterráneo	0	0	0	0
70	*El País Cat*	0	0	0	0
71	Madridiario.es	0	0	0	0
72	*El País*	2008	1000	186	3194
73	*20minutos.es*	2372	431	296	3099

Note: Own elaboration based on data from Fanpage Karma.

A total of 73 publications were obtained for the period studied. For space reasons, it is not possible to transcribe all of these publications, but they are placed in the Supplementary Files, with the complete post of each publication. These media publish only on Facebook, except for post number 17, from the newspaper *La Razón*, which was made on Instagram. There are national newspapers that are disseminated in print and digital (*ABC*, *El Mundo*, *El País*, *El Periódico*, *La Razón*, *La Vanguardia*), those that are published only in digital (*El Confidencial*, *El Español*, *infoLibre*, *Público*, *The Objective*), and those that have regional circulation (*El Periódico de Aragón*, *Diario Córdoba*, *Diario de Navarra*, *Diario de Sevilla*, *Diario SUR*, *El Día*, *El Periódico de Aragón*, *El Periódico Extremadura*, *Huelva Información*, *La Opinión de Murcia*, *La Opinión de Zamora*, *Levante*, *La Voz de Galicia*, *Madridiario.es*, etc.). There are also television channels (Antena 3, Cuatro, La Sexta, RTVE), which publish posts from their news programmes and, in the case of Antena 3's Espejo Públic, also from their TV shows. In the free press, *20minutos.es* stands out, with several publications on social media, and as for the magazines, only *Esquire Spain* and the *EC Cultura* supplement of the digital newspaper *El Confidencial* have a presence in social media.

Regarding the content of the posts, we found seven quotes in the headlines from interviewees referring to that Russophobia: two quotes from Putin; two opinion letters or editorials; and three allusions to expressions, in quotation marks ("social justice," "political formula," and "evil Russian"), referring to television fiction, specifically to the stereotypes in the Netflix series *Stranger Things*. Furthermore, there were only five hashtags (from most to least engagement): #TrincheraCultural, #Hearstforpeace, #Esquireforpeace, #pensamiento, and #opinión.

Only half of the posts have significant engagement, as interactions decline to insignificance from post 40 onwards. Regarding the nature of these interactions, the most used is the "like," doubling and tripling the number of interactions per comment, in all cases. The number of times posts are shared is also strikingly low, and only in the first two posts is it more than a hundred. The most important post is from *El País*:

> It can't be that people here hate us just because we are Russians. The big problem is that they identify Russians with Putin and we are not to blame for his decisions. Least of all the children. The Russian community in Catalonia denounces insults against them and even death threats.

It refers to an article with interviews with several Russian people living in Catalonia, which, as mentioned, is home to the largest Russian community in Spain.

The second search focuses only on "news" and its results are collected in Table 2. Some posts coincide with those in Table 1 ("media"), but new ones are now incorporated as Fanpage Karma cataloged them exclusively as "news" and not as entertainment products. It is important to point out that we differentiate and use both categories—"news" and "media"—because "media" includes accounts dedicated to news, entertainment, analysis, and fiction, among others, while "news" refers only to accounts that disseminate information. Similarly to Table 1, the name of the account, the number of "likes," the number of public comments, the number of times the post was shared, and the number of

Table 2. News publications in Spanish and Spain, on Facebook and Instagram.

Order of priority	Account	Number of "likes"	Number of comments	Number of times shared	Number of total reactions
1	El País	2008	1000	186	3194
2	*20minutos.es*	2372	431	296	3099
3	*ABC.es*	920	413	33	1366
4	*La Vanguardia*	735	501	65	1301
5	El País	941	224	76	1241
6	*La Voz de Galicia*	320	480	16	816

Table 2. (Cont.) News publications in Spanish and Spain, on Facebook and Instagram.

Order of priority	Account	Number of "likes"	Number of comments	Number of times shared	Number of total reactions
7	*Público*	315	420	30	765
8	*20minutos.es*	468	246	15	729
9	*El Mundo*	353	267	34	654
10	*Ultima Hora Mallorca*	225	209	16	450
11	*Público*	124	259	20	403
12	Antena 3 Noticias	241	132	8	381
13	*ABC.es*	134	162	15	311
14	Antena 3 Noticias	129	155	7	291
15	*La Razón*	206	43	0	249
16	*elDiario.es*	119	72	23	214
17	*El Mundo*	78	119	9	206
18	RTVE	102	69	3	174
19	*El Independiente*	132	15	27	174
20	*Levante-EMV*	147	18	7	172
21	*elDiario.es*	100	50	8	158
22	*Público*	98	28	21	147
23	*infoLibre*	56	61	26	143
24	*Crónica Global*	38	74	5	117
25	*20minutos.es*	36	77	0	113
26	*El Confidencial*	50	50	13	113
27	*Público*	24	59	6	89
28	*elDiario.es*	57	16	9	82
29	*La Nueva España Digital*—LNE.es	59	15	2	76
30	NAIZ	53	11	9	73
31	*Gigantes del Basket*	30	22	1	53
32	*Diario de Avisos*	17	5	3	25
33	*El Español*	11	14	0	25
34	Esquire Spain	13	3	4	20
35	*La Opinión de Murcia*	9	6	2	17
36	*Larazon.es*	4	12	0	16
37	*El Periódico de Aragón*	3	5	2	10
38	*La Voz de Galicia*. Edición Lugo	2	4	0	6
39	*Diario Córdoba*	3	2	0	5
40	RTVE Noticias	4	1	0	5
41	*Informacion.es*	3	1	0	4
42	*El Periódico Extremadura*	4	0	0	4
43	*Larazon.es*	0	2	2	4
44	*El Día*	3	0	0	3
45	*Diario de Sevilla*	2	1	0	3
46	*Diario de Navarra*	1	2	0	3
47	*El Periódico Extremadura*	1	2	0	3
48	*eldiariocv.es*	2	0	0	2
49	*Leonoticias Diariodigital*	2	0	0	2
50	*El Periódico de Aragón*	2	0	0	2
51	*The Objective*	1	0	1	2
52	*The Objective*	1	1	0	2
53	*EC Cultura*	1	0	1	2
54	*Larazon.es*	2	0	0	2
55	*Crónica Global*	0	2	0	2
56	*eldiariocv.es*	1	0	0	1
57	*Leonoticias Diariodigital*	1	0	0	1
58	*Diario SUR*	0	0	1	1
59	*La Opinión de Zamora*	1	0	0	1
60	*Huelva Información*	0	0	0	0
61	*Economía Digital*	0	0	0	0

Table 2. (Cont.) News publications in Spanish and Spain, on Facebook and Instagram.

Order of priority	Account	Number of "likes"	Number of comments	Number of times shared	Number of total reactions
62	El País Cat	0	0	0	0
63	*Crónica Global*	0	0	0	0
64	El País Cat	0	0	0	0
65	Madridiario.es	0	0	0	0
66	El País	2008	1000	186	3194
67	*20minutos.es*	2372	431	296	3099

Note: Own elaboration based on data from Fanpage Karma.

overall reactions (the sum of "like," "love," "haha," "sad," "angry," and "wow," in the case of Facebook; and "like" in the case of Instagram) are collected.

A total of 67 publications were obtained for the period studied. For space reasons, it is not possible to transcribe all of these publications, but they are placed in the Supplementary Files, with the complete post of each publication. As in Table 1, the media publish only on Facebook, except for post number 15 (post number 17 in Table 1), from the newspaper *La Razón*, which was made on Instagram (an opinion article titled "The nonsense of Russophobia"). Just as in Table 1, there are national newspapers that are distributed in print and digital, those that are published only in digital, and those that have regional circulation. As for the television channels, this time only posts from their news programs and not from other entertainment shows were considered.

About the content of the posts, we find four quotes in the headlines from interviewees referring to that Russophobia: two quotes from Putin and two allusions to expressions in quotation marks (again, "social justice" and "evil Russian"). There are also six hashtags, although not entirely the same as in the first table (from most to least engagement): #TrincheraCultural, #Hearstforpeace, #Esquireforpeace, #pensamiento, #opinión, and #madridiario.

Only about half of the posts have significant engagement, as interactions decline to insignificance from post 39 onwards. Regarding the nature of these interactions, the most used is once again "like," always multiplying the number of interactions per comment. The number of times posts are shared is also low, although the first five posts exceed one thousand reactions. When compared to Table 1, the news stories generate more engagement than entertainment products. The most important post is the same that occupied the first position in Table 1, from *El País*. The second and third most important posts are also the same. The one from *20minutos.es* is a Facebook Watch post, with a video (1 minute and 10 seconds long) of Putin presented with the text "Putin's tough message to the West about sanctions that harm Russia." *ABC*'s is a Facebook post with the text "The French brand, after the closure of its stores in Russia, has blocked the marketing of its products to Russian customers who intend to take them to their country," two photographs of Russian women smashing a Chanel bag, and the link to the article of the same newspaper, with the headline "Russian influencers smash their Chanel bags of 8,000 euros after the ban on sale: 'This is Russophobia.' "

5. Conclusions and Discussion

The first objective of this research was to explore and comment on the state of the art on freedom of expression in social networks and its deontological limitations to prevent hatred against nationalities (EU legislation, scientific research, and Twitter and Instagram deontological limitations). The analysis was divided between a detailed study of European legislation, as the geographical and political scope of the corpus of the work, and the most recent and cutting-edge research on the ethical problems of Twitter and Instagram, the two social networks owned by Meta. Regarding European regulations, the wide tradition of directives, recommendations, and institutions, with more than 80 years of experience, before and after the European Constitution, has been identified and commented on. Considering the dynamic and changing nature of communication and new technologies, the range is closed in a consultation of European citizens, currently awaiting their comments and in the spirit of continuing to ensure a pluralistic, transparent, and democratic media spectrum in the countries of the European Union (European Commission, 2022). It is clear that, although the legislation is numerous and has a long tradition in time, it needs more speed to adjust quickly to changes and social problems, and that more institutional and media publicity is also needed, so that the public, of all strata, is aware of it and does not comply with the hatred that continues to exist in social networks.

Regarding Twitter and Instagram, the main lines of their community rules have been commented on, citing freedom of expression and hate speech, although, in parallel, the two networks openly state their desire to promote dialogue and freedom (Instagram, 2022; Twitter, 2022). At the same time, the academic tradition of the last five years focuses on certain forms of unnecessary censorship, especially on Instagram, which would unfairly allow certain inappropriate posts and censor publications. Progress on this censorship comes from social denunciation and originates advances, albeit slow, in social networks, which would still need more monitoring and surveillance by Meta, and more digital

literacy on the part of users so that they detect hate and can proactively denounce and expedite these changes (Bernárdez-Rodal et al., 2021; Díaz-Altozano et al., 2021; Kermani & Hooman, 2022). This theoretical analysis also shows a broad, rigorous, and adequate academic tradition to the problems of each moment: Academic texts do cover many problems and details, even before ethics or legislation. But academic texts also recommend, in many of their conclusions, a greater monitoring by social networks, that, under the guise of promoting dialogue, sometimes allow hateful publications that no one denounces and that should not exist.

The second objective of the work was to study the emergence of possible cases of Russophobia, in a mediatized form, through the news of Spanish media and the comments they generated on their Facebook and Instagram sites, thanks to the digital analysis platform Fanpage Karma. The results have been offered in two tables (Tables 1 and 2), one on "media" and the other on "news," in Spain and Spanish, published between January 1, 2022, and October 20, 2022. The resulting posts have been very similar and it is striking that for such a long period, the number of posts on Russophobia did not reach a hundred units. Accordingly, and considering the number of media analyzed, the media coverage of Russophobia is scarce, with an average of 1 news item per media, and with 2 news items in very few cases. One might think that this coverage does not help to raise awareness of the problem, thinking quantitatively, in terms of the number of news items. However, given the public's trust in the regional media, this idea would require another full in-depth study just to address this issue. The reason for the potential value of such a study is the ample existence of news and testimonies in Spanish regional media. According to Instituto Nacional de Estadística's (2022) official data, the Russian population in Spain is mostly divided between Catalonia, Valencia, and Andalusia, and numerous media from these communities do appear in the two tables, showing a greater commitment than the national media.

Analyzing the posts in-depth, it is also striking that in such a long period, only six hashtags are used. This circumstance leads us to conclude that the contents of the posts are very similar, in their language, to the news or article they refer to. Only in exceptional cases does the post innovate compared to the informative content that it intends to share and that has been previously published in a newspaper, television channel, or magazine. This being so, and thinking about Facebook's audience, it can be guessed that the media targets an audience of adult generations (millennials, generation X, and boomers). The centennial and alpha generations, as digital natives, consume and write publications with more hashtags and emojis, and these barely have a place in the Facebook posts detected and analyzed. Thus, these posts are designed for young adults and adults, and the language is adapted to their uses and customs. This fact is backed by the unique appearance of a single post on Russophobia on Instagram (from the newspaper *La Razón*), while all the others are from Facebook.

As for limitations and prospects in this study, there are numerous possibilities that this work cannot cover because it is framed within specific spatial and temporal guidelines. The first limitation comes from the Fanpage Karma tool, which only allows the analysis of the last year at the time of access. It would therefore be interesting to compare this analysis with another study from years before 2022, to detect whether the problem of Russophobia already existed and, if it did, it received media coverage at the time. Another limitation is that Fanpage Karma only offers its free Discovery tool for analyzing Twitter, Facebook, and Instagram. Only with a professional account, with a high paid fee, can one access the analysis of more social networks, including TikTok. However, this analysis of more social media is only offered on the profiles of the account holder or owner of that social media, while Discovery allows for the analysis of accounts that are not owned by the searcher, i.e., public accounts, as also discussed in the methodological framework. This limitation could lead to other prospective research, comparing the results of this work with other research that would also analyze YouTube and TikTok. However, as these social networks more typically target alpha and centennial audiences, it can be argued that their sample would still be much smaller.

Another limitation refers to the reality of Russophobia in Spain. This article has dealt with how Russophobia appears in the Spanish media, but it would be very interesting to know whether this Russophobia does exist and what crime rates are involved.

Acknowledgments

This article is part of the framework of a Concilium project (931.791) of the Complutense University of Madrid: Validation of Communication Models, Business, Social Networks and Gender. This research is part of a project supported by the Madrid Government (Comunidad de Madrid, Spain) under the Multiannual Agreement with Universidad Complutense de Madrid in the line Research Incentive for Young PhDs, in the context of the V PRICIT (Regional Programme of Research and Technological Innovation; Call PR/27/21, titled "Traceability, Transparency, and Access to Information: Study and Analysis of the Dynamics and Trends in the Area," reference PR27/21–017, duration 2022–2024, funding of 43,744.22 euros).

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 authors (unedited).

References

Abuín-Penas, J., & Abuín-Penas, R. (2022). Social media and interest in official information in times of pandemic: Analysis of the communication of the European ministries of health on Facebook during Covid-19. *Revista de Comunicación y Salud*, *12*, 59–76. https://doi.org/10.35669/rcys.2022.12.e303

Almendros, P. V., & Rojano, F. J. P. (2022). The work of verification of news from the communication department. A case study in the agri-food sector. *Revista de Ciencias de la Comunicación e Información*, *27*, 1–23. https://doi.org/10.35742/rcci.2022.27.e238

Barrios-Rubio, A., & Pedrero-Esteban, L. M. (2021). The transformation of the Colombian media industry in the smartphone era. *Journal of Creative Communications*, *16*(1), 45–60. https://doi.org/10.1177/0973258620981347

Bernárdez-Rodal, A., López-Priego, N., & Padilla-Castillo, G. (2021). Culture and social mobilisation against sexual violence via Twitter: The case of the "#LaManada" court ruling in Spain. *Revista Latina de Comunicación Social*, *79*, 237–262. www.doi.org/10.4185/RLCS-2021-1502

Calderón-Garrido, D., León-Gómez, A., & Gil-Fernández, R. (2019). The use of the social networks between the students of teacher's degree in an environment exclusively online. *Revista de Comunicación*, *147*, 23–40. http://doi.org/10.15178/va.2019.147.23-40

Caldevilla-Domínguez, D., Barrientos-Báez, A., & Padilla-Castillo, G. (2021a). Twitter as a tool for citizen education and sustainable cities after Covid-19. *Sustainability*, *13*(6), Article e3514. https://doi.org/10.3390/su13063514

Caldevilla-Domínguez, D., Barrientos-Báez, A., Pérez-García, A., & Gallego-Jiménez, M. G. (2021b). The use of social media and its relation to the tourist's purchase decision. *Revista de Comunicación*, *154*, 443–458. http://doi.org/10.15178/va.2021.154.e1360

Charter of Fundamental Rights of the European Union. (2000). *Official Journal of the European Union*, C 326.

Chen, S., Schreurs, L., Pabian, S., & Vandenbosch, L. (2019). Daredevils on social media: A comprehensive approach toward risky selfie behavior among adolescents. *New Media & Society*, *21*(11/12), 2443–2462. https://doi.org/10.1177/1461444819850112

Chen, Y., Sherren, K., Smit, M., & Lee, K. Y. (2021). Using social media images as data in social science research. *New Media & Society*. Advance online publication. https://doi.org/10.1177/14614448211038761

Chung, M., & Wihbey, J. (2022). Social media regulation, third-person effect, and public views: A comparative study of the United States, the United Kingdom, South Korea, and Mexico. *New Media & Society*. Advance online publication. https://doi.org/10.1177/14614448221122996

Correyero-Ruiz, B., & Sánchez Martínez, J. (2022). Access to "truth" through the "fiction" in Chernobyl. The delirium of the story in hypermodernity. *Historia y Comunicación Social*, *27*(1), 253–265. https://doi.org/10.5209/hics.72862

Demuner Flores, M. R. (2021). The use of social media in micro-enterprises in the face of Covid-19 effects. *Revista de Comunicación de la SEECI*, *54*, 97–118. https://doi.org/10.15198/seeci.2021.54.e660

Díaz-Altozano, P., Padilla-Castillo, G., & Requeijo-Rey, P. (2021). Girls' sexualization on social networks: The need for semantic intelligence on Instagram. *Investigaciones Feministas*, *12*(1), 31–45. https://doi.org/10.5209/infe.69559

Directive 2012/29/EU of the European Parliament and of the Council of 25 October 2012 establishing minimum standards on the rights, support and protection of victims of crime, and replacing Council framework decision 2001/220/Jha. (2012). *Official Journal of the European Union*, L 315.

Eriksson, M. (2018). Pizza, beer and kittens: Negotiating cultural trauma discourses on Twitter in the wake of the 2017 Stockholm attack. *New Media & Society*, *20*(11), 3980–3996. https://doi.org/10.1177/1461444818765484

European Commission. (2022). *Safeguarding media freedom in the EU—New rules*. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13206-Safeguarding-media-freedom-in-the-EU-new-rules_en

Fanpage Karma. (2022). *Features*. www.fanpagekarma.com/es/features

García, S. (2022, March 16). Ximo Puig apoya a la UV, pero advierte de no caer en la "rusofobia" [Ximo Puig supports the UV, but warns against falling into "Russophobia"]. *ESdiario Comunidad Valenciana*. www.esdiario.com/valencia/398514049/ximo-puig-rusofobia-universidad-valencia.html

Ging, D., & Garvey, S. (2018). "Written in these scars are the stories I can't explain": A content analysis of pro-ana and thinspiration image sharing on Instagram. *New Media & Society*, *20*(3), 1181–1200. https://doi.org/10.1177/1461444816687288

Gmiterek, G. (2021). Polish university libraries social networking services during the Covid-19 pandemic spring term lockdown. *The Journal of Academic Librarianship*, *47*(3), 1–7. https://doi.org/10.1016/j.acalib.2021.102331

Gómez Díaz, L. (2022, April 3). Las caras de la rusofobia en España: De la discriminación por "justicia social" al sentimiento de culpa por la guerra [The faces of Russophobia in Spain: From discrimination for "social justice" to guilt over the war]. *RTVE.es*. www.rtve.es/noticias/20220403/caras-rusofobia-espana-discriminacion-justicia-social-sentimiento-culpa-guerra/2326641.shtml

Gong, Q., Verboord, M., & Wang, Y. (2022). Media usage and political trust among young adults in China:

The role of media credibility, trust in sources and political membership. *Global Media and Communication*, *18*(3), 301–321. https://doi.org/10.1177/17427665221125553

Hernandez-Cruz, N. (2021). Mapping the thematic evolution in communication over the first two decades from the 21st century: A longitudinal approach. *Iberoamerican Journal of Science Measurement and Communication*, *1*(3), 1–10. https://doi.org/10.47909/ijsmc.88

IAB Spain. (2022). *Top tendencias digitales 2022* [Top digital trends 2022]. https://iabspain.es/estudio/top-tendencias-digitales-2022

Instagram. (2022). *Terms and policies*. https://help.instagram.com

Instituto Nacional de Estadística. (2022). *Población extranjera por nacionalidad, comunidades, sexo y año* [Foreign Population by nationality, communities, sex and year] [Data set]. www.ine.es/jaxi/Tabla.htm?path=/t20/e245/p08/l0/&file=02005.px&L=0

Jacobs, K., Sandberg, L., & Spierings, N. (2020). Twitter and Facebook: Populists' double-barreled gun? *New Media & Society*, *22*(4), 611–633. https://doi.org/10.1177/1461444819893991

Kermani, H., & Hooman, N. (2022). Hashtag feminism in a blocked context: The mechanisms of unfolding and disrupting #rape on Persian Twitter. *New Media & Society*. Advance online publication. https://doi.org/10.1177/14614448221128827

La Justicia rusa declara "extremista" Meta y prohíbe Facebook e Instagram en el país [Russian judiciary declares Meta "extremist" and bans Facebook and Instagram in Russia]. (2022, March 21). *El Confidencial*. www.elconfidencial.com/mundo/2022-03-21/justicia-rusa-prohibe-facebook-instagram-por-extremismo_3395304

Manca, S., Passarelli, M., & Rehm, M. (2022). Exploring tensions in Holocaust museums' modes of commemoration and interaction on social media. *Technology in Society*, *68*, 1–13. https://doi.org/10.1016/j.techsoc.2022.101889

Martínez-Fresneda Osorio, H., & Sánchez Rodríguez, G. (2022). The influence of Twitter in the agenda setting of communication media. *Revista de Ciencias de la Comunicación e Información*, *27*, 1–21. https://doi.org/10.35742/rcci.2022.27.e136

Martínez-Sánchez, M. E., Nicolas-Sans, R., & Bustos-Díaz, J. (2021). Analysis of the social media strategy of audio-visual OTTs in Spain: The case study of Netflix, HBO and Amazon Prime during the implementation of Disney +. *Technological Forecasting and Social Change*, *173*, 1–7. https://doi.org/10.1016/j.techfore.2021.121178

Matamoros-Fernández, A., & Farkas, J. (2021). Racism, hate speech, and social media: A systematic review and critique. *Television & New Media*, *22*(2), 205–224. https://doi.org/10.1177/1527476420982230

Organization for Security and Co-Operation in Europe. (n.d.). *What we do*. https://www.osce.org/odihr/what-we-do

Osuna, P. (2022, March 3). El presidente de la Casa de Rusia en Alicante denuncia casos de rusofobia en los colegios [The president of the Casa de Rusia in Alicante denounces cases of Russophobia in schools]. *Viva Valencia Noticias*. https://vivavalencia.net/valencia/1033547/el-presidente-de-la-casa-de-rusia-en-alicante-denuncia-casos-de-rusofobia-en-los-colegios

Padilla-Castillo, G. (2020). Violencia de género como delito de odio en redes sociales [Gender-based violence as a hate crime on social networks]. In A. Bernárdez (Ed.), *Desafíos Feministas. Temas y tramas para pensar en un mundo en crisis* (pp. 109–140). Tirant Lo Blanch.

Padilla-Castillo, G., & Rodríguez-Hernández, J. (2022). Sustainability in TikTok after Covid-19. The viral influencers in Spanish and their micro-actions. *Estudios sobre el Mensaje Periodístico*, *28*(3), 573–585. https://doi.org/10.5209/esmp. 81133

Quintana-Pujalte, L., & Pannunzio, M. F. (2021). Fact-checking in Latin America. Typology of viral contents dismissed during the coronavirus pandemic. *Revista de Ciencias de la Comunicación e Información*, *26*, 27–46. https://doi.org/10.35742/rcci.2021.26.e178

RNE. (2022, March 10). El Museo Ruso de Málaga: Abierto, sin renovar exposiciones [The Russian Museum in Malaga: Open, without renewing exhibitions]. *24 Horas*. www.rtve.es/play/audios/24-horas/museo-ruso-malaga-abierto/6430868

Rodríguez-Pérez, C., Jivkova-Semova, D., Pérez-Vara, E., Asadi, N., & Kreutler, M. (2022). Unravelling migration media coverage: How migrants and immigration were portrayed in Spain during the refugee crisis (2015–2018). *Historia y Comunicación Social*, *27*(1), 161–173. https://doi.org/10.5209/hics.77537

Rodríguez-Vidales, Y., & Padilla-Castillo, G. (2018). The pure political cinema: The fiction as inspiration–reflection of political spectacularization. *Revista de Comunicación de la SEECI*, *45*, 15–42. https://doi.org/10.15198/seeci.2018.45.15-42

Rusia abre una causa penal contra Meta por llamados a la violencia en Facebook [Russia opens criminal case against Meta over Facebook calls for violence]. (2022, March 11). *Euronews*. https://es.euronews.com/2022/03/11/ucrania-crisis-meta-rusia

Rusia y Ucrania: Rusia bloqueará Instagram después de que Meta anunciara que permitirá los llamados a la violencia contra Putin y las tropas rusas [Russia and Ukraine: Russia to block Instagram after Meta announces it will allow calls for violence against Putin and Russian troops]. (2022, March 11). *BBC News Mundo*. https://www.bbc.com/mundo/noticias-60707988

Russia's financial monitoring agency adds Meta to "extremists" list–agencies. (2022, October 11). *Reuters*. www.reuters.com/technology/russias-

financial-monitoring-agency-adds-meta-extremists-list-ifax-2022-10-11
Ryzhova, A. (2022). Motivated by political beliefs, not only by language: How Russian speakers in Germany compose their transnational news repertoires. *Journalism*. Advance online publication. https://doi.org/10.1177/14648849221130557
Savolainen, L., Uitermark, J., & Boy, J. D. (2022). Filtering feminisms: Emergent feminist visibilities on Instagram. *New Media & Society*, *24*(3), 557–579. https://doi.org/10.1177/1461444820960074
Segado-Boj, F., Díaz-Campo, J., & Navarro-Sierra, N. (2020). Emotions and news on social media about climate change sharing. Moderating role of habits, previous attitudes and uses and gratifications among university students. *Revista Latina de Comunicación Social*, *75*, 245–269. http://doi.org/10.4185/RLCS-2020-1425
Segura-Mariño, A. G., Paniagua-Rojano, F. J., & Fernández-Sande, M. (2020). Methodology to evaluate university communication on Facebook and Twitter. *Revista Prisma Social*, *28*, 127–144. https://revistaprismasocial.es/article/view/3357
Soto Ivars, J. (2022, March 6). Rusofobia: Cortar lazos culturales con Rusia es fortalecer a Putin [Russophobia: Cutting cultural ties with Russia strengthens Putin]. *El Confidencial*. www.elconfidencial.com/cultura/2022-03-06/rusofobia-cortar-lazos-culturales-fortalecer-putin_3386523
Statista. (2022). *Redes sociales con mayor número de usuarios activos a nivel mundial en enero de 2022 (en millones)* [Social networks with the largest number of active users worldwide in January 2022 (in millions)] [Data set]. https://es.statista.com/estadisticas/600712/ranking-mundial-de-redes-sociales-por-numero-de-usuarios
Teatro Real. (2022, March 4). *El Teatro Real cancela la actuación del Ballet Bolshói* [Teatro Real cancels Bolshoi Ballet performance] [Press Release]. www.teatroreal.es/en/node/18065
The Social Media Family. (2022). *VII informe del uso de las redes sociales en España 2021* [VII report on the use of social networks in Spain 2021]. https://thesocialmediafamily.com/informe-redes-sociales
Treaty Establishing a Constitution for Europe. (2004). *Official Journal of the European Union*, C 310.
Trunfio, M., & Della Lucia, M. (2019). Engaging destination stakeholders in the digital era: The best practice of Italian regional DMOs. *Journal of Hospitality & Tourism Research*, *43*(3), 349–373. https://doi.org/10.1177/1096348018807293
Tsuriel, K., Dvir Gvirsman, S., Ziv, L., Afriat-Aviv, H., & Ivan, L. (2021). Servant of two masters: How social media editors balance between mass media logic and social media logic. *Journalism*, *22*(8), 1983–2000. https://doi.org/10.1177/1464884919849417
Twitter. (2022). *The Twitter rules*. https://help.twitter.com/en/rules-and-policies/twitter-rules
United Nations. (1966). *International Covenant on Civil and Political Rights*. www.ohchr.org/SP/ProfessionalInterest/Pages/CCPR.aspx
United Nations' Office of the High Commissioner for Human Rights. (2022). *Freedom of expression and opinion*. www.ohchr.org/en/topic/freedom-expression-and-opinion
Wang, S., & Ding, R. (2022). "Business inquiries are welcome": Sex influencers and the platformization of non-normative media on Twitter. *Television & New Media*, *23*(8), 822–839. https://doi.org/10.1177/15274764221077666
We Are Social. (2022). *Digital Report 2022*. https://wearesocial.com/es/blog/2022/01/digital-2022

About the Authors

David Caldevilla-Domínguez is a doctor in information sciences (audiovisual communication) from the Complutense University of Madrid and has a diploma in teaching from the University of Zaragoza. Is a professor at the faculty of information sciences of the Complutense University of Madrid, Universida Europea de Madrid, IED, ESERP, and IPAM (Oporto, Portugal). David is the secretary general of the Spanish Society for Ibero-American Communication Studies (SEECI), of the History of Information Systems (HISIN), and of the International Forum on Communication and Public Relations (Forum XXI). He is also the principal researcher of the Complutense Concilium validated research group (931.791).

Almudena Barrientos-Báez is a doctor with Cum Laude in Education. She is an assistant doctor at the Complutense University of Madrid and was the director of the Enabling Master's Degree in Teacher Training at the Universidad Europea de Madrid. Almudena holds a master's in management of protocol, production, organization, and design of events—communication area (Universidad Camilo José Cela), a master's in management of tourist accommodation (Universidad de Girona), and degrees in tourism (Universidad La Laguna) and teaching (Universidad de Valencia). Her research work is linked to neuromarketing, neurocommunication, gender, use of social networks, communication, PR, and ICT. She is part of the public audiovisual media project Before the Ecosystem of Platforms: Models of Management and Evaluation of the Public Value of Reference for Spain (PID2021-122386OB-I00), financed by the MCIN, AEI, and FEDER, EU. She is member of the Complutense Concilium validated research group (931.791).

Graciela Padilla-Castillo is coordinator of the degree in journalism and lecturer at the Faculty of Information Sciences, Complutense University of Madrid. She is the director of the journal *Investigaciones Feministas*. She has completed her postdoctoral training at the University of California, Los Angeles (UCLA), with a full scholarship, and has traveled as an expert in gender studies to the Benemérita Universidad Autónoma de Puebla (BUAP, Mexico). She has participated in three European projects, and more than 40 research projects with competitive grants (R&D), private grants, and teaching innovation projects, both as a researcher and as a PI. Graciela is member of the Institute for Feminist Research of Complutense University of Madrid, of the research group 671 (an interdisciplinary group of feminist research), and a member of the research group 962041 (Complutense group of brand management and integrated communication).

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 160–174
https://doi.org/10.17645/pag.v11i2.6328

Article

Spanish Political Communication and Hate Speech on Twitter During the Russian Invasion of Ukraine

Antonio José Baladrón-Pazos [1], Beatriz Correyero-Ruiz [2,*], and Benjamín Manchado-Pérez [3]

[1] Department of Communication Sciences and Sociology, Rey Juan Carlos University, Spain
[2] Department of Journalism, Universidad Católica de Murcia, Spain
[3] Department of Financial and Actuarial Economics & Statistics, Complutense University of Madrid, Spain

* Corresponding author (bcorreyero@ucam.edu)

Submitted: 28 October 2022 | Accepted: 18 January 2023 | Published: 17 May 2023

Abstract
Since the Russian invasion of Ukraine on February 24, 2022, the polarized discourse of the West vs. Russia seems to have escalated to levels similar to those during the Cold War period. The aim of this article, which is centered on the case of Spain, is to discover to what extent communication from political parties contributed to such polarization by encouraging hate speech. To this end, messages sent by the political parties represented in the Spanish parliament, over the social network Twitter during the first 60 days of the Russian invasion of Ukraine, were analyzed: A total of 3,186 tweets from 67 official accounts, both from these political parties and their main leaders, were coded. The results showed that despite social networks in general—and Twitter in particular—being a favorable environment for the promotion of hate speech, the communication of Spanish political parties was generally characterized by political correctness and moderation. The presence of the main indicators of hate speech analyzed (threats, criticism, ridicule, or insults) was very minor. The present article associates this finding to other variables such as the tone of the tweets (informative, opinionated), their scope (international, national), and engagement (replies, quotes, retweets), among other factors. It is concluded that Spanish political authorities had a socially-responsible behavior in the case analyzed, reinforcing the importance of public diplomacy to counteract hate speech.

Keywords
hate speech; polarization; political communication; political parties; political social responsibility; Russia; Spain; Twitter; Ukraine

Issue
This article is part of the issue “Hate Speech, Demonization, Polarization, and Political Social Responsibility” edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

The invasion of Ukraine by Russia in February of 2022 once again underscored the role of social networks in modern wars, as the conflict was to be fought both physically and digitally. The Ukraine war is, in fact, the first viral war, broadcasted in real-time through fragments of images disseminated in social networks. Twitter, Facebook, Instagram, YouTube, and TikTok have become a battlefield in which battles and threats from both sides of the war are being reported instantaneously. Thus, for example, when the invasion was first detected by Google Maps, the newspaper *The Washington Post* tracked the first movement of the Russian troops by using videos uploaded to TikTok by users in Ukraine, and Twitter was the medium selected by president Zelensky to disseminate a video in which he announced that he was not leaving the country. The phrase “We are here. We are in Kyiv. We are protecting Ukraine” (Applebaum, 2022) was found on his Twitter profile.

Social networks have also echoed the general public's rejection of this war, as the polarized discourse of the West vs. Russia seems to have increased to levels similar to those during the Cold War period (Bluhm, 2023; Schwartz, 2022). "Hate speech" has also been used as a tool to shape the image of the invading country: At the international level, Russia faced the fast condemnation by many world leaders, and EU and NATO allies, and many Western companies stopped operating in the country; Vladimir Putin was personified as the "bad guy" (Garner, 2022), becoming the person with the worst reputation on the internet in 2022 (Observatorio Español de Internet, 2022). According to Milosevich-Jurasti (2022), a turning point in the current West–Russia relationship was the definition of the Russian leader as "a war criminal" by Western observers. The first ones to qualify Putin in this manner were the American President Joe Biden and the High Representative of European Union Foreign Policy Josep Borrell.

As opposed to "the villain," the international community qualified Zelensky as "the hero" (González-Martín, 2022; Pereira & Reevell, 2022) and, in this sense, Spanish public opinion during the first 30 days of the war was aligned with the positions of the EU and the national government. According to a poll published in April by the Spanish Center for Sociological Research (Centro de Investigaciones Sociológicas [CIS], 2022a), the Spanish people gave a score of 7.1 to acts perpetrated by Ukraine in the conflict, while Russia scored the worst, 1.5 points—on a scale ranging from 1 (*strongly disagree*) to 10 (*strongly agree*). Spain obtained a simple pass (5.06), as only 39.7% of the population mentioned being "very" or "strongly" in agreement with the position of the Spanish government on the conflict (CIS, 2022b).

As for the role of world leaders in the conflict, the Ukrainian president obtained the best score (7.35) and the Russian leader the worst (1.35). The leader of the Spanish government, Pedro Sanchez, was found second-to-last, ahead only of the Chinese president Xi Jinping (3.44) and Putin himself (CIS, 2022b).

Considering that Spain is the most polarized country in Europe (Gidron et al., 2019) and that political ideology is the third highest cause of discrimination and hate crimes in the country (right after "racism and xenophobia" and "sexual orientation and gender identity"; see Ministerio del Interior, 2021), the present article aims to uncover up to what point communication from Spanish political parties on Twitter contributed to fomenting polarization associated to the Russian-Ukrainian conflict in the first few months of the war, and to analyze if it promoted hate speech.

The phenomenon of political polarization is widespread across many democratic countries due to the predominance of "post-truth communication" (Waisbord, 2018), in which reality is often distorted and emotion prevails over reason, as a way of arguing (Pérez-Escolar & Noguera-Vivo, 2022). This increase in hate was observed in the last report from the Anti-Defamation League (2022), at both the international level and in the case of Spain, the latter within the results from the second national poll of affective polarization (Centro de Estudios Murcianos de Opinión Pública [CEMOP], 2022).

Since January 2020, the government of Spain has been based on a coalition between the Spanish Socialist Workers Party (PSOE), with a center-left ideology, and Unidas Podemos (UP), with a left-to-extreme-left ideology. The government is currently led by Pedro Sanchez, leader of the PSOE and president of the country since 2018, who, after winning the elections in November 2019, came to an agreement with UP that resulted in the first central government coalition in present-day Spanish democracy. Until then, Spanish governments had either been led by the PSOE or the Popular Party (PP), the latter with a center-to-center-right ideology. The PP is, in fact, the main opposition party, followed by Ciudadanos (center-to-center-right), VOX (right-to-extreme right) far behind, and lastly, regional and nationalist parties of varied ideologies. Therefore, we can conclude that, as of today, the Spanish government is tied to a leftist ideology (with the inclusion of members of the Communist Party among its ranks for the first time), while the parliamentary opposition maintains an ideology tied to the right. Table 2 shows specific information on the ideological position of the rest of the parties with representation in the Spanish parliament.

It must be mentioned that all Spanish political parties manifested their opposition to the Ukraine war, although not fully agreeing with some of the measures taken by the government. Thus, for example, the largest volume of tweets of the selected sample corresponded to those published on March 2nd (as shown in Figure 1), the day it was first announced that Spain would send weapons to Ukraine. This decision was met with strong opposition from the leftist parties and generated controversy in the coalition government due to the opposition of the Minister of Social Rights, Ione Belarra, and the Minister of Equality, Irene Montero, both members of UP ("España envía armas," 2022). Belarra qualified the parties who agreed to send weapons as "war parties," and criticized, in a video disseminated on social networks, that the decision could lead to "an uncertain and dangerous scenario of world war" (Podemos, 2022).

The political tension felt during these early days of the conflict can be summarized in tweets such as those from VOX leaders (with their catchphrase: "Putin's allies are in the government"; see VOX, 2022), as well as from the Government Action Secretary of Podemos Pablo Echenique:

> Listening to editorials and talk-show speakers from some media, there's the feeling they have not yet decided if the enemy is Putin (who assassinates Ukrainians) or Unidas Podemos (who works for the cease-fire and peace). Or even worse—that they have decided. (Echenique, 2022a)

Comments such as these show that the greater the ideological polarization—understood as the distance between the ideological positions of the parties—the greater the affective polarization—that is, the distance between the affection we feel towards those with whom we share our political ideas and positions, and feelings of rejection towards supporters of other parties who defend opposing positions (Egea & Garrido, 2022, p. 17).

To set the context of the present research study we can highlight the boycott led by the leftist political parties of President Zelensky's intervention in the Congress of Deputies on April 5, 2022, the day in which the greatest volume of tweets was observed (n = 117) in that month (see Figure 1). On their social networks, two deputies from UP announced they would not attend the event due to their "antifascist conscience." Also, the General Secretary of the Communist Party in Madrid Álvaro Aguilera called out Zelensky on Twitter as a "danger against peace" (Aguilera, 2022). These strategies of confrontation on social media are directly linked to strategic motivations to promote hate speech.

Hate speech can be defined as denigrating language towards people or collectives motivated by race, gender, religion, ideology, etc. (Nockleby, 2000; Waldron, 2012; Whillock & Slayden, 1995). Gagliardone et al. (2015) expand the concept to expressions that promote prejudices, as these can indirectly contribute to the creation of a climate of hostility and polarization that can lead to the use of language at the same level, resulting in a toxic spiral (Amores et al., 2021; Guerrero-Solé & Philippe, 2020). In this sense, the Pyramid of Hate created by the Anti-Defamation League is a useful tool for understanding how hate advances through a continuum, as already mentioned by Allport (1954). Thus, the normalization of biased or prejudiced behaviors (creation of stereotypes, micro-aggressions, insults, etc), can evolve and become the seed of more severe problems such as discrimination, violence based on prejudices, or even genocide. Using the base of the pyramid as the starting point for this study, our research used the following four categories as indicators of hate speech: threats, criticism, ridicule, and insults. Many other tools and lines of research for the detection and analysis of hate speech can be found in the literature (Amores et al., 2021; Pereira-Kohatsu et al., 2019).

Studies have described the role of social platforms as environments that promote hate speech—especially Twitter—due to their active role in the creation of public opinion (e.g., Campos-Domínguez, 2017; Colleoni et al., 2014; Himelboim et al., 2013; Parmelee & Bichard, 2012; Soedarsono et al., 2020). The work presented here provides a point of view centered on Spain, for the study of the use of Twitter in the Russian-Ukraine conflict (Chen & Ferrara, 2022; Donofrio et al., 2022; Morejón-Llamas et al., 2022; Polyzos, 2022; Shevtsov et al., 2022; Smart et al., 2022). An additional aim is to analyze the influence of Twitter on parliamentary communication, a topic that has been the object of analysis in many contemporary studies (e.g., Campos-Domínguez et al., 2022; Dubois & Gaffney, 2014; Esteve Del Valle et al., 2021; Martínez-Rolán & Piñeiro-Otero, 2016).

2. Objectives and Method

The objective of the present study is to identify and categorize the communication of political parties represented in the Spanish parliament on Twitter and to discover to what extent it has promoted hate speech and polarization about one of the most important international events of 2022: the Russian invasion of Ukraine on February 24. The selected timeline comprised the first 60 days of the invasion, from February 24 to June 24, and a total of 67 official accounts from the different political groups and their main leaders (secretary, president, or speaker) were analyzed. In line with this general aim, our specific research objectives were:

Objective 1: To identify messages on Twitter related to the Ukraine invasion published by the political parties represented in the Spanish parliament and their main leaders, and to determine their temporal distribution during the first 60 days of the war.

Objective 2: To code these Twitter messages regarding tone, subject areas, and the type of engagement created—understood as the "index of responses of the users to the comments sent through social networks" (Fontenla-Pedreira et al., 2020, p. 4), that is, the analysis of the replies, quotes, and retweets.

Objective 3: To identify hate speech indicators (threats, criticism, ridicule, insults) in these Twitter messages.

To obtain this data, the free software t-hoarder (Congosto et al., 2017) was used, as well as the set of keywords shown in Table 1. The final sample consisted of 3,186 tweets published during the period analyzed in

Table 1. List of keywords.

Keywords	
arkov	moscu
árkov	moscú
bucha	odesa
dombá	odessa
donb	otan
donet	putin
guerra	rusa
invasi	rusia
kiev	ruso
kremlin	russ
kyiv	ucran
mariupol	ukra
mariúpol	zelen

all the official languages of Spain, containing the aforementioned keywords. The texts were coded following a double-verification method to guarantee the intercoder reliability between two raters, with an average Cohen's kappa coefficient of 0.83. The IBM SPSS v27 program was used for data analysis. Table 2 shows the distribution of the 3,186 tweets from our sample for each of the 67 accounts analyzed during this period, regardless if they were original messages, replies, quotes, or retweets.

3. Results

Starting with Objective 1, we can observe a decrease in the temporal distribution of tweets about the Russian invasion of Ukraine during the period analyzed, as shown in Figure 1; nevertheless, 37.9% of the publications were concentrated in the first 10 days, and this number reached 59% when we look at the first 20 days. In any case, it is also observed that this trend fluctuated downward on a daily basis week after week, without considering if this was due to relevant or newsworthy events associated with the evolution of the conflict. An exception, however, was day 41 of the period analyzed (April 5, when Ukraine President Zelensky addressed the Spanish Congress remotely).

Arranged by political party (see Table 3), we see that the ruling party PSOE was the most active party on Twitter, responsible for 20.8% of these messages. The PSOE was followed by Ciudadanos (15.2%), the center-to-center-right Basque Nationalist Party (or PNV, responsible for 7.4% of these messages), VOX (5.9%), and the left-to-extreme left EH Bildu (5.8%). The main opposition party, PP, was only responsible for 3.7% of these tweets, a percentage similar to that of other parties with a minority representation in the Spanish parliament. UP was not very active on Twitter on this subject, accounting for only 5% of these messages.

Objective 2 of our research was to characterize communications on Twitter about the Russian invasion of Ukraine during the first 60 days of the conflict in terms of the tone of the messages, subject area, and the type of engagement they created. In this respect, we were limited to the type of tweets and we did not consider their engagement rate. To characterize the tone of the messages, a differentiation was made between informative, opinionated, and mixed tweets: Informative messages were defined as messages in which the Twitter user—political party or political leader—presented or recounted events about the Russian invasion of Ukraine; opinionated messages were those in which the user offered his or her assessment or position about the events. Thus, it must be clarified that the mere presence of an assessment or opinion did not imply an opinionated message (tone), as long as the opinion did not correspond to that of the user, who only disseminated it for informative purposes (for example, when a political party publishes a declaration or an opinion from a third

Table 2. Twitter accounts of the parties and politicians analyzed.

Account	*N*	%	Account	*N*	%	Account	*N*	%
@CiudadanosCs	461	15	@NestorRego	41	1.3	@prcantabria	15	0.5
@PSOE	457	14	@Santi_ABASCAL	36	1.1	@andoniortuzar	13	0.4
@eajpnv	168	5.3	@Ortega_Smith	35	1.1	@FeijooGalicia	13	0.4
@socialistes_cat	157	4.9	@KRLS	31	1	@Pdemocratacat	13	0.4
@ehbildu	113	3.5	@cupnacional	30	0.9	@EnComu_Podem	12	0.4
@PabloEchenique	103	3.2	@gabrielrufian	30	0.9	@MarinaBS_Cs	12	0.4
@CristinaNarbona	99	3.1	@ionebelarra	29	0.9	@mariadolorsa	11	0.3
@JuntsXCat	83	2.6	@TeresaRodr_	29	0.9	@InesArrimadas	10	0.3
@populares	82	2.6	@FClavijoBatlle	27	0.8	@upn_navarra	8	0.3
@Esquerra_ERC	81	2.5	@Yolanda_Diaz_	27	0.8	@ArnaldoOtegi	7	0.2
@vox_es	77	2.4	@anioramas	24	0.8	@AdelanteAND	6	0.2
@aramateix	65	2	@compromis	24	0.8	@josep_rius	6	0.2
@MertxeAizpurua	65	2	@cucagamarra	24	0.8	@Elisendalamany	5	0.2
@obloque	62	1.9	@EnComun_Gal	23	0.7	@adrianpumares	4	0.1
@coalicion	58	1.8	@enricmorera	22	0.7	@davidbonvehi	4	0.1
@AITOR_ESTEBAN	55	1.7	@salvadorilla	22	0.7	@junqueras	4	0.1
@Hectorgomezh	54	1.7	@ierrejon	20	0.6	@navarra_suma	4	0.1
@Adrilastra	53	1.7	@Ferran_Bel	17	0.5	@TeruelExiste_	4	0.1
@monicaoltra	52	1.6	@FOROAsturias	16	0.5	@JoseMariaMazon	3	0.1
@Nueva_Canarias	45	1.4	@joanbaldovi	16	0.5	@SobiranistesCat	3	0.1
@ivanedlm	41	1.3	@anaponton	15	0.5	@CarmenMoriyon	2	0.1
@jessicaalbiach	41	1.3	@MasPais_Es	15	0.5	@jordisanchezp	1	0
						@RevillaMiguelA	1	0
						TOTAL	3186	100

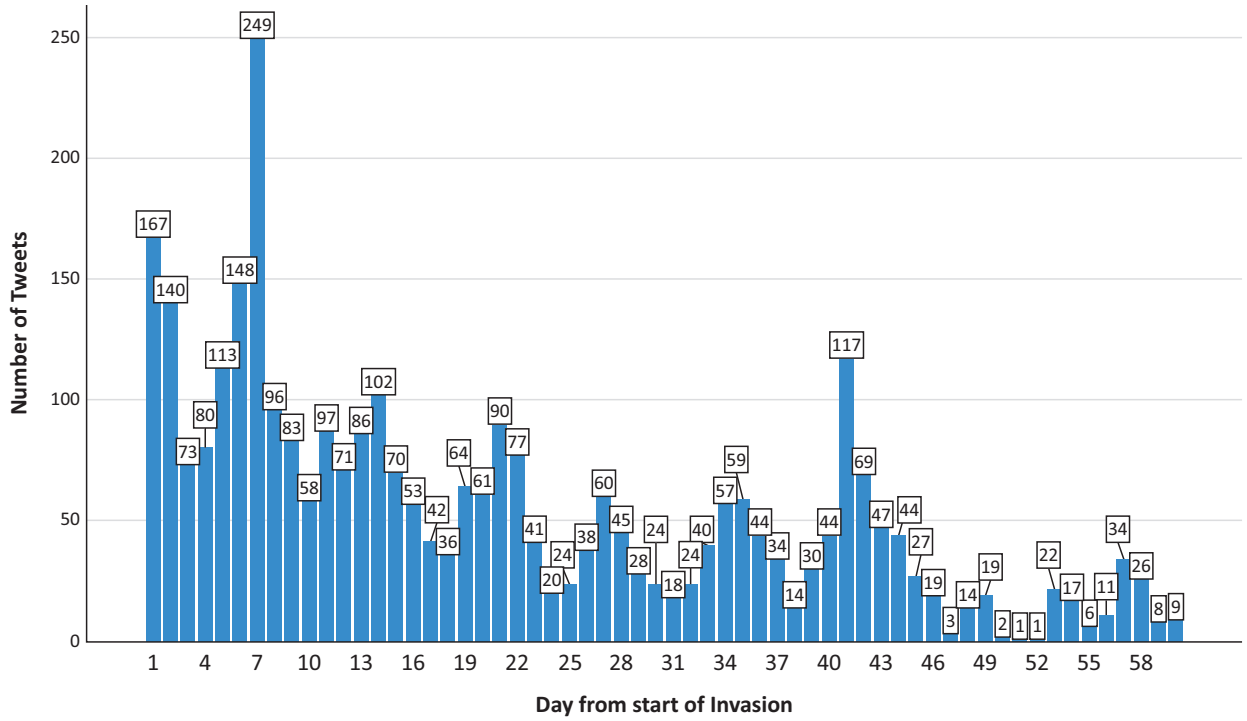


Figure 1. Temporal distribution of the publication of tweets.

Table 3. Distribution of the tweets per political party.

Denomination	Acronym	Ideological position	*N*	%
Partido Socialista Obrero Español	PSOE	center-left	663	20.80%
Ciudadanos	CS	center to center-right	483	15.20%
Partido Nacionalista Vasco	PNV	center-right to center-left	236	7.40%
VOX	VOX	right to extreme right	189	5.90%
EH Bildu	EHB	left to extreme left	185	5.80%
Partido de los Socialistas de Cataluña	PSC	center-left	179	5.60%
Unidas Podemos	UP	left to extreme left	159	5.00%
Junts per Catalunya	JXC	center-right	121	3.80%
Partido Popular	PP	center-right to right	119	3.70%
Bloque Nacionalista Galego	BNG	left	118	3.70%
Esquerra Republicana de Catalunya	ERC	center-left	115	3.60%
Compromís	COMPR	left	114	3.60%
Coalición Canaria	CCAN	center to center-right	109	3.40%
Candidatura de Unidad Popular	CUP	left to extreme-left	106	3.30%
En Común Podem	ECPOD	left	53	1.70%
Nueva Canarias	NCAN	center-left to left	45	1.40%
Más País	MASP	center-left to left	35	1.10%
Adelante Andalucía	AAND	left	35	1.10%
Partido Demócrata Europeo Catalán	PDCAT	center to center-right	34	1.10%
Galicia en Común	ECGAL	left	23	0.70%
Foro Asturias	FORO	center-right to right	22	0.70%
Partido Regionalista de Cantabria	PRC	center to center-left	19	0.60%
Sobiranistes	SOBIR	left	8	0.30%
Unión del Pueblo Navarro	UPN	center-right to right	8	0.30%
Teruel Existe	TER	combined	4	0.10%
Navarra Suma	NAVSUM	center-right to right	4	0.10%
TOTAL			3186	100.00%

party between quote marks). The "mixed" category was used for messages in which the presentation/recounting of events and the assessment or positioning of the broadcaster were combined in a single message. Our conclusion was that the tone of the tweets was more opinionated (45.6%) than informative (41%), also considering that, in the remaining 13.4% of the "mixed tweets," the aim of the user went beyond providing information (they also often provided their opinions), and that some form of opinion or assessment could already be found in plenty of informative tweets. When separating the data according to political party, the differences were substantial.

When we take into account the 10 political parties with the greatest representation in the Spanish parliament, it can be clearly observed that some of them kept their messages informative in tone, while others used a more opinionated tone (see also Table 4). For the PSOE, 51.1% of its tweets were informative (as compared to the mean of 41% for all parties). The PSOE was followed by the main opposition party, the PP (51.3%), and the PNV (68.2%). A more opinionated tone was found in the

Table 4. Tone of the tweets.

		Tone			
Party		Informative	Opinionated	Mixed	Total
UP	*N*	22	119	18	159
	% PART_NUM	13.80%	74.80%	11.30%	100.00%
	% TONE	1.70%	8.20%	4.20%	5.00%
PP	*N*	61	54	4	119
	% PART_NUM	51.30%	45.40%	3.40%	100.00%
	% TONE	4.70%	3.70%	0.90%	3.70%
VOX	*N*	50	132	7	189
	% PART_NUM	26.50%	69.80%	3.70%	100.00%
	% TONE	3.80%	9.10%	1.60%	5.90%
ERC	*N*	45	64	6	115
	% PART_NUM	39.10%	55.70%	5.20%	100.00%
	% TONE	3.40%	4.40%	1.40%	3.60%
PSOE	*N*	339	247	77	663
	% PART_NUM	51.10%	37.30%	11.60%	100.00%
	% TONE	25.90%	17.00%	18.00%	20.80%
CS	*N*	128	235	120	483
	% PART_NUM	26.50%	48.70%	24.80%	100.00%
	% TONE	9.80%	16.20%	28.10%	15.20%
EHB	*N*	82	67	36	185
	% PART_NUM	44.30%	36.20%	19.50%	100.00%
	% TONE	6.30%	4.60%	8.40%	5.80%
PNV	*N*	161	63	12	236
	% PART_NUM	68.20%	26.70%	5.10%	100.00%
	% TONE	12.30%	4.30%	2.80%	7.40%
MASP	*N*	8	23	4	35
	% PART_NUM	22.90%	65.70%	11.40%	100.00%
	% TONE	0.60%	1.60%	0.90%	1.10%
JXC	*N*	49	57	15	121
	% PART_NUM	40.50%	47.10%	12.40%	100.00%
	% TONE	3.70%	3.90%	3.50%	3.80%
Other parties	*N*	362	391	128	881
	% PART_NUM	41,09%	44,38%	14,53%	100,00%
	% TONE	27,80%	27,00%	30,20%	27,70%
Total of all parties	*N*	1307	1452	427	3186
	% PART_NUM	41,00%	45,60%	13,40%	100,00%
	% TONE	100,00%	100,00%	100,00%	100,00%

more ideologically-extreme parties, such as UP (74.8% as compared to a mean of 45.6% for all parties) and VOX (69.8%). Ciudadanos deserves special mention, as it used mixed tweets to the detriment of merely informative tweets (24% as compared to the mean of 13.4% for all parties). Through a simple correspondence analysis, a visual representation was obtained of the informative or opinionated tones of all the involved parties, as shown in Figure 2.

In terms of subject, the tweets were categorized as international, national, or mixed. International tweets were centered on worldwide politics or realities, referring to the war in the international arena or the consequences of the conflict beyond Spain. National tweets were focused on Spanish national politics, or the Spanish reality, and addressed the consequences of the war as they were felt in Spain (on many occasions from the economic point of view or referring to national issues that were somehow associated with the conflict). The "mixed" category was reserved for tweets that combined both national and international dimensions. The results from the top 10 Spanish political parties (see Table 5), point out that the subject area was predominantly international: 79.2% of the cases, as compared to 19.7% of the tweets referring to national matters, and a marginal 1.1% with a mixed focus. No significant differences were found in the subject area of the messages regarding tone. Also, no great differences were found in the subject area addressed by the tweets from each political party analyzed, although the parties that addressed the national reality to a greater extent were PP, VOX, PSOE, and EH Bildu, while the ones that frequently focused their messages on the international arena were Ciudadanos, Más País (center-left to left), Esquerra Republicana de Catalunya (center-left), and PNV.

With respect to the type of engagement created, it is possible to observe that 57.7% corresponded to retweets and 36.2% to original messages, with the remaining responses or cited tweets only reaching 6.2%. Likewise, regarding tone, we found that, in the original messages, the informative tone was more frequent than the opinionated one and, conversely, the opinionated tone was more frequent on the retweets. Lastly, the results indicated that when correlating the type of engagement and subject area of the tweets (see Table 6), retweets were less frequent for messages centered on the national reality (46.0% of the cases, as compared to 57.7% of the mean) and that, on the contrary, original messages were more predominant than the mean in the case of tweets focused on the national arena (46.8% and 36.2% respectively).

Objective 3 of the present research was to identify hate speech indicators in Twitter communications from political parties with representation in the Spanish parliament during the first 60 days of the conflict. To this end, the following four indicators were considered: threats (speech that states the desire to harm someone or something); criticism (messages that speak badly of someone

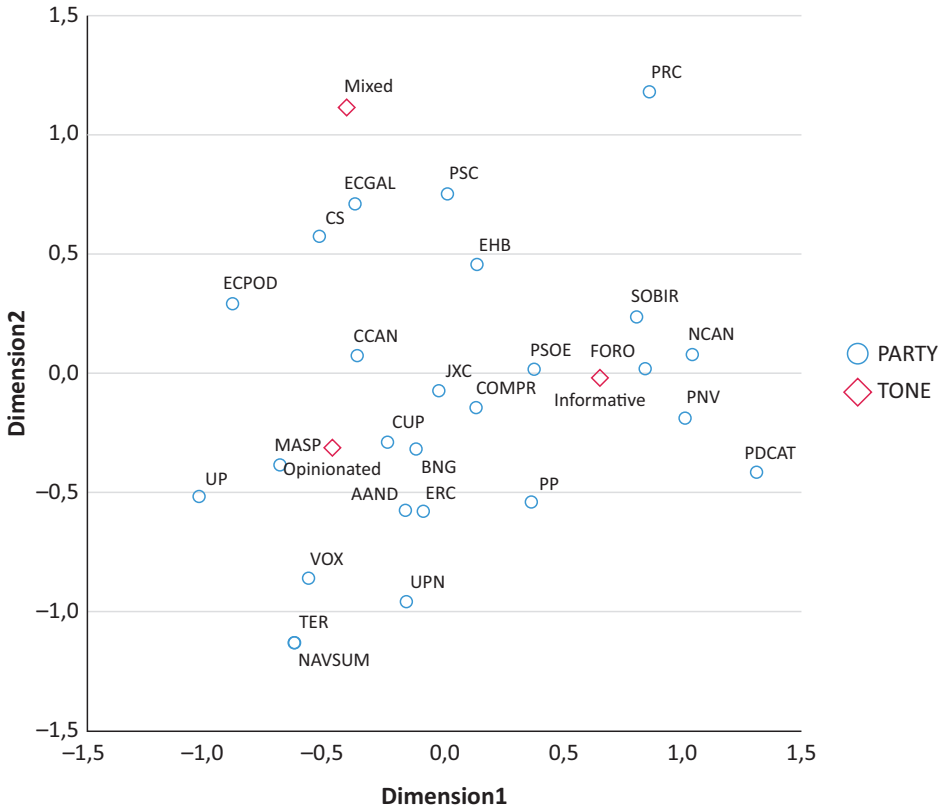


Figure 2. Simple correspondence analysis of the tones found in the messages.

Table 5. Subject area of the tweets.

Party		Area			
		National	International	Mixed	Total
UP	*N*	32	121	6	159
	% PART_NUM	20.10%	76.10%	3.80%	100.00%
	% AREA	5.10%	4.80%	16.70%	5.00%
PP	*N*	28	90	1	119
	% PART_NUM	23.50%	75.60%	0.80%	100.00%
	% AREA	4.50%	3.60%	2.80%	3.70%
VOX	*N*	44	138	7	189
	% PART_NUM	23.30%	73.00%	3.70%	100.00%
	% AREA	7.00%	5.50%	19.40%	5.90%
ERC	*N*	19	96	0	115
	% PART_NUM	16.50%	83.50%	0.00%	100.00%
	% AREA	3.00%	3.80%	0.00%	3.60%
PSOE	*N*	152	504	7	663
	% PART_NUM	22.90%	76.00%	1.10%	100.00%
	% AREA	24.20%	20.00%	19.40%	20.80%
CS	*N*	42	438	3	483
	% PART_NUM	8.70%	90.70%	0.60%	100.00%
	% AREA	6.70%	17.40%	8.30%	15.20%
EHB	*N*	39	145	1	185
	% PART_NUM	21.10%	78.40%	0.50%	100.00%
	% AREA	6.20%	5.70%	2.80%	5.80%
PNV	*N*	43	192	1	236
	% PART_NUM	18.20%	81.40%	0.40%	100.00%
	% AREA	6.80%	7.60%	2.80%	7.40%
MASP	*N*	4	31	0	35
	% PART_NUM	11.40%	88.60%	0.00%	100.00%
	% AREA	0.60%	1.20%	0.00%	1.10%
JXC	*N*	6	115	0	121
	% PART_NUM	5.00%	95.00%	0.00%	100.00%
	% AREA	1.00%	4.60%	0.00%	3.80%
Other parties	*N*	219	652	10	881
	% PART_NUM	24,86%	74,01%	1,14%	100,00%
	% AREA	34,90%	25,80%	27,80%	27,70%
Total all parties	*N*	628	2522	36	3186
	% PART_NUM	19,70%	79,20%	1,10%	100,00%
	% AREA	100,00%	100,00%	100,00%	100,00%

or something, inflicting damage, but without insults or ridicule); ridicule (speech ridiculing someone or something mockingly, but without using insults); and insults (messages that offend someone or something, with hurtful expressions or words). Considering the meaning granted to these four indicators in the coding process, the presence of ridicule was found in 2.5% of the tweets, insults in 1.8%, criticism in 1%, and threats in 0.2%. Thus, the presence of hate speech was small, without these indicators being found in 94.5% of the 3,186 tweets about the Russian invasion of Ukraine disseminated by political parties with representation in the Spanish parliament.

When separating the data according to the tone of the messages (see Table 7), it was observed that in informative and mixed tweets, there was a smaller presence of hate speech indicators than in opinionated ones; if these indicators were not found in 94.5% of the total messages, in the case of the opinionated ones, this value was reduced to 89.5%. As for hate speech indicators per subject area, no significant differences were found; a higher percentage of these indicators were

Table 6. Engagement and subject area of the tweets.

		Area			
Engagement		National	International	Mixed	Total
Original	*N*	294	848	11	1153
	% ENG	25.50%	73.50%	1.00%	100.00%
	% AREA	46.80%	33.60%	30.60%	36.20%
Reply	*N*	29	64	8	101
	% ENG	28.70%	63.40%	7.90%	100.00%
	% AREA	4.60%	2.50%	22.20%	3.20%
Quote	*N*	16	78	1	95
	% ENG	16.80%	82.10%	1.10%	100.00%
	% AREA	2.50%	3.10%	2.80%	3.00%
Retweet	*N*	289	1532	16	1837
	% ENG	15.70%	83.40%	0.90%	100.00%
	% AREA	46.00%	60.70%	44.40%	57.70%
Total	*N*	628	2522	36	3186
	% ENG	19.70%	79.20%	1.10%	100.00%
	% AREA	100.00%	100.00%	100.00%	100.00%

found in messages that were mixed in nature, although we must consider the small number of this type of tweets. Likewise, we must also take into account that we cannot confirm that the use of hate speech indicators led to a greater or lesser level of engagement, specifically due to the reduced number of tweets with these indicators, which could not provide us with conclusive results (see Table 8).

The indicators of hate speech according to its use by political parties can be found in Table 9. The political parties with the smallest percentage of hate speech indicators in their tweets were the following, in descending order: PSOE (98.6% versus a mean of 94.5%), EH Bildu (97.8%), Más País (97.1%), PP (96.6%), PNV (96.6%), and Junts Per Catalunya (95.9%). On the contrary, the ones that published a smaller number of tweets without hate indicators (always below the total mean), were the following, in ascending order: VOX (74.1%), UP (83.6%), and Ciudadanos (91.3%).

Through a simple correspondence analysis, a visual representation was obtained for each of the four hate speech indicators analyzed (threat, criticism, ridicule, and insult), and each of the political parties in the Spanish parliament, independently of the number of seats (see Figure 3). This figure confirms, in general terms, what was previously mentioned about the scarce presence of these indicators, thereby adding a more qualitative perspective. However, it must be pointed out that UP and Sobiranistes were closer to ridicule, while VOX was closer to insult and criticism.

Table 7. Hate speech indicators according to the tone of the messages.

		Speech					
Tone		Threat	Criticism	Ridicule	Insult	None	Total
Informative	*N*	0	0	1	7	1299	1307
	% TONE	0.00%	0.00%	0.10%	0.50%	99.40%	100.00%
	% SPEECH	0.00%	0.00%	1.30%	12.30%	43.10%	41.00%
Opinionated	*N*	0	30	77	46	1299	1452
	% TONE	0.00%	2.10%	5.30%	3.20%	89.50%	100.00%
	% SPEECH	0.00%	96.80%	96.30%	80.70%	43.10%	45.60%
Mixed	*N*	6	1	2	4	414	427
	% TONE	1.40%	0.20%	0.50%	0.90%	97.00%	100.00%
	% SPEECH	100.00%	3.20%	2.50%	7.00%	13.70%	13.40%
Total	*N*	6	31	80	57	3012	3186
	% TONE	0.20%	1.00%	2.50%	1.80%	94.50%	100.00%
	% SPEECH	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 8. Indicators of hate speech according to engagement.

		Speech					
Engagement		Threat	Criticism	Ridicule	Insult	None	Total
Original	*N*	3	3	25	13	1109	1153
	% ENG	0.30%	0.30%	2.20%	1.10%	96.20%	100.00%
	% SPEECH	50.00%	9.70%	31.30%	22.80%	36.80%	36.20%
Reply	*N*	0	3	7	1	90	101
	% ENG	0.00%	3.00%	6.90%	1.00%	89.10%	100.00%
	% SPEECH	0.00%	9.70%	8.80%	1.80%	3.00%	3.20%
Quote	*N*	0	4	8	7	76	95
	% ENG	0.00%	4.20%	8.40%	7.40%	80.00%	100.00%
	% SPEECH	0.00%	12.90%	10.00%	12.30%	2.50%	3.00%
Retweet	*N*	3	21	40	36	1737	1837
	% ENG	0.20%	1.10%	2.20%	2.00%	94.60%	100.00%
	% SPEECH	50.00%	67.70%	50.00%	63.20%	57.70%	57.70%
Total	*N*	6	31	80	57	3012	3186
	% ENG	0.20%	1.00%	2.50%	1.80%	94.50%	100.00%
	% SPEECH	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 9. Indicators of hate speech according to its use by political parties.

		Speech					
Party		Threat	Criticism	Ridicule	Insult	None	Total
UP	*N*	0	1	21	4	133	159
	% PART_NUM	0.00%	0.60%	13.20%	2.50%	83.60%	100.00%
	% SPEECH	0.00%	3.20%	26.30%	7.00%	4.40%	5.00%
PP	*N*	0	0	2	2	115	119
	% PART_NUM	0.00%	0.00%	1.70%	1.70%	96.60%	100.00%
	% SPEECH	0.00%	0.00%	2.50%	3.50%	3.80%	3.70%
VOX	*N*	0	13	24	12	140	189
	% PART_NUM	0.00%	6.90%	12.70%	6.30%	74.10%	100.00%
	% SPEECH	0.00%	41.90%	30.00%	21.10%	4.60%	5.90%
ERC	*N*	0	0	3	3	109	115
	% PART_NUM	0.00%	0.00%	2.60%	2.60%	94.80%	100.00%
	% SPEECH	0.00%	0.00%	3.80%	5.30%	3.60%	3.60%
PSOE	*N*	2	1	4	2	654	663
	% PART_NUM	0.30%	0.20%	0.60%	0.30%	98.60%	100.00%
	% SPEECH	33.30%	3.20%	5.00%	3.50%	21.70%	20.80%
CS	*N*	3	9	11	19	441	483
	% PART_NUM	0.60%	1.90%	2.30%	3.90%	91.30%	100.00%
	% SPEECH	50.00%	29.00%	13.80%	33.30%	14.60%	15.20%
EHB	*N*	0	4	0	0	181	185
	% PART_NUM	0.00%	2.20%	0.00%	0.00%	97.80%	100.00%
	% SPEECH	0.00%	12.90%	0.00%	0.00%	6.00%	5.80%
PNV	*N*	0	1	3	4	228	236
	% PART_NUM	0.00%	0.40%	1.30%	1.70%	96.60%	100.00%
	% SPEECH	0.00%	3.20%	3.80%	7.00%	7.60%	7.40%
MASP	*N*	1	0	0	0	34	35
	% PART_NUM	2.90%	0.00%	0.00%	0.00%	97.10%	100.00%
	% SPEECH	16.70%	0.00%	0.00%	0.00%	1.10%	1.10%

Table 9. (Cont.) Indicators of hate speech according to its use by political parties.

Party		Speech					
		Threat	Criticism	Ridicule	Insult	None	Total
JXC	*N*	0	1	1	3	116	121
	% PART_NUM	0.00%	0.80%	0.80%	2.50%	95.90%	100.00%
	% SPEECH	0.00%	3.20%	1.30%	5.30%	3.90%	3.80%
Other parties	*N*	0	1	11	8	861	881
	% PART_NUM	0,00%	0,11%	1,25%	0,91%	97,73%	100,00%
	% SPEECH	0,00%	3,40%	13,50%	14,00%	28,70%	27,70%
Total all parties	*N*	6	31	80	57	3012	3186
	% PART_NUM	0,20%	1,00%	2,50%	1,80%	94,50%	100,00%
	% SPEECH	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

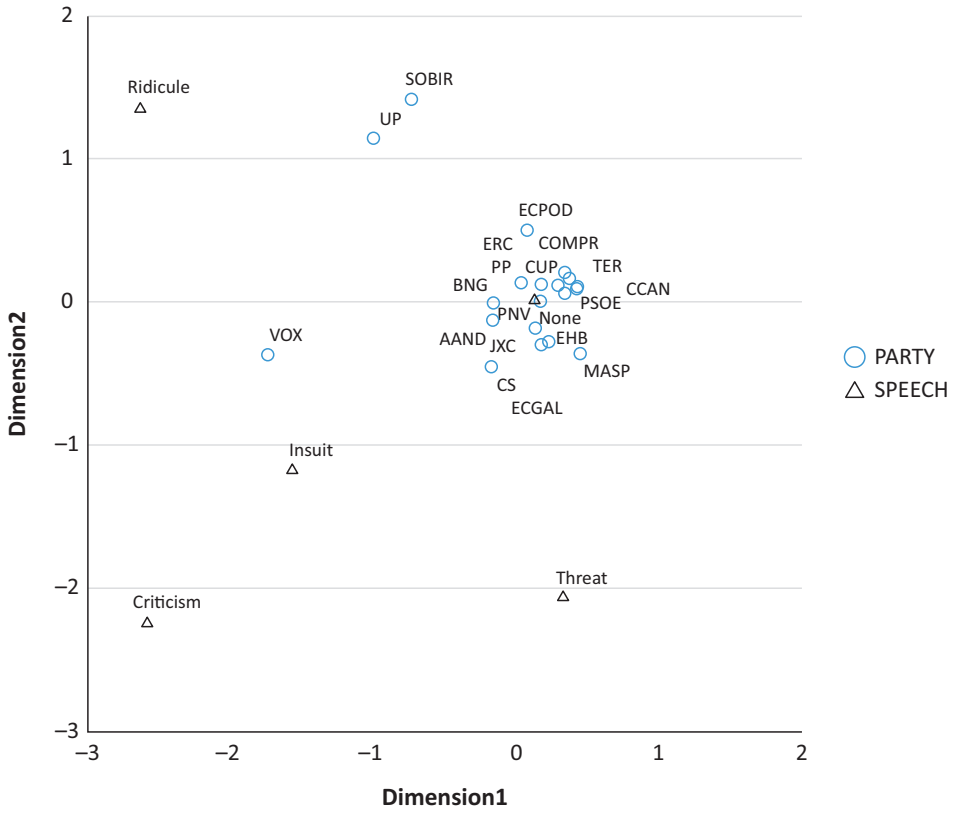


Figure 3. Simple correspondence analysis of the hate speech indicators according to the parties.

4. Conclusions

Our analysis shows that the Twitter debate on the Russian-Ukrainian conflict, between Spanish political parties, had a low level of toxicity during the first 60 days of the war (Guerrero-Solé & Philippe, 2020). Therefore, it can be deduced that most Spanish political leaders/forces, including the governing party, showed a high level of social responsibility to avoid tensions.

Spanish political parties used Twitter with a greater intention to opine than inform. All of these political parties took to Twitter to express their opposition to the war in Ukraine, although at the same time defending their different ideological positions. Thus, the main parties in the country granted more importance to sharing informative messages on Twitter, while parties with the most extreme ideologies attributed greater significance to sharing opinion pieces, which demonstrates the positive association between ideological polarization and affective polarization: Greater ideological polarization results in greater affective polarization (Orriols, 2021; Torcal & Comellas, 2022). Most of the tweets analyzed tended to offer a negative image of Russia, given its conflicting positioning against European and democratic interests. Those who openly identified their position did so in favor of Ukraine. Most of the messages analyzed

also addressed the humanitarian component of the conflict and argued for the need to end the confrontations through a peace agreement. We can conclude, then, that the opinion of politicians was mostly aligned with that of Spanish society in general (CIS, 2022a). These types of dichotomous discourses were also observed in other studies that analyzed the dissemination of political messages in armed conflicts (Moreno-Mercado et al., 2022; Orhan, 2020). We also concluded that Twitter messages published by the more extreme parties focused on matters adjacent to the war itself, which mostly dealt with national political polarization.

In general, Spanish political parties with more extreme ideologies promoted hate speech to a greater degree, although the study revealed that these types of messages did not generate a greater level of engagement as compared to those that did not promote that type of speech; citizens for whom these messages were intended also wagered for moderate and socially responsible communication.

More than 80% of the tweets analyzed referred to the international arena (these messages created the greatest engagement), and only 17% alluded to the national impact of the war. Although correlations were not found between the national/international dimension of the messages and each of the political parties, some significant differences were observed, especially when referring to the national character of specific events—as a result of the ideological position of each party—which shows how different parties and their leaders bid for framing public interest matters through ideology and the use of social networks.

The results obtained also demonstrate that Spanish polarization was less related to partisanship than to ideological blocs, which translates into aversion or animosity towards the parties of the opposite bloc (Garrido et al., 2021, p. 277), as shown in tweets that tended to disseminate unverified rumors or accusations between political adversaries (see, e.g., Echenique, 2022b; VOX, 2022). Thus, the polarizing discursive mechanisms that were observed in the sample of tweets analyzed correspond to some of the seven strategies proposed by Marín-Albaladejo (2022), namely: dichotomous simplification; demonization; victimization; reporting of conspiracies; disinformation; and promotion of subjects that contribute towards increasing tension. These strategies of confrontation on the internet encourage hate speech.

Although the present article did not address all the possible indicators of hate speech, as it only reflected on four of them (threats, criticism, ridicule, and insult), the study could be broadened to encompass other types of parameters addressed in other methodologies. This would help us to more strongly corroborate the scarce presence of hate speed in the sample analyzed.

As for the extent to which communication from Spanish political parties on Twitter contributed towards promoting polarization associated with the Russian-Ukrainian conflict in the first 60 days of the war, the results obtained were very reasonable, especially if we consider the international dimension of the event. The months following these first 60 days should also be analyzed, especially considering the impact of the war on Spain—on the national economy, as a recurring subject in the media, and in the Spanish political agenda. Future studies should be conducted to draw meaningful inferences on foreign policy matters.

Conflict of Interests

The authors declare no conflicts of interest.

References

Aguilera, A. [@AlvaroAgFa]. (2022, April 5). *Lo siento pero no. Esto no. Zelenski es un peligro para la paz y su pueblo. Heredero de un golpe* [I'm sorry, but no. Not this. Zelenski is a danger against peace and his people. Heir to a coup] [Tweet]. Twitter. https://twitter.com/AlvaroAgFa/status/1511373707314438149

Allport, G. W. (1954). *The nature of prejudice.* Addison-Wesley.

Amores, J. J., Blanco-Herrero, D., Sánchez-Holgado, P., & Frías-Vázquez, M. (2021). Detecting ideological hatred on Twitter. Development and evaluation of a political ideology hate speech detector in tweets in Spanish. *Cuadernos*, *49*, 98–124. https://doi.org/10.7764/cdi.49.27817

Anti-Defamation League. (2022). *Online hate and harassment: The American experience 2022*. https://www.adl.org/resources/report/online-hate-and-harassment-american-experience-2022

Applebaum, A. [@anneapplebaum]. (2022, February 25). *Zelensky: We are here, we are in Kyiv, we are defending Ukraine* [Tweet]. Twitter. https://twitter.com/anneapplebaum/status/1497260281214423046

Bluhm, M. (2023, January 24). Are we in a new Cold War? *Vox*. https://www.vox.com/23568071/are-we-in-a-new-cold-war-russia-ukraine

Campos-Domínguez, E. (2017). Twitter y la comunicación política [Twitter and political communication]. *Profesional de la información*, *26*(5), 785–794. https://doi.org/10.3145/epi.2017.sep.01

Campos-Domínguez, E., Esteve Del Valle, M., & Renedo-Farpón, C. (2022). Rhetoric of parliamentary disinformation on Twitter. *Comunicar*, *72*, 47–58. https://doi.org/10.3916/C72-2022-04

Centro de Estudios Murcianos de Opinión Pública. (2022). *II Encuesta Nacional de Polarización Política* [II national survey on political polarization]. https://www.cemopmurcia.es/estudios/ii-encuesta-nacional-de-polarizacion-politica

Centro de Investigaciones Sociológicas. (2022a). *Encuesta sobre cuestiones de actualidad. Marzo 2022* [Survey on current issues. March 2022] (Working Paper No. 3358). https://www.cis.es/cis/export/

sites/default/-Archivos/Marginales/3340_3359/3358/es3358mar.pdf

Centro de Investigaciones Sociológicas. (2022b). *Barómetro de marzo 2022* [Barometer March 2022] (Working Paper No. 3355). https://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3340_3359/3355/es3355mar.pdf

Chen, E., & Ferrara, E. (2022). *Tweets in time of conflict: A public dataset tracking the twitter discourse on the war between Ukraine and Russia*. ArXiv. https://doi.org/10.48550/arXiv.2203.07488

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication*, *64*(2), 317–332. https://doi.org/10.1111/jcom.12084

Congosto, M., Basanta-Val, P., & Sánchez-Fernández, L. (2017). T-Hoarder: A framework to process Twitter data streams. *Journal of Network and Computer Applications*, *83*, 28–39. https://doi.org/10.1016/j.jnca.2017.01.029

Donofrio, A., Rubio, Á. L., & Abellán, C. (2022). Rusia–Ucrania, un análisis comparativo de la audiencia en Twitter de los perfiles del Gobierno de la Federación Rusa y la Oficina del Presidente de Ucrania [Russia–Ukraine, a comparative analysis of the Twitter audience of the profiles of the Government of the Russian Federation and the Office of the President of Ukraine]. *Revista Latina de Comunicación Social*, *81*, 18–43. https://www.doi.org/10.4185/RLCS-2022-1819

Dubois, E., & Gaffney, D. (2014). The multiple facets of influence: Identifying political influentials and opinion leaders on Twitter. *American Behavioral Scientist*, *58*(10), 1260–1277. https://doi.org/10.1177/0002764214527088

Echenique, P. [@PabloEchenique]. (2022a, March 4a). *Escuchando los editoriales y a los tertulianos de algunos medios de comunicación, da la sensación de que todavía no han decidido* [Listening to editorials and talk-show speakers from some media, there's the feeling they have not yet decided] [Tweet]. Twitter. https://twitter.com/PabloEchenique/status/1499686800738168832

Echenique, P. [@PabloEchenique]. (2022b, March 4b). *La relación entre VOX y Putin es clarísima—Misma ideología: anti-LGTBI, ultranacionalista, antifeminista* [The relationship between VOX and Putin is very clear—Same ideology: anti-LGTBI, ultranationalist, anti-feminist] [Tweet]. Twitter. https://twitter.com/pabloechenique/status/1499838555832606723?lang=es

Egea, M., & Garrido, A. (2022). La grieta: Polarización ideológica y afectiva en el debate político español [The rift: Ideological and affective polarization in the Spanish political debate]. In E. Arroyas-Langa, P. L. Pérez-Díaz, & M. Pérez-Escolar (Eds.), *El debate público en la red: Polarización, consenso y discursos del odio* [Online public debate: Polarization, consensus and hate speech] (pp. 13–31). Comunicación Social Ediciones y Publicaciones.

España envía armas a Ucrania en medio de las discrepancias en el Gobierno y dentro de Unidas Podemos [Spain sends arms to Ukraine amid disagreements in the government and within Unidas Podemos]. (2022, March 4). *RTVE*. https://www.rtve.es/noticias/20220304/espana-envia-armas-ucrania-medio-discrepancias-gobierno-dentro-unidas-podemos/2301929.shtml

Esteve Del Valle, M., Broersma, M., & Ponsioen, A. (2021). Political interaction beyond party lines: Communication ties and party polarization in parliamentary twitter networks. *Social Science Computer Review*, *40*(3), 736–755. https://doi.org/10.1177/0894439320987569

Fontenla-Pedreira, J., Rúas-Araújo, J., & Conde-Vázquez, E. (2020). Social media audience of televised debates in the general elections of April 2019. *Revista Latina de Comunicación Social*, *76*, 1–16. http://doi.org/10.4185/RLCS-2020-1434

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. UNESCO Publishing.

Garner, I. (2022, March 4). How is the war going for Putin on social media? Not great. *The Washington Post.* https://www.washingtonpost.com/outlook/2022/03/04/russian-social-media-ukraine-vk-propaganda

Garrido, A., Martínez-Rodríguez, M. A., & Mora-Rodríguez, A. (2021). Polarización afectiva en España [Affective polarization in Spain]. *Más Poder Local*, 45, 21–40. https://www.maspoderlocal.com/index.php/mpl/article/view/polarizacion-afectiva-en-espana-mpl45

Gidron, N., Adams, J., & Horne, W. (2019, October 18). *How ideology, economics and institutions shape affective polarization in democratic polities* [Paper presentation]. Annual Seminar on the State and the Capitalism, Center for European Studies Harvard, Cambridge, MA, USA. https://ces.fas.harvard.edu/uploads/files/events/GAH-Affective-Polarization-in-Democratic-Polities.pdf

González-Martín, A. (2022, May 11). ¿Es Volodímir Zelenski el héroe necesario o no? [Is Volodymyr Zelensky the necessary hero or not?] (IEEE Paper No. 32). IEEE. https://www.ieee.es/Galerias/fichero/docs_analisis/2022/DIEEEA32_2022.pdf

Guerrero-Solé, F., & Philippe, O. (2020). La toxicidad de la política española en Twitter durante la pandemia de la Covid-19 [The toxicity of Spanish politics on Twitter during the Covid-19 pandemic]. *HTXT*, 2020(21), 133–139. https://doi.org/10.31009/hipertext.net.2020.i21.12

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communi-*

cation, *18*(2), 154–174. https://doi.org/10.1111/jcc4.12001

Marín-Albaladejo, J. A. (2022). La polarización discursiva como estrategia de comunicación en las cuentas de líderes y partidos políticos [Discursive polarization as a communication strategy in the accounts of political leaders and parties]. In E. Arroyas-Langa, P.L. Pérez-Díaz, & M. Pérez-Escolar (Eds.), *El debate público en la red: polarización, consenso y discursos del odio* [Online public debate: Polarization, consensus and hate speech] (pp. 51–68). Comunicación Social Ediciones y Publicaciones.

Martínez-Rolán, X., & Piñeiro-Otero, T. (2016). The use of memes in the discourse of political parties on Twitter: Analysing the 2015 state of the nation debate. *Communication & Society*, *29*(1), 145–160. https://doi.org/10.15581/003.29.35935

Milosevich-Jurasti, M. (2022, April 26). El balance actual de la guerra en Ucrania y las perspectivas de un acuerdo de paz. *Real Instituto el Cano.* https://www.realinstitutoelcano.org/analisis/el-balance-actual-de-la-guerra-en-ucrania-y-las-perspectivas-de-un-acuerdo-de-paz

Ministerio del Interior. (2021). *Informe de Evolución de los Delitos de Odio en España* [Report on the evolution of hate crimes in Spain]. https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/publicaciones-periodicas/informe-sobre-la-evolucion-de-los-delitos-de-odio-en-Espana/Informe_evolucion_delitos_odio_Espana_2021_126200207.pdf

Morejón-Llamas, N., Martín-Ramallal, P., & Micaletto-Belda, J. P. (2022). Twitter content curation as an antidote to hybrid warfare during Russia's invasion of Ukraine. *Profesional de la información*, *31*(3). https://doi.org/10.3145/epi.2022.may.08

Moreno-Mercado, J. M., García-Marín, J., & Luengo, Ó. G. (2022). Conflictos armados y la construcción de narrativas a través de Twitter. El caso de la guerra entre Armenia y Azerbaiyán [Armed conflicts and the construction of narratives on Twitter: The case of the war between Armenia and Azerbaijan]. *Revista Española de Ciencia Política*, *59*, 89–114. https://doi.org/10.21308/recp.59.04

Nockleby, J. T. (2000). Hate speech. In L. W. Levy, K. L. Karst, & A. Winkler (Eds.), *Encyclopedia of the American Constitution* (2nd ed., pp. 1277–1299). Macmillan.

Observatorio Español de Internet. (2022, March 30). *Vladímir Putin ya es la marca política con peor reputación en Internet* [Vladimir Putin is already the political brand with the worst online reputation]. http://observatorioespanoldeinternet.com/?s=putin

Orhan, D. (2020). Making foreign policy through Twitter: An analysis of Trump's Tweets on Iran. In E. Esiyok (Ed.), *Handbook of research on new media applications in public relations and advertising*. IGI Global. https://doi.org/10.4018/978-1-7998-3201-0.ch022

Orriols, L. (2021). *La polarización afectiva en España: Bloques ideológicos enfrentados* [Affective polarization in Spain: opposing ideological blocs] (Report No. 28). EsadeEcPol. https://itemsweb.esade.edu/research/Policy-insight-polarizacion-afectiva.pdf

Parmelee, J. H., & Bichard, S. L. (2012). *Politics and the Twitter revolution: How tweets influence the relationship between political leaders and the public.* Lexington Books.

Pereira, I., & Reevell, P. (2022, February 25). What to know about Ukrainian president Volodymyr Zelenskyy. *ABC News*. https://abcnews.go.com/International/ukrainian-president-volodymyr-zelenskyy/story?id=83085078

Pereira-Kohatsu, J. C., Quijano-Sánchez, L., Liberatore, F., & Camacho-Collados, M. (2019). Detecting and monitoring hate speech in twitter. *Sensors*, *19*(21). https://doi.org/10.3390/s19214654

Pérez-Escolar, M., & Noguera-Vivo, J. M. (Eds.). (2022). *Hate speech and polarizationin participatory society*. Routledge.

Podemos [@PODEMOS]. (2022, March 3). *La única elección actual es entre negociación y diplomacia o un conflicto mundial con consecuencias imprevisibles* [The only choice today is between negotiation and diplomacy or a global conflict with unpredictable consequences] [Tweet]. Twitter. https://twitter.com/PODEMOS/status/1499426679256670213?ref_src=twsrc%5Etfw

Polyzos, E. S. (2022). *Escalating tension and the war in Ukraine: Evidence using impulse response functions on economic indicators and Twitter sentiment*. *SSRN*. http://doi.org/10.2139/ssrn.4058364

Shevtsov, A., Tzagkarakis, C., Antonakaki, D., Pratikakis, P., & Ioannidis, S. (2022). *Twitter dataset on the Russo-Ukrainian war*. ArXiv. https://doi.org/10.48550/arXiv.2204.08530

Smart, B., Watt, J., Benedetti, S., Mitchell, L., & Roughang, M. (2022). *#IStandWithPutin versus #IStandWithUkraine: The interaction of bots and humans in discussion of the Russia/Ukraine war*. ArXiv. https://doi.org/10.48550/arXiv.2208.07038

Soedarsono, D. K., Mohamad, B. B., Akanmu, D. M., & Putri, I. P. (2020). Political leaders and followers' attitudes: Twitter as a tool for political communication. *Journal of Critical Reviews*, *7*(8), 1245–1252.

Schwartz, M. (2022, April 3). Cold War 2.0 has already begun—and it's going to be even scarier than the first one. *Business Insider.* https://www.businessinsider.com/cold-war-russia-china-ukraine-america-nuclear-scarier-foreign-policy-2022-3

Torcal, M., & Comellas, J. M. (2022). Affective polarization in times of political instability and conflict: Spain from a comparative perspective. *South European Society and Politics*, *27*(1), 1–26. https://doi.org/10.1080/13608746.2022.2044236

VOX. [@vox_es]. (2022, February 25). *Lastra, Enrique Santiago, Irene Montero….Los aliados de Putin están en el Gobierno y Pedro Sánchez debe expulsarlos inmediatamente* [Lastra, Enrique Santiago, Irene Montero….Putin's allies are in the government and Pedro Sánchez must expel them immediately] [Tweet]. Twitter. https://twitter.com/vox_es/status/1497185372467568663

Waisbord, S. (2018). The elective affinity between post-truth communication and populist politics. *Communication Research and Practice*, *4*(1), 17–34. https://doi.org/10.1080/22041451.2018.1428928

Waldron, J. (2012). *The harm in hate speech*. Harvard University Press.

Whillock, R. K., & Slayden, D. (1995). *Hate speech*. SAGE.

About the Authors

Antonio José Baladrón-Pazos (PhD) holds a doctorate in advertising and public relations and two bachelor's degrees: one in journalism from the Pontifical University of Salamanca and one in advertising and public relations from the University of Vigo. He has received an Extraordinary Doctorate Award and a Research Award on mass communication from the Consell de l'Audiovisual de Cataluña, and has more than twenty years of experience in teaching and research. He has been a visiting professor at different foreign universities, both European and American, and participated in numerous academic activities.

Beatriz Correyero-Ruiz (PhD) is vice-dean of the Journalism Programme at the Catholic University of Murcia (UCAM) and a member of the UCAM research group on Communication, Politics and Image. She holds a doctorate in journalism from the Complutense University of Madrid and has more than twenty years of experience in teaching and research. Her lines of research are the study of state propaganda through tourism, radio, and multimedia journalism.

Benjamín Manchado-Pérez (PhD) earned a doctorate from the Department of Financial and Actuarial Economics & Statistics at the Complutense University of Madrid, where he is currently collaborating as an external researcher. He has co-authored several scientific articles on econometrics and advertising research, using bibliometric methods and multivariate analysis. He has over 25 years of professional experience in different sectors: university teaching, private banking, asset management, construction, aerospace engineering, agrifood, and in the healthcare industry.

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 175–186
https://doi.org/10.17645/pag.v11i2.6369

Article

Voter's Perception of Political Messages Against the Elite Classes in Spain: A Quasi-Experimental Design

Juan Enrique Gonzálvez-Vallés [1,*], José Daniel Barquero-Cabrero [2], and Natalia Enseñat-Bibiloni [2]

[1] Communication Theories and Analysis, Complutense University of Madrid, Spain
[2] Department of Research, ESERP Business & Law School, Spain

* Corresponding author (jegonzalvez@ucm.es)

Submitted: 29 October 2022 | Accepted: 13 March 2023 | Published: 17 May 2023

Abstract
Demonization constitutes the construction of a simplified symbolic reality through the prism of hatred. The elite classes have been framed within the discourse of far-left parties in Spain to make them antagonistic to their target audience. In this area, research tends to use explicit measures, but few have used implicit measures. This article aims to assess both types of voter perceptions of anti-elite class messages in Spain. Eighty Spanish voters belonging to left- and right-wing ideologies participated in this study. Implicit attitudes were measured through the implicit association test. The result shows that demonization is deepening in various layers of society in order to provoke a confrontation with the elite classes. Social networks have emerged as a means of social control, but they are not perceived as such by the subjects but rather as an additional informational or purely propagandistic element. But even among the public who are aware of the latter, the more their perception of the elite classes is limited to demonization, the more extreme their social vision is.

Keywords
demonization; elite classes; hate speech; polarization; political discourse; social networks

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

The current development of technologies, as well as their potential, provides great advances and benefits that cannot be ignored in almost any field, including the media (Tuesta-Panduro, 2021). However, the media has been expanding and continues to expand its contact channels to provide information and, especially within the framework of web 2.0, this has led to harmful effects such as the creation and subsequent viralization of fake news and the emergence of hate messages (Rani et al., 2022). Such pernicious information is created by combining typographic and visual elements that accentuate the manipulative intention (Rodríguez Ferrándiz & Sorolla-Romero, 2021).

However, disinformation is not only framed from the point of view of the media; it extends to other areas such as corporate communication, institutional communication, and political communication (Mut Camacho & Rueda Lozano, 2021). What sets a common pattern in all the aforementioned areas is the possibility of building communicative strategies that favor disinformation, boost hate speech, and contribute to the polarization of public opinion (Gidron et al., 2020; López del Castillo Wilderbeek, 2021). Thus, new ecosystems of information manipulation are generated where society is unaware of reality and fixes the framework through pernicious messages (Larrondo-Ureta et al., 2021).

When polarization occurs, the center of the board is usually emptied, and a great deal of stress is put on

the ends. It occurs in any aspect where a communication strategy is developed, but due to its impact on public opinion, it reaches higher levels in political communication (Romero-Rodríguez et al., 2021). Through messages that demonize opponents and hate speech, society becomes fragmented and self-realities are constructed in which only those with the same opinion have a place. The stereotype is transformed into a normative element that is cultural, social, and religious in nature, cornering the political opponent through disqualifications and making them an enemy in all the previously mentioned areas (Baptista et al., 2021).

It is at this point that hate speech emerges. Moreover, it has found social networks an excellent space to become visible and expand exponentially. These two circumstances mean that, in this aspect, web 2.0 stands as a particularly sensitive loudspeaker when it comes to political issues (Urcola Eceiza & Elezgarai Ibañez, 2021). Groups of individuals, sometimes acting in a coordinated fashion, give value to hate messages that include aggressive, violent, and discriminatory terms. The low quality of information on social networks is a complex issue for society as it affects certain groups and weakens democracy (Abdullah et al., 2022).

While there is no universal definition of hate speech, given the moral and legal implications it may have, it refers to expressions that foster prejudice or intolerance and indirectly contribute to a hostile climate that may ultimately foster discrimination or even violent behavior or attack (Civila et al., 2021). These discourses single out a group of individuals based on certain characteristics by attributing to their addressee a set of constitutive qualities that are considered highly undesirable. They tend not necessarily to resort to violence but may also use more refined forms, such as jokes, insinuations, or ambiguous images (Romero Rodríguez et al., 2021).

Spain is the most polarized country in Europe, with high ideological polarization but low emotional polarization. More than two-thirds of citizens believe that traditional political parties and politicians do not care about people like them (Alonso, 2021). In this sense, it has been observed that the degree of polarization of citizens is often associated with belonging to certain social groups and the consumption of political news, which may occur due to contagion effects and polarization of acceptance by political actors (Jaráiz et al., 2021).

Hate speech is also one of the most important discursive tactics used by populist movements to promote their actions and agendas (Alonso-Muñoz & Casero-Ripollés, 2021; Bartha et al., 2020; Olivas Osuna, 2021). As such, elites are at the center of hate speech in far-left parties in Europe (Marcos-Marne, 2021). In the Spanish case, the attack focuses on discrediting the business class as a social agent, attacking its political representatives, which are identified with the right wing or radical right, and denying the violence of the authoritarian and populist states of the far left (Coll, 2021; Fischer, 2020; Iglesias, 2022; Ruiz Coll, 2022). It is also worth noting that Twitter's activity on the issue of caste has had a lot of activity and interactions, as recent studies have shown, helping to link issues such as the alleged link between elite, caste, business, and disenfranchisement (Montessori & Morales-López, 2019; Tierno, 2019).

A theoretical assumption of this study is that frequent contact between voters and people belonging to the elite classes can lead to changes in their beliefs and behaviors. In essence, this model adopts the contact hypothesis, one of the most effective strategies in social psychology for reducing prejudice (Brown & Hewstone, 2005). The contact hypothesis states that contact between members of different groups can generate positive intergroup relations and a noticeable decrease in prejudice and negative stereotypes (Allport, 1954). The literature has shown that the quantity and quality of contact positively affect perceptions when the person with whom contact is maintained is perceived as representative of their outgroup (Brown & Hewstone, 2005).

1.1. Use of the Political Discourse Analysis

As a socio-political phenomenon, social polarization has aroused the interest of a number of social science researchers. In this framework, some research has investigated electoral scenarios: elections in Germany and the US (Barberá, 2015). They have also focused on analyzing political and party divisions, such as those between left and right in Spain (Barberá, 2015; Ben-David & Fernández, 2016). These studies highlight how political, economic, and social conjunctures become a fertile space for the emergence and study of polarization (Banks et al., 2021; Pérez-Escolar & Noguera-Vivo, 2022).

Other studies have investigated voter perceptions from various perspectives. One of these analyzes them through the strategies implemented by political parties (Nasr, 2020). However, other research focuses on the analysis of these perceptions focused on the figure of the party leader and candidate (Fernández-Vázquez & Somer-Topcu, 2019). Recent studies delve into the latter line with new variables such as segmentation by race (Chirco & Buchanan, 2023) and the entry of new leaders into traditional parties that had previously enjoyed stable leadership (Bridgewater, 2023).

Other studies warn of the need to establish the subtle difference between social polarization and society's simple use of ideological stereotypes (such as left or right) to make sense of reality from established referents. That is, not all people who use ideological labels to designate political events are polarised; some only find in these labels a way of signifying experiences from social frameworks (Banks et al., 2021; Schmid et al., 2022). However, when it is indeed polarisation, it is associated with ruptures in the social fabric and naturalized violence in everyday relationships, as suggested by previous research in Spain (Ben-David & Fernández, 2016; Pérez-Escolar & Noguera-Vivo, 2022).

The media are important because they are the means by which the parties involved in the conflict present a positive image of themselves and a devalued image of others. They deploy tactics such as disinformation, propaganda, and explicit social control mechanisms to manipulate information, preventing a diverse range of perspectives on the phenomenon from being explored. Moreover, they can promote mobilizations in favor of or against one of the parties, exacerbate emotions, mold patriotic attitudes and identifications with a group of victims, and be shapers of the construction of subjective reality (Dunmire, 2012).

Other authors argue that the media directs public opinion's attention towards particular issues and objects on the political and social scene, setting public agendas and manipulating the political scene by influencing public behavior to create a certain climate for action (Wang, 2016). In this sense, media influence shifts attention, conversations, and political decisions from a socio-political to a socio-psychological level, where the focus is on citizens' attitudes. Emotions are mobilized, and it is hoped that, based on these emotions, people will support or oppose a political project (Xu & Ballena, 2022).

Other authors do not reach definitive conclusions regarding how the media can influence the configuration of attitudes, beliefs, or narratives regarding political actors (Brown & Klein, 2020). However, other research concludes that, while the media do not make people more polarized, they do increase the rigidity of those who are. Thus, people are drawn closer to information that reinforces their beliefs and affiliation to a political group in a selective exposure that solidifies the polarization of the general public, not only of their television audience, but also in those who receive those messages, both personally, in discussions and conversations, and through their social networks (Shokrollahi et al., 2021).

Moreover, media reports that have taken sides in political situations are not primarily intended to communicate facts but to help people make sense of the world, given particular predispositions that are beyond the possibility of dialogue with divergent arguments. Polarization is built by reaffirming these acquired positions and beliefs through the messages that are broadcast and strengthened in the absence of counterarguments. This pushes people toward ideological extremes and polarises attitudes (Pavlichenko, 2022; Rachman & Yunianti, 2017).

For his part, Manheim (2004) states that the mainstream media, controlled by large economic groups, replicate the ideological alignments of civil society with the elites. This allows for control, homogenization of the narrative, and construction of public events in general, to the detriment of pluralism of information. In the same sense, Fanoulis and Guerra (2021) define them as ideological apparatuses that have the capacity to determine the thinking, behavior, and action of the audience, creating semantics that focuses on a single enemy, causing civil society to take sides. This handling of information influences the construction of citizen attitudes in order to mobilize possible endorsement and legitimization or, on the contrary, rejection and delegitimization. Thus, the media are not just mere informants but active actors that construct realities and, in some cases, have become polarising agents contributing, through their discourses, to the construction of the enemy imaginary (Eroğlu Utku & Köroğlu, 2020).

1.2. Hate Speech and Its Relation to Disinformation Against the Elite Classes

Most of the works that analyze the political identity of far-left parties in Spain through their discourse give it one main characteristic: that it is populist because it is based on the people–elite dichotomy (Barbeito Iglesias & Iglesias Alonso, 2021; Custodi, 2021; Stoehrel, 2016).

Until 2021, the representation at the national level of these formations fell to Podemos. A formation founded in January 2014, with the immediate objective of running in the European elections of May 25th. Unexpectedly, it gained almost 8% of the votes, causing a political commotion (Stobart, 2014). In those elections, the combined vote of the Partido Popular and Partido Socialista did not reach 50% of the votes for the first time. Podemos' discourse, especially in televised debates, went on the offensive, setting the political agenda and introducing new terms in everyday political language, such as caste, to refer to the elites. However, in renewing his message against the elitist classes, Podemos positioned the new-lower democracy in opposition to the old-arrow oligarchy (Brandariz Portela, 2021).

The hate speech is emphasized from the first declarations of the leader of the purple party, Pablo Iglesias, who affirms before the media that corruption makes this system work since it allows the economic elites, "incapable and predatory," to use the political elites for their interests. More than a political system, this order is a regime, a closed system dominated by a caste that excludes the majority (Ivaldi et al., 2017; Podemos València, 2014).

Podemos will associate the homeland with the defense of the people and the welfare state, of what equalizes and unites, opposing the right, which it identifies with divisive symbols: the flag, the Castilian language, or traditions. The right does not defend the public but privatizes, evades taxes, and hoards capital in tax havens (Manetto, 2014). The far-left party proposes an up–down axis that represents the dichotomy between democracy as the government of the people and democracy as competition between elites. The elites and their representatives appear in the discourse as a source of hatred since they are constituted as usufructuaries of institutions for their interests and which are not representative of reality since they are in other locations physically (Europa Press, 2021; Fernández Holgado et al., 2022). This presupposes that these institutions are radically favorable to the elites and conducive to the hijacking of democracy (Jansa & Vilasero, 2015).

The concept of homeland is redefined to pit some Spaniards against others. They construct a narrative where stereotypes about elites in general and the right wing, in particular, are supported by calling them unpatriotic (Fernández Vázquez, 2022). The leaders of the splits within the extreme left amplify their discourse by providing misinformation about who is patriotic and who should not be. In this sense, national identity excludes the elites who decide what happens in Spain without going through the ballot box (Errejón, 2021; Mudde & Rovira Kaltwasser, 2013).

2. Objectives and Hypotheses

Assuming, therefore, the existence of an increasingly intense bias of hatred towards the elite classes on the part of the extreme-left parties and that it is polarized in the receivers of the messages, who have the possibility of exercising their right to vote, the present research has the following objectives:

- O1: To analyze the implicit bias in the reception of messages emitted by politicians of the Spanish extreme left that demonize the elite classes;
- O2: To make a significant contribution to the study of hate speech as an informational disorder, providing an assessment of the prejudices, stereotypes, and affections perceived by voters about the elite classes.

The three starting hypotheses are:

- H1: Left-wing voters, when receiving messages through heuristics, exhibit more implicit attitudes toward people in the elite classes compared to right-wing voters.
- H2: Left-wing voters will show more negative biases, stereotypes, and affection toward the elite classes compared to their right-wing peers.
- H3: Voters who report more contact, both in quantity and quality, with people belonging to the elite classes will show less explicit negative attitudes toward this group.

3. Materials and Methods

A 2 × 2 quasi-experimental design was used to frame the discourse of extreme left-wing political parties in Spain (left-wing voters vs. right-wing voters) × type of stimulus (photos of messages on Twitter from the elite classes [ingroup] vs. photos of messages on Twitter by extreme left-wing politicians [outgroup]). To make up the sample, a non-probabilistic method was used to count 40 right-wing and 40 left-wing voting subjects. Their ages ranged from 18 to 75 years (A = 44.3), 67.5% being women (n = 54) and 32.5% men (n = 26).

On arrival at the experiment, all participants declared whether they were left- or right-wing voters and this was recorded in the software used. All participants signed an informed consent form at the start of the study. All were informed that the objective was to evaluate voters' perceptions of different social classes and discourses produced about them, assuring them that their responses would be anonymous. After the study, they were thanked for their collaboration.

The research focused on the study of public conversation on the social network Twitter. The most representative terms from previous research (caste, homeland, rich) were used to search the messages of the extreme-left politicians. For the messages of the elite classes, we searched for the messages of corporate events that had the most interaction. The FreeIAT program was used for the design and, to carry out the measurement, each subject was placed in front of a laptop computer. The stimuli were: six photographs of Twitter messages by people from the elite classes, six photographs of Twitter messages by politicians of the extreme left (see Figure 1), seven words related to honesty, and seven

Figure 1. Examples of the Twitter stimuli used in the implicit association test (IAT).

words inciting hatred from the messages or speeches of extreme left-wing politicians.

The subjects of the elite class were chosen through their inclusion in the Merco (2022) ranking and the extreme left-wing politicians as the members with the highest number of followers in social networks of the Podemos and Más País formations since their foundation and who had obtained a seat in the Congress of Deputies. For the choice of words, we used the *Diccionario de Sinónimos y Antónimos María Moliner* (Moliner et al., 2012; see Figure 2). The scheme was structured in seven blocks, with five training blocks (1, 2, 3, 5, and 6) with 20 stimuli each; and two central blocks (3 and 6) composed of 24 stimuli.

Figure 2. Examples of words used in the IAT. Notes: *Honrado*, *íntegro*, and *decente* translates to honoured, upright, and decent, respectively; *corrupto*, *inmoral*, and s*invergüenza* translates to corrupt, immoral, and shameless, respectively.

In Block 1, the task consisted of discriminating between photos of people belonging to the elite classes and politicians of extreme left-wing parties. In Block 2, words related to honesty and those related to hate were classified. In Block 3, the preceding stimuli were combined as a test. The purpose was to choose between pairs of combined categories: pictures of people belonging to the elite classes and words related to honesty, and pictures of extreme left-wing politicians and words related to hate. The same stimuli were responded to in Block 4, but these responses were taken for the final analysis. In Block 5, the order of presentation of the stimuli from Block 1 was altered, the task consisting in discriminating between photos belonging to people from the elite classes and politicians from extreme left-wing formations. Blocks 6 and 7 again presented combinations of categories, but in inverted order. Block 6 presented photos of extreme left-wing politicians and words related to honesty, and photos of people from the elite classes and words related to hate. Finally, in Block 7, this pattern was repeated, but the responses were considered for the final analysis. In each of the seven blocks, participants responded to test items that were not computed in the final analysis. The interval between the individual's response and the appearance of the next stimulus was 250 ms.

At the culmination of the IAT application, subjects were asked to complete a questionnaire with different self-report scales with items adapted from different instruments. The scales were: prejudice, stereotypes, and affection. A total of 15 statements (five for each scale) from Twitter messages of extreme-left politicians were used, and the response was measured by a seven-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The alpha for prejudice was 0.72, for stereotypes 0.73, and for affection 0.79. To assess the quantity and quality of the scale "contact with people in the elite classes" (α = 0.71), five other statements were established and also measured using a seven-point Likert scale ranging from 1 (*not at all*) to 7 (*very much*).

Before proceeding with statistical analyses to determine the presence or absence of implicit bias, a purification of the database obtained by the IAT was performed. Extremely fast or very slow responses were reconverted, and the percentage of erroneous responses was calculated. The response times were subjected to a logarithmic transformation to normalize their distribution. Following the strategy of Greenwald et al. (1998), very fast responses (below 300 ms) appear when the participant presses the key before the stimulus appears on the screen, and very slow responses (above 3,000 ms), that correspond to brief periods of participant distraction, were recategorized (Nosek et al., 2007).

4. Results

Table 1 shows the IAT reaction times of subjects of both political tendencies. To test the hypothesis referring to right-wing voters not evidencing implicit attitudes towards far left-wing politicians compared to left-wing voters, a repeated measures ANOVA was performed. The data allowed us to accept the working hypothesis since one sample showed an implicit bias toward the elite classes (F = 1.104; sig. = 0.003). On the other hand, the main effect referred to the political tendency was significant (F = 1.054; sig. = 0.004), and also the interaction between the IAT and the persons shown (F = 1.078; sig. = 0.03).

Table 1. Average reaction times.

Type of persons	Associations	Incompatible associations	Compatible IAT effect
Elitist class people	873.43	801.11	72.32
	(213.25)	(132.65)	(187.84)
Far-left politicians	187.85	160.32	27.53
	(254.47)	(231.32)	(175.01)

Notes: N = 80, p < 0.05; standard deviations are shown in parentheses.

Table 2 shows the results for all the scales used in this study. The results show that the two samples present disparate results. First, it can be seen that left-wing voters were characterized by expressing a great deal of prejudice towards people from the elite classes, while no differences were found, however, in the prejudice shown towards extreme left-wing politicians. Differences were found in the stereotype variable, in which right-wing voters rated far left-wing politicians less strongly in this regard compared to the sample of far left-wing voters regarding people from the elite classes. In the affections towards the outgroup, differences were found in the group of left-wing voters, who expressed less affection towards people from the elite classes compared to right-wing voters.

Table 3 records the correlations between explicit and implicit measures for each sample. First, the correlations between the IAT and the explicit measures were significant. In the sample of left-wing voters, it stands out that affection is negatively associated with stereotypes toward people in the elite classes. Similarly, affection is negatively associated with prejudice toward people in the elite classes. In the sample of right-wing voters, a striking result is observed: a weak correlation of affection with the perception of stereotyping, as well as with a weak attitude of prejudice towards extreme left-wing politicians. It is also interesting to note that stereotypes correlated positively with prejudice.

Multiple regression analyses were performed to assess the impact of contact variables and voters' ideological orientation on implicit (IAT score) and explicit (prejudice, stereotypes, and affection) attitudes toward people in the elite classes. A single index was computed for stereotypes and affection toward people in the elite classes. The hate (stereotypes) and aversion (affection) scores were inverted to align them in a positive direction. Initially, the Durbin-Watson statistic was checked to evaluate the assumption of independence of the errors. In all cases, the values were close to two, verifying the fulfillment of this assumption. Furthermore, it was verified that the explanatory variables had a variance inflation factor of less than three, which ruled out the presence of multicollinearity. The predictors were centered before fitting the regression model.

The data indicated that, on the one hand, voter type, hate, aversion, and prejudice, independently, and, on the other hand, the hate-by-aversion interaction, showed significant effects on some of the criterion variables. Left-wing voters presented greater prejudice ($\beta = 0.34$, $t = 2.00$, $p < 0.05$) and greater hatred ($\beta = 0.75$, $t = 2.93$, $p < 0.05$) toward people in the elite classes. On the other hand, it was observed that participants who reported having hatred toward people in the elite classes expressed more dislike toward members of this group ($\beta = 0.46$, $t = 1.37$, $p < 0.05$). Likewise, left-wing voters who indicated having more prejudice toward people in the elite classes felt more dislike toward these people ($\beta = 0.28$, $t = 1.91$, $p < 0.05$). Finally, voters who reported greater prejudice toward the elite classes attributed greater amounts of hatred toward them ($\beta = 0.65$, $t = 3.45$, $p < 0.05$).

5. Discussion and Conclusions

These findings do not contradict other studies on hate speech and its impact on society (Amores et al., 2021; Correcher Mira, 2021; Montero et al., 2022; Verloo, 2018), but they do complement them from a different point of view. First, the study of perception using neuro-communication tools is innovative and allows us to assess the impact of these discourses on implicit and explicit attitudes toward elite classes. In this way, the framework of the political discourse of extreme left-wing

Table 2. Averages, standard deviations, and ANOVA tests for both types of voters.

	Left-wing voters (n = 40)		Right-wing voters (n = 40)		
Scales	Average	Standard deviation	Average	Standard deviation	F
Prejudice	3.56	0.32	1.22	0.33	15.84
Stereotypes	5.02	0.83	2.35	1.02	9.84
Affections	4.65	1.13	1.89	0.78	65.53

Note: $p < 0.05$.

Table 3. Correlations between the IAT and the scales according to voter type.

	IAT	Prejudice	Stereotypes	Affections
IAT	—	0.521	0.635	0.532
Prejudice	0.560	—	0.621	−0.531
Stereotypes	0.488	0.839	—	−0.403
Affections	0.653	0.137	0.139	—

Notes: Correlations above the diagonal belong to left-wing voters and correlations below the diagonal to right-wing voters (n = 40 for each sample, $p < 0.05$).

parties in Spain is constructed, taking into account all the participating agents.

The first hypothesis tested in this research was that left-wing voters when receiving messages through heuristics would present more implicit attitudes towards people from the elite classes compared to right-wing voters. The analyses point to the presence of differences between the groups, so it can be affirmed that the presence of implicit bias is not independent of the group of voters to which one belongs. This finding leads this research towards the question of whether these negative implicit associations had already been internalized by left-wing voters, even before being subjected to the communicational impacts present on Twitter. The literature on implicit social cognition indicates that from an early age and through various sources of socialization, people harbor simplified representations of minority groups (Serrano Villalobos & Cuesta Díaz, 2017). Thus, in the face of implicit experimental techniques such as the IAT, negative associations regarding these groups often emerge (Berberena & Wirzberger, 2021). As several authors point out (Cox & Devine, in press; Huber, 2020), a long time must elapse and profound societal changes must occur for these implicit attitudes to be substantially altered.

As the results at the implicit level show differences, one should not lose sight of the discussion on whether most of the decisions humans make tend to be rational and deliberative (Ross, 1981; Turiel, 2010). Thus, the scrutiny of implicit and explicit attitudes allows us to observe what could be the dangers of hate speech generated toward the elite classes in this attitudinal dimension. In this sense, the hypothesis that left-wing voters would show a more negative implicit and explicit evaluation of people from the elite classes compared to right-wing voters was fully accepted.

In general, left-wing voters obtained high levels of prejudice, stereotypes, and negative affection toward people from the elite classes. These data are not surprising since extreme left-wing parties in Spain have displayed hate-related terms such as privileged caste, corrupt, and unpatriotic among their communicative policies. However, differences were found in two affective variables on which we wish to focus the analysis: hate and aversion. Research shows that hatred is a very important affective variable in the prediction of prejudice (Gonzálvez Vallés, 2013). In general, when members of the majority interact with people who possess some stigma, one of the most notorious affective reactions is usually hatred, which can result in a brake for a positive interaction with people who do not correlate with their ideology (Amores et al., 2021; Rodríguez Terceño et al., 2016). The results reveal that there is a greater presence of hatred among left-wing voters associated with the attribution of more aversion towards people from the elite classes.

In terms of prejudice, it was observed that left-wing voters showed more prejudice towards people belonging to the elite class. This result could appear intuitive in the first instance; however, the interpretation changes when weak correlations between prejudice and affection are observed in right-wing voters. This belief would be expressing an ambivalent content of prejudices towards people of the elite classes. In this regard, several authors (Escobar Beltrán, 2016; Gomes Barbosa, 2022) have pointed out that many social groups are described in positive and negative terms at the same time. When interacting with social groups perceived as warm but not competent, people do not usually express openly hostile stereotypes, but rather these appear mixed with attributes linked to pity and compassion. In this sense, the right-wing voters in the study would be attributing these co-religionist stereotypes to people from the elite classes, as compared to their left-wing voting peers, and according to the messages by which they have been impacted. On the other hand, the latter, having a more biased view towards the elites, may judge them as an enemy or adversary.

Finally, the hypothesis that analyzes the contact dimensions (quantity and quality) was also accepted. Left-wing voters have less contact with the elites, and more prejudice and hatred towards these types of subjects were seen. However, it could not be established which aspects of the contact could be causing this hate effect. These data are consistent with other studies conducted in sociological contexts, in which it has been shown that the quantity, first, and then the quality of contact contributes to the generation of attitudes compatible with hatred and aversion (Tropp et al., 2022).

A result that confirms the hypothesis was the inverse relationship between the perception of greater contact with the category and the attribution of stereotypes toward people in the elite class. This data possibly reflects the responses of voters who do not regularly interact with the elite. In general, it has been observed that when there is partial knowledge of individuals in an outgroup, in-group dynamics can contaminate the contact situation by invoking negative reactions (Meleady, 2021). Supporting the above, it is observed that when there is a high quantity of quality contact, it is possible to generalize stereotypes to members of the category as a whole, as posited by Brewer and Gaertner's (2001) contact model.

The limitations of this study include the use of a cross-sectional design, which prevents us from clarifying the causal direction of contact toward attitude change. Also, no data was available to ascertain the previous attitudes of voters toward the elite classes; such data would have allowed us to know the extent to which contact with elites affects intergroup attitudes. The tendency toward social polarization may have affected the data in both samples (Kingzette et al., 2021). However, the more frequent contact with elites among right-wing voters may have generated responses based on a more accurate understanding of this class's characteristics, making them feel that they better understand this class as a whole.

This research is the first to use the IAT to measure implicit attitudes toward the elite classes generated by

the demonizing discourse produced by extreme left-wing parties in Spain. Although these results are framed in the area of intergroup knowledge and are not intended to be generalizable to all voters of both tendencies with very different realities, this research presents useful elements to guide the identification of hate speech coming from the extreme left in Spain.

Acknowledgments

The authors would like to thank ESERP Business & Law School and its research observatory for providing their facilities to collect all the data necessary for this research.

Conflict of Interests

The authors declare no conflict of interests.

References

Abdullah, N. H., Hassan, I., Azura Tuan Zaki, T. S., Ahmad, M. F., Hassan, N. A., Mohd Zahari, A. S., Ismail, M. M., & Azmi, N. J. (2022). Examining the relationship between factors influencing political information seeking-behaviour through social media among youths in Malaysia. *Revista de Comunicación de la SEECI*, *55*, 1–15. https://doi.org/10.15198/seeci.2022.55.e746

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Alonso, A. (2021, March 21). ¿Es España el país más polarizado de Europa? [Is Spain the most polarized country in Europe?]. *El Independiente*. https://tinyurl.com/yhhzv63y

Alonso-Muñoz, L., & Casero-Ripollés, A. (2021). ¿Buscando al culpable? La estrategia discursiva en Twitter de los actores políticos populistas europeos en tiempos de crisis [Looking for the guilty? The discursive Twitter strategy of European populist political actors in times of crisis]. *Cultura, Lenguaje y Representación*, *26*, 29–45.

Amores, J. J., Blanco-Herrero, D., Sánchez-Holgado, P., & Frías-Vázquez, M. (2021). Detectando el odio ideológico en Twitter: Desarrollo y evaluación de un detector de discurso de odio por ideología política en tuits en español [Detecting ideological hate on Twitter: Development and evaluation of a detector of hate speech by political ideology in Spanish-language tweets]. *Cuadernos.info*, *2021*(49), 98–124.

Banks, A., Calvo, E., Karol, D., & Telhami, S. (2021). #PolarizedFeeds: Three experiments on polarization, framing, and social media. *The International Journal of Press/Politics*, *26*(3), 609–634.

Baptista, J. P., Correia, E., Gradim, A., & Piñeiro-Naval, V. (2021). Partidismo: ¿El verdadero aliado de las fake news? Un análisis comparativo del efecto sobre la creencia y la divulgación [Partisanship: The true ally of fake news? A comparative analysis of the effect on belief and disclosure]. *Revista Latina de Comunicación Social*, *79*, 23–47. https://doi.org/10.4185/RLCS-2021-1509

Barbeito Iglesias, R., & Iglesias Alonso, Á. (2021). Political emotions and digital political mobilization in the new populist parties: The cases of Podemos and Vox in Spain. *International Review of Sociology*, *31*(2), 246–267.

Barberá, P. (2015, September 3–6). *How social media reduces mass political polarization: Evidence from Germany, Spain, and the US* [Paper presentation]. 2015 APSA Conference, San Francisco, CA, United States.

Bartha, A., Boda, Z., & Szikra, D. (2020). When populist leaders govern: Conceptualizing populism in policy making. *Politics and Governance*, *8*(3), 71–81.

Ben-David, A., & Fernández, A. M. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, *10*, 1167–1193.

Berberena, T., & Wirzberger, M. (2021). Unveiling unconscious biases and stereotypes in students: The necessity of self-reflection in higher education. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*, 3488. https://escholarship.org/uc/item/6k97t362

Brandariz Portela, T. (2021). El debate electoral del 10-N de 2019 en España: Los candidatos políticos y RTVE [The 10-N 2019 election debate in Spain: Political candidates and RTVE]. *Revista de Ciencias de la Comunicación e Información*, *26*, 1–25. https://doi.org/10.35742/rcci.2021.26.e130

Brewer, M., & Gaertner, S. L. (2001). Toward reduction of prejudice: Intergroup contact and social categorization. In R. Brown & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 451–476). Blackwell Publishers.

Bridgewater, J. (2023). Leader change, time in office and the determinants of voter perceptions. *Parliamentary Affairs*, *76*(1), 146–161.

Brown, M., & Klein, C. (2020). Whose data? Which rights? Whose power? A policy discourse analysis of student privacy policy documents. *The Journal of Higher Education*, *91*(7), 1149–1178.

Brown, R., & Hewstone, H. (2005). An integrative theory of intergroup contact. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 255–343). Academic Press.

Chirco, P., & Buchanan, T. M. (2023). We the people. Who? The face of future American politics is shaped by perceived foreignness of candidates of color. *Analyses of Social Issues and Public Policy*. Advance online publication. https://doi.org/10.1111/asap.12341

Civila, S., Romero-Rodríguez, L. M., & Aguaded, I. (2021). Symbolic-discursive violence and new media: An epistemological perspective. *Catalan Journal of Communication & Cultural Studies*, *13*(1), 43–61.

Coll, V. (2021, July 13). Podemos no condena la represión en Cuba porque "no es una dictadura." *El Mundo*. https://tinyurl.com/26y2342y

Correcher Mira, J. (2021). La banalización del discurso del odio: Una expansión de los colectivos ¿vulnerables? [The trivialization of hate speech: An expansion of vulnerable groups?]. *InDret*, *2021*(2), 86–149. https://doi.org/10.31009/indret.2021.i2.04

Cox, W., & Devine, P. (in press). Changing implicit bias vs empowering people to address the personal dilemma of unintentional bias. In J. A. Krosnick, T. H. Stark, & A. L. Scott (Eds.), *The Cambridge handbook of implicit bias and racism*. Cambridge University Press.

Custodi, J. (2021). Nationalism and populism on the left: The case of Podemos. *Nations and Nationalism*, *27*(3), 705–720.

Dunmire, P. L. (2012). Political discourse analysis: Exploring the language of politics and the politics of language. *Language and Linguistics Compass*, *6*(11), 735–751.

Eroğlu Utku, D., & Köroğlu, N. Ö. (2020). Anti-immigration vs anti-EU: Political discourse analysis of Brexit decision of the UK. *Journal of Divinity Faculty of Hitit University*, *19*(1), 1–34.

Errejón, Í. (2021). *Con todo: De los años veloces al futuro* [All in all: From the fast years to the future]. Planeta.

Escobar Beltrán, S. A. (2016). Del odio al prejuicio: Reflexiones sobre la subjetividad y su prueba en los instrumentos penales antidiscriminación [From hatred to prejudice: Reflections on subjectivity and its proof in anti-discrimination criminal instruments]. *Estudios Socio-Jurídicos*, *18*(2), 175–202.

Europa Press. (2021, April 11). Iglesias sobre vídeo de Ayuso corriendo: "Te falta mucho Madrid, tía, podías ir a que te graben a barrios del Sur" [Iglesias on video of Ayuso running: "You're missing Madrid a lot, aunt, you could go to be filmed in southern neighborhoods"]. *Heraldo*. https://tinyurl.com/4ucpkmzr

Fanoulis, E., & Guerra, S. (2021). Aporias in Greece and Spain's left-wing inclusionary populism. *New Political Science*, *43*(3), 339–354.

Fernández Holgado, J. Á., Puentes-Rivera, I., & Fontenla-Pedreira, J. (2022). Relaciones públicas, comunicación política y narrativa audiovisual: Spot electoral del Partido Popular en las elecciones a la Asamblea de Madrid de 2021 [Public relations, political communication and audiovisual narrative: Electoral spot of the Partido Popular in the elections to the Assembly of Madrid in 2021]. *Revista de Ciencias de la Comunicación e Información*, *27*, 45–60. https://doi.org/10.35742/rcci.2022.27.e253

Fernández Vázquez, G. (2022). La resemantización de la idea de patria en Podemos: Un ejercicio inacabado [The resemantization of the idea of homeland in Podemos: An unfinished exercise]. *Más Poder Local*, *2022*(49), 60–77.

Fernández-Vázquez, P., & Somer-Topcu, Z. (2019). The informational role of party leader changes on voter perceptions of party positions. *British Journal of Political Science*, *49*(3), 977–996.

Fischer, A. M. (2020). The dark sides of social policy: From neoliberalism to resurgent right-wing populism. *Development and Change*, *51*(2), 371–397.

Gidron, N., Adams, J., & Horne, W. (2020). *American affective polarization in comparative perspective*. Cambridge University Press.

Gomes Barbosa, M. P. (2022). *El discurso contra la población LGTB en los medios sociales y su relación con los delitos de odio en España* [The discourse against the LGTB population in social media and its relation to hate crimes in Spain]. Manuscript in preparation. https://gredos.usal.es/bitstream/handle/10366/149995/Marcos%20Gomes%20Barbosa-rep.pdf?sequence=1&isAllowed=y

Gonzálvez Vallés, J. E. (2013). *Redes sociales y lo 2.0 y lo 3.0* [Social networks and the 2.0 and the 3.0]. Vision Libros.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

Huber, M. (2020). Steered by numbers: How quantification differentiates the reform of a German university. *Politics and Governance*, *8*(2), 26–35.

Iglesias, P. [@PabloIglesias]. (2022, May 9). Sobre los insultos de Sánchez Galán (Iberdrola) a los ciudadanos, las regañinas de nuestro gobierno y los hechos de Andrés Manuel López Obrador #LaBaseSánchezGalán [On Sánchez Galán's (Iberdrola) insults to citizens, our government's scoldings and Andrés Manuel López Obrador's deeds #LaBaseSánchezGalán] [Tweet]. Twitter. https://tinyurl.com/2p9jevxs

Ivaldi, G., Lanzone, M. E., & Woods, D. (2017). Varieties of populism across a left–right spectrum: The case of the Front National, the Northern League, Podemos and Five Star Movement. *Swiss Political Science Review*, *23*, 354–376. https://doi.org/10.1111/spsr.12278

Jansa, M., & Vilasero, M. (2015, January 31). Así hemos contado la Marcha del Cambio de Podemos en Madrid [This is how we have told the story of the March for Change of Podemos in Madrid]. *El Periódico*. https://tinyurl.com/3ttjntfn

Jaráiz, E. G., Martín, Á. C., López, M. P., & Castaños, S. P. (2021). El auge de la extrema derecha en España [The rise of the extreme right in Spain]. *Revista Española de Investigaciones Sociológicas*, *2021*(176), 161–163.

Kingzette, J., Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2021). How affective polarization undermines support for democratic norms. *Public Opinion Quarterly*, *85*(2), 663–677.

Larrondo-Ureta, A., Fernández, S.-P., & Morales-i-Gras, J. (2021). Desinformación, vacunas y Covid-19: Análisis de la infodemia y la conversación digital en Twit-

ter [Disinformation, vaccines, and Covid-19: Analysis of the infodemic and digital conversation on Twitter]. *Revista Latina de Comunicación Social*, *79*, 1–18. https://doi.org/10.4185/RLCS-2021-1504

López del Castillo Wilderbeek, F. L. (2021). El seguimiento sobre las fake news en medios institucionales durante el coronavirus en España [Monitoring of fake news in institutional media during the coronavirus in Spain]. *Vivat Academia: Revista de Comunicación*, *154*, 1–12. https://doi.org/10.15178/va.2021.154.e1253

Manetto, F. (2014, Ocotber 18). Iglesias apela a la mayoría para "ocupar la centralidad del tablero" [Iglesias appeals to the majority to "occupy the center of the board"]. *El País*. https://tinyurl.com/2jpxjukx

Manheim, J. B. (2004). *Biz-war and the out-of-power elite: The progressive-left attack on the corporation*. Psychology Press.

Marcos-Marne, H. (2021). A tale of populism? The determinants of voting for left-wing populist parties in Spain. *Political Studies*, *69*(4), 1053–1071.

Meleady, R. (2021). "Nudging" intergroup contact: Normative social influences on intergroup contact engagement. *Group Processes & Intergroup Relations*, *24*(7), 1180–1199.

Merco. (2022). *Ranking Merco companies Spain*. https://www.merco.info/es/ranking-merco-empresas

Moliner, M., Esteban, J. D., Sánchez, A. M., & Milbeau, A. J. (2012). *Diccionario de sinónimos y antónimos* [Dictionary of synonyms and antonyms]. Gredos.

Montero, A. I., Bullido, N. L., & Hervàs, D. A. (2022). Discursos de odio: Revisión sistemática de la producción científica y consideraciones educativas [Hate speech: Systematic review of scientific production and educational considerations]. *Revista Fuentes*, *24*(2), 222–233.

Montessori, N. M., & Morales-López, E. (2019). The articulation of "the people" in the discourse of Podemos. In J. Zienkowski & R. Breeze (Eds.), *Imagining the peoples of Europe: Populist discourses across the political spectrum* (pp. 123–147). Johns Benjamins Publishing Company.

Mudde, C., & Rovira Kaltwasser, C. (2013). Exclusionary vs. inclusionary populism: Comparing contemporary Europe and Latin America. *Government and Opposition*, *48*(2), 147–174. https://doi.org/10.1017/gov.2012.11

Mut Camacho, M., & Rueda Lozano, A. (2021). Las empresas ante la desinformación: La necesidad de un nuevo enfoque metodológico [Companies facing disinformation: The need for a new methodological approach]. *Vivat Academia: Revista de Comunicación*, *155*, 113–129. https://doi.org/10.15178/va.2022.155.e1327

Nasr, M. (2020). Voter perceptions of parties' left–right positions: The role of party strategies. *Electoral Studies*, *68*, Article 102239. https://doi.org/10.1016/j.electstud.2020.102239

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The implicit association test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). Psychology Press.

Olivas Osuna, J. J. (2021). Populismo en España: Fundamentos teóricos y relatos dominantes [Populism in Spain: Theoretical foundations and dominant narratives]. *Araucaria*, *23*(47), 371–401.

Pavlichenko, L. V. (2022). Polarization in media political discourse on the war in Ukraine: Critical discourse analysis. *Alfred Nobel University Journal of Philology*, *2*(24), 214–223. https://doi.org/10.32342/2523-4463-2022-2-24-18

Pérez-Escolar, M., & Noguera-Vivo, J. M. (2022). *Hate speech and polarization in participatory society*. Routledge.

Podemos València. (2014, May 14). *Intervención Pablo Iglesias inicio campaña en Valencia* [Intervention Pablo Iglesias campaign kickoff in Valencia] [Video]. YouTube. https://www.youtube.com/watch?v=a8RGcUpK6fE

Rachman, A., & Yunianti, S. (2017). Critical discourse analysis in Donald Trump presidential campaign to win American's heart. *Tell: Teaching of English Language and Literature Journal*, *5*(2), 8–17. http://dx.doi.org/10.30651/tell.v5i2.388

Rani, N., Das, P., & Bhardwaj, A. K. (2022). Rumor, misinformation among web: A contemporary review of rumor detection techniques during different web waves. *Concurrency and Computation: Practice and Experience*, *34*(1), Article e6479.

Rodríguez Ferrándiz, R., & Sorolla-Romero, T. (2021). El análisis de las imágenes en la era de las fake news [The analysis of images in the era of fake news]. *AdComunica*, *2021*(22), 23–26. https://doi.org/10.6035/2174-0992.2021.22.2

Rodríguez Terceño, J., Gonzálvez Vallés, J. E., & Caldevilla Domínguez, D. (2016). Narrativa digital en el entorno de la web 2.0 dentro del sector de la comunicación universitaria [Digital narrative in the web 2.0 environment within the university communication sector]. *F@ro: Revista Teórica del Departamento de Ciencias de la Comunicación*, *1*(23), 168–180.

Romero Rodríguez, L. M., Razo, A. L. V., & Abdul, B. C. (2021). Fake news de humor y sátira y actitudes hacia la política: Análisis comparativo del realismo percibido y los sentimientos de eficacia, alienación y cinismo en estudiantes de comunicación [Humor and satire fake news and attitudes toward politics: Comparative analysis of perceived realism and feelings of efficacy, alienation, and cynicism in communication students]. *OBETS: Revista de Ciencias Sociales*, *16*(2), 465–480.

Romero-Rodríguez, L. M., Tejedor, S., & Pabón Montealegre, M. V. (2021). Actitudes populistas y percepciones de la opinión pública y los medios de comu-

nicación: Estudio correlacional comparado entre España y Colombia [Populist attitudes and public opinion and media perceptions: A comparative correlational study between Spain and Colombia]. *Revista Latina de Comunicación Social*, *79*, 43–66. https://doi.org/10.4185/RLCS-2021-1507

Ross, L. (1981). The "intuitive scientist" formulation and its developmental implications. In J. H. Flavell & L. Ross (Eds.), *Social cognitive development: Frontiers and possible futures* (pp. 1–42). Cambridge University Press.

Ruiz Coll, M. A. (2022, July 19). Podemos pide prisión para los directivos de bancos y eléctricas que trasladen al cliente el nuevo impuesto [Podemos calls for imprisonment for bank and electricity managers who pass on the new tax to customers]. *El Español*. https://tinyurl.com/2p97t5hf

Schmid, U. K., Kümpel, A. S., & Rieger, D. (2022). How social media users perceive different forms of online hate speech: A qualitative multi-method study. *New Media & Society*. Advance online publication. https://doi.org/10.1177/14614448221091185

Serrano Villalobos, O., & Cuesta Díaz, V. (2017). Tecnoadicción al sexo en la población juvenil: Propuesta del diseño de una escala de detección temprana [Technosex addiction in the youth population: Proposal for the design of an early detection scale]. *Revista de Comunicación y Salud*, *71*, 119–149. https://doi.org/10.35669/revistadecomunicacionysalud.2017.7(1).119-149

Shokrollahi, O., Hashemi, N., & Dehghani, M. (2021). *Discourse analysis of Covid-19 in Persian Twitter social networks using graph mining and natural language processing*. ArXiv. https://doi.org/10.48550/arXiv.2109.00298

Stobart, L. (2014, November 5). Understanding Podemos (1/3): 15-M & counter-politics. *Left Flank*. https://left-flank.org/2014/11/05/explaining-podemos-1-15-m-counter-politics

Stoehrel, R. F. (2016). The regime's worst nightmare: The mobilization of citizen democracy. A study of Podemos' (aesthetic) populism and the production of affect in political discourse. *Cultural Studies*, *31*(4), 543–579. https://doi.org/10.1080/09502386.2016.1264004

Tierno, P. C. (2019). Populismo de izquierda en Europa: Una comparación entre Podemos y la Francia Insumisa [Left-wing populism in Europe: A comparison between Podemos and Unsubmissive France]. *Revista de Estudios Políticos*, *2019*(184), 41–65.

Tropp, L. R., White, F., Rucinski, C. L., & Tredoux, C. (2022). Intergroup contact and prejudice reduction: Prospects and challenges in changing youth attitudes. *Review of General Psychology*, *26*(3), 342–360.

Tuesta-Panduro, J. A. (2021). Las tecnologías de la información y comunicación, competencias investigativas y docencia universitaria: Revisión sistemática [Information and communication technologies, research skills and university teaching: A systematic review]. *Maestro Y Sociedad*, *18*(2), 440–456.

Turiel, E. (2010). Snap judgment? Not so fast: Thought, reasoning, and choice as psychological realities. *Human Development*, *53*, 105–109.

Urcola Eceiza, E., & Elezgarai Ibañez, I. (2021). La polarización y echo chambers en el debate de la #LeydeEutanasia en Twitter [Polarization and echo chambers in the #LeydeEutanasia debate on Twitter]. *Revista de Comunicación de la SEECI*, *2021*(54), 187–204. https://doi.org/10.15198/seeci.2021.54.e723

Verloo, M. (2018). Gender knowledge, and opposition to the feminist project: Extreme-right populist parties in the Netherlands. *Politics and Governance*, *6*(3), 20–30.

Wang, J. (2016). New political and communication agenda for political discourse analysis: Critical reflections on critical discourse analysis and political discourse analysis. *International Journal of Communication*, *10*, 2766–2784.

Xu, X., & Ballena, C. T. (2022). Political interviews with Donald Trump: Uncovering power and context through critical political discourse analysis. *Philippine Social Science Journal*, *5*(1), 9–24.

About the Authors

Juan Enrique Gonzálvez-Vallés holds a PhD in information sciences from the Complutense University of Madrid. He currently belongs to the Department of Theories and Analysis of Communication at the Complutense University of Madrid, where he lectures on research methods in advertising and public relations, psychology of communication, and hypermedia narrative. Previously, he belonged to the Department of Audiovisual Communication and Advertising at CEU San Pablo University.

José Daniel Barquero-Cabrero obtained five PhDs in the area of economic and social sciences from the International University of Catalonia, the Camilo José Cela University of Madrid, the Autonomous University of Coahuila de México, and the Universities of Malaga, Huelva, Cádiz and Seville (interuniversity program). For his contributions to the academic world, he has been awarded the title of doctor honoris causa by universities in America, Europe, and Asia. He is a professor at Lasalle University and UOC University.

Natalia Enseñat-Bibiloni is the director of marketing and communication at ESERP Business School. She obtained a university degree in business management and a diploma in business sciences. She also completed an official master's degree in marketing management and a postgraduate degree in high specialisation in management techniques at the University of Barcelona. She obtained her doctorate in legal and economic sciences with the qualification of excellent cum laude. She is a member of the European Higher Council of Doctors.

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 187–197
https://doi.org/10.17645/pag.v11i2.6374

Article

Hate Speech and Polarization Around the "Trans Law" in Spain

Patricia Sánchez-Holgado *, Carlos Arcila-Calderón, and Marcos Gomes-Barbosa

Department of Sociology and Communication, University of Salamanca, Spain

* Corresponding author (patriciasanc@usal.es)

Submitted: 30 October 2022 | Accepted: 2 January 2023 | Published: 17 May 2023

Abstract
The approval of the law for the real and effective equality of trans people and for the guarantee of the rights of LGTBI people (the Trans Law) has been one of the most controversial legislative actions in recent years; however, there is very little knowledge about the public's perception of the issue and how they express their opinions in the public sphere. Supporters of the law consider that the free determination of gender is essential to end discrimination against trans people, while critics express hate speech that can be a precursor to violent actions. This work aims to fill a gap, studying the relationship between the perceptions of a controversial and polarizing issue and their public expression through social media. The main objective is to analyze the public's perception, with special attention to age, gender, and political ideology, and compare it with the hate speech posted on Twitter. The methodology presents a survey of citizens and the use of computational methods to analyze Twitter messages with a machine-learning algorithm that classifies them as hate or not hate. The results indicate that the majority (80.1%) support the approval of the law, while those close to left and center political ideologies favor the law more than those who define themselves as right-wing. There are no significant differences according to age. Of the messages analyzed, 9.7% were classified as hate.

Keywords
hate speech; LGTBI; polarization; public perception; Spain; Trans Law; transphobia; Twitter

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

The proposal of the law for the real and effective equality of trans people and for the guarantee of the rights of lesbian, gay, trans, bisexual and intersex people (hereinafter LGTBI; popularly known as the Trans Law) is framed as one of the most controversial legislative actions in recent years due to the public debate it has generated. In Spanish society, progress has been made in the rights and protection of LGBTI people, but achieving real equality is a challenge. Discrimination based on sexual orientation and gender identity or expression continues to be a cause for concern, especially in the workplace, which can lead to verbal attacks, in the form of hate speech and physical violence, which can lead to hate crime (Córdoba, 2021).

With this in mind, we developed the following research questions:

RQ1: Does the public support the processing of the Trans Law, and if so, do the people who show their support for the law also support its main points?

RQ2: Does greater support shown for the proposed Trans Law result in less hate speech on social networks?

In addition, a second research question (RQ2) leads us to consider whether society's support for the processing of the Trans Law may be related to the presence of hate speech on social networks.

There is a wide range of opinions from the different sectors of Spanish society: Supporters of the Trans

Law consider that free gender self-determination is essential to end discrimination against trans people (Rodríguez-Ruiz & Mestre i Mestre, 2021), while critics express anti-trans rights opinions that may be precursors of hate speech and violent actions (Carratalá, 2020).

The main objective of this work is to analyze public opinion on the processing of the Trans Law proposal and its main statements, with special attention to age, gender, and political ideology, and compare it with the presence of hate speech messages regarding gender identity or expression on the social network Twitter. This work aims to fill a gap, studying the relationship between the perception of a controversial and polarizing issue, the approval of the Trans Law, and the public expression of said polarization through social networks.

2. Theoretical Background

The Trans Law is promoted by the Ministry of Equality and aims to protect the rights of LGTBI people and their families. The Council of Ministers approved the law on June 29, 2022, allowing it to go through the legislative process and be approved in December 2022. The proposed law aims to comply with the principle of formal equality, proclaimed in Article 14 of the Spanish Constitution (Cortes Generales, 1978), the principle of real and effective equality (Art. 9.2), and the consideration of the dignity of the person and the free development of the personality as the foundation of the political order and social peace (Art. 10.1).

The Trans Law proposes an establishment of a minimum system of protection against discrimination based on sexual orientation and identity, gender expression, or uniform sexual characteristics throughout the state territory, which protects equality laws; prohibits conversion therapies; allows access to assisted reproduction services for trans people with the ability to gestate; legislates for intersex human rights, avoiding medical-surgical interventions during the first year of life when their reproductive anatomy has characteristics of both sexes (except for health reasons); promotes the employment for trans women; and includes other important subjects for the protection of LGTBI people and their environment. The law also defines other elements aimed at improving the quality of life of LGTBI people, such as requiring that companies with more than 50 workers have action protocols for addressing harassment or violence against LGTBI people. It also includes measures to prevent homelessness among LGTBI people and measures on adoption and foster care. Concerning rural Spain, it includes the term "*sexilio*" (exile for reasons of gender) and indicates that LGTBI people from rural environments must be able to access the same resources under the same conditions as in big cities.

The public receives information about this type of legal procedure that depends to a large extent on the media, public opinion, and the conversation around the issues that make up the current agenda. Citizen participation is challenging in our societies (Naser et al., 2021). Knowing how citizens perceive the proposed law and its content is relevant to understanding the support that should be offered and how it can affect future decisions by government entities (Bayo Fernández et al., 2018). Some previous studies have made advances in analyzing public support for government actions, especially in the case of the health crisis (Chon & Park, 2021; Jørgensen et al., 2021). For this reason, we put forward a first hypothesis:

> H1: The higher the level of information declared about the Trans Law, the greater the support shown.

According to data from the Centro de Investigaciones Sociológicas (2021), 93.9% of Spaniards declare themselves to be heterosexual. However, in the younger age group, 18–24 years old, only 82.7% identify as heterosexual, while 16.4% identify with other sexual orientations. The Instituto de la Juventud (2020) also points in a similar direction: 16% of young people (15–29 years old) declare themselves to be non-heterosexual, and around 75% identify with one of the masculine or feminine poles, while 25% identify with non-binary positions. That is why it can be understood that young people are more likely to show support for the Trans Law, so we state the following hypothesis:

> H2: The younger the person, the more support they show for the proposed Trans Law.

In the first decade of the 21st century, Spain achieved substantial progress in LGBT rights. The legislation governing family formation, filiation, and marriage included LGBT families, which meant the visibility of non-heterosexual nuclear family structures and new models of parenthood (Imaz, 2017). Calvo and Trujillo (2011) reported that changes in the Civil Code allowed homosexual couples to marry, and full adoption rights were also granted in 2005. Platero-Méndez (2008) indicates that, historically, the center and the left-wing political ideologies have made a greater effort to legislate toward the rights of sexual minorities. The left-wing political groups have included gender identity issues in their politics and have worked to achieve social improvements. We can therefore propose a new hypothesis:

> H3: The stronger the leftist ideology that a person has, the greater their support for the proposal of the Trans Law.

By conducting a review of the main points of conflict and controversy in Spanish public opinion, we can check whether, compared to people who do not support the Trans Law, people who do express support for it will show agreement with the approval of these points individually.

Regardless of the important advances with Law 03/2007, the LGTBI+ organizations continued to point

out the limitations to trans people's rights, such as the previous professional diagnosis to the legal gender change and the limitation by age (only in 2019 were minors considered mature enough to change their legal gender). Therefore, the inclusion of gender self-determination in the law represents a substantial improvement in the situation of transgender people in Spain, and it is an old demand of LGTBI people (Esteve-Alguacil & Nonell i Rodríguez, 2021). Despite being one of the most controversial issues in the bill, 14 autonomous communities have already recognized free gender self-determination before the debate on the Trans Law. However, the results of a prior public consultation carried out by the Ministry of Equality between June 26 and July 15, 2020, in which 1,044 contributions from citizens were received, revealed that 58% of the contributions submitted rejected the principle of gender self-determination. That is why it is a point of conflict, and the following hypothesis can be stated:

> H4: Compared to people who do not support the Trans Law, those who support it will be more likely to agree with trans people legally changing their gender without the need for a medical diagnosis (self-determination).

Conservative sectors of Spanish politics have attacked gender self-determination. For example, the president of the extreme-right party VOX, Santiago Abascal, signed an opinion article published in the newspaper *El Mundo* declaring that if the Trans Law were approved, any citizen (rapist or abuser) would be able to choose their legal sex, which would be a threat to the lives of women and children (Abascal, 2021). On the other hand, there is an evident conflict between some radical feminist groups, who consider that free gender self-determination erases women, and trans groups, who consider that pathologization and the bureaucratic process is a stigma (Corral-Díaz, 2021); they tend to disagree with the trans-exclusionary radical feminism (TERF) groups because of their opposition to transsexual theses (Álvarez-Rodríguez, 2022). Within the framework of public discussions about the Trans Law, the use of the hashtag #ContraElBorradoDeLasMujeres has been reaffirmed as the position contrary to the autonomous identity of transsexual people (Ferré-Pavia & Zaldívar, 2022). Since it is another controversial point, we state the following hypothesis:

> H5: Compared to people who do not support the Trans Law, those who support it will agree that feminism, in general, should be inclusive toward trans women.

Some LGBTI+ organizations have also criticized the Trans Law, but in this case, despite recognizing that the law presents significant advances for LGBTI rights, such as gender self-determination, they have concerns that certain limitations have yet to be considered. According to prominent organizations such as FELGTBI+, Chrysallis, and Fundación Triángulo (FELGTBI+ et al., 2021), one of these claims would be a third box on the national identity document so that non-binary people can mark it. Similar policies have been adopted in the Catalonia Autonomous Community. In 2019, the Catalan Department of Labor, Social Affairs, and Families announced that official documents in Catalonia would include "non-binary" as a third option. In 2022, Catalonia incorporated the reference to non-binarism, referring to "sense" and family diversity in administrative documents. Based on this point, we establish the following hypothesis:

> H6: Compared to people who do not support Trans Law, those who support it will agree that non-binary people's rights should be included.

The public discussion about the legislation on sexual orientation and gender identity generates such polarized opinions that public demonstrations, in support and rejection of the law, are observed during legislative processing, which open new spaces for extreme discourses and, therefore, for harmful content targetting this historically vulnerable group. Sponholz and Christofoletti (2019) point out that the media system, with strong commercial ties, especially television and the internet, can empower public figures who give speeches that contain violence against minority groups to guarantee their presence and media prominence. It is of great importance on the internet since homogeneous discussion spaces, echo chambers, and intentional misinformation contribute to creating a toxic atmosphere (Cinelli et al., 2021). Torregrosa et al. (2020) conclude that extremist political groups use violent discourse to maintain their relevance in the networks and amplify their opinions following an echo chamber effect.

In addition, it is important to highlight that the Trans Law was proposed in a context of concern about the increase in hate crimes against the LGTBI community. According to the report on the evolution of hate crime in Spain 2021, published by the Ministry of the Interior (López Gutiérrez, Sánchez Jiménez, et al., 2021), the category of sexual orientation and gender identity is the subject of the second largest number of registered hate crimes (466 cases), 25.86% of the total cases. It represents an increase of 68.23% compared to the previous year (277 cases in 2020). On the other hand, the same report indicates that 87.10% of the victims of hate crimes based on sexual orientation/gender identity did not file a complaint, meaning that the real number of hate crimes against the LGBTI population may even be higher than those registered and that the volume of under-reporting is high. On the other hand, data from the European Union Agency for Fundamental Rights (2020) show that transsexual and intersex people suffer the most violence, intolerance, and discrimination in Spain. Forty-one percent of LGTBI people suffered some type of harassment

due to their condition, and 8% of those had suffered some physical or sexual assault in the last five years. However, harassment affects 51% of transgender people and 52% of intersex people in Spain, while physical or sexual assaults affect 15% and 19%, respectively.

Hate speech is any form of expression that spreads, incites, promotes, or justifies any form of hate based on intolerance (Arcila-Calderón et al., 2021). It can be transmitted through many forms of expression, language, and media. Furthermore, in the context of public discussion, hate speech can provoke, attract, and empower people who oppose socially rejected individuals. Beyond the dissemination of discriminatory perspectives, hate speech can also evolve into behaviors that give rise to other types of violent acts (Allport, 1954) and political radicalization (Abuín-Vences et al., 2022; Bilewicz & Soral, 2020; Soral et al., 2018).

We are currently immersed in a context where the creation and dissemination of content in digital media and social networks are growing ever faster. Under this circumstance, hate speech can trigger hate crimes (Arcila-Calderón et al., 2020). Müller and Schwarz (2020) suggest that there is a significant relationship between online hate speech and offline violent attacks. Consistent with this thesis, Regehr (2022) concludes that the internet's echo chamber can contribute to creating a loop that transforms digital hatred into physical violence. In addition, there is evidence that if a group receives great support from society, the level of hate speech online is inversely proportional (Arcila-Calderón et al., 2022). For all these reasons, we pose RQ2.

3. Methods

The methodology presents a public opinion survey among citizens, which will serve to answer RQ1. The methodology uses computational methods and content analysis to analyze Twitter messages with an automatic learning algorithm that classifies them as hate or non-hate, and that will answer RQ2.

3.1. Public Opinion Survey

A telephone survey was carried out on November 12, 2021. An anonymous questionnaire survey was carried out by a trained team and applied to Spanish citizens between 16 and 89 years old, according to the seventh article of the organic law on personal data protection and guarantee of digital rights (Jefatura del Estado, 2018). People under 16 or who did not live in Spanish territory were not considered for the interview. The data was collected from a sample of participants using Diet Random Diary (a prefix 6xx and adding five random numbers). The total sample was n = 181. Regarding the gender of the participants, 52.5% were women, and 45.9% were men. The most represented autonomous communities in the study case are Andalusia, with 21% of the respondents; Madrid, with 18.8%; and Catalonia, with 17.7%.

The research is not representative. The total sample was n = 181. A power analysis indicated that the sample size was adequate. Specifically, the a priori analysis using G*Power 3.1.9.7 (Faul et al., 2009) determined that a medium effect size of 0.3 (Cohen, 1988) can be detected with at least 67 individuals (assuming $1 - B = 0.80$ and alpha level $\alpha = 0.05$), which means that our sample was satisfactory. Regarding the gender of the participants, 52.5% were women and 45.9% were men.

The survey was divided into two sections. The first included the following sociodemographic categories:

- Gender (*male*, *female*, *non-binary*);
- Age (in years);
- Region of residence in Spain;
- Political ideology (measured on a scale from 1 = *extreme left* to 10 = *extreme right*);
- Level of information declared about the Trans Law and its content (measured on a scale from 1 = *little information* to 5 = *much information*).

In the second section, the following variables and measures were used:

- Support for the Trans Law: Do you agree with the approval of the law for the real and effective equality of trans people and for the guarantee of the rights of LGTBI people, known as the Trans Law? *Yes*, *No*, *I prefer not to answer*.
- Support for transgender people to legally change their gender without the need for a medical diagnosis: Do you agree that transgender people can legally change their gender without the need for a medical diagnosis (self-determination)? *Yes*, *No*, *I prefer not to answer*.
- Feminism should include trans women: Do you agree that feminism, in general, includes trans women? *Yes*, *No*, *I prefer not to answer*.
- Support for the incorporation of non-binary people into the Trans Law: Do you agree that non-binary people are included in the Trans Law? *Yes*, *No*, *I prefer not to answer*.

The responses obtained were anonymized. Data were analyzed using the IBM SPSS program, performing an exploratory and descriptive analysis, including cross tables and correlation for independent variables. A hypothesis contrast of relationships between two variables was performed, using chi-square tests.

3.2. Computational Methods and Content Analysis

In the second stage, hate speech detection was carried out on Twitter. The work was divided into two main phases, one to create a training corpus and another to generate predictive models. The creation of a specific corpus for the study allows us to overcome the limitations of previously developed prototypes that

used dictionaries or general databases (Arcila-Calderón et al., 2020).

To create the training corpus, data collection was carried out between July 1 and November 30, 2021, dates selected as they were closest to the publication of the Trans Law. The download of the messages from the social network Twitter was carried out in the integrated environment Pycharm, connected to the application programming interface. The downloaded messages were filtered by the Spanish language and the keywords: #NoLeyTrans, #NoALaNiñezTrans, #Transfobia, #Homofobia, #LaLeyTransEsMisoginia, #NOalaLeyTrans, #TERF, #Transodio, #20Nov, #LeyTransAlCongreso, #LeyTrans, #LGTBI, #Feminismo. The database was cleaned, and all tweets that used the keywords in another context or language, those that contained incomplete messages, or did not directly refer to the topic were manually deleted. Although a large number of messages were downloaded, the final sample consisted of 5,000 tweets, which were then manually sorted by a team of trained coders.

The definition of hate speech used for coding was based on the three levels of cyberhate included in the taxonomy proposed by Miró-Llinares (2016) and included all types of hate based on gender and sexual orientation that could constitute a crime. The description and classification procedures for hate speech developed by Arcila-Calderón et al. (2021) were followed.

All tweets were manually classified into subsamples of 500 tweets by two human coders, indicating whether they contained *hate* (1) or *no hate* (0) in a binary way and discarding all those that did not meet the established requirements. To check the reliability of the intercoders, only the classifications with total agreement were accepted; the rest were discarded.

From the initial sample of 5,000 tweets, 3,756 (75.12%) were validated, while 1,244 messages (24.88%) did not meet the validation parameter and were discarded. Of the validated messages used to create the training corpus, 3,394 tweets were classified as non-hate (90.36%), and 362 were classified as hate (9.64%; Table 1).

The generated corpus was then used to train the predictive model with machine learning algorithms based on the prototype for the automatic detection of hate speech on Twitter in Spanish specifically motivated by gender and sexual orientation from Arcila-Calderón et al. (2021). The machine learning algorithm was used to detect hate speech on Twitter based on a stance for or against trans laws. Examples were provided to the model (70% training, 30% test), and eight predictive models were validated; six of them were generated by using surface learning algorithms (original Naïve Bayes, Naïve Bayes for multinomial models, Naïve Bayes for Bernoulli's multivariate models, logistic regression, linear classifiers with stochastic gradient descent training, and support vector machines), another model was generated from feedback from previous models, and a final model was generated using deep learning. For the evaluation of the predictive model, evaluation metrics for supervised machine learning were applied. This validation returned acceptable quality evaluation metrics for the hate class, with the following results: accuracy = 0.8510, precision = 0.7124, recall = 0.6979, and F1 score = 0.7051. To provide a metric calculated from precision and recall, we use the ROC-AUC score = 0.8688, which defines the threshold classification model function (0.70) and a loss of 0.4066.

4. Results

4.1. Public Opinion Survey

In this section, we will address the first research question (RQ1), which asked whether the public support the processing of the Trans Law and, if so, whether its supporters also support the Law's main points. In addition, in a second research question (RQ2), we will review the level of social support for the Trans Law and address all the established hypotheses.

Taking the total sample, the results indicate that 80.1% of those surveyed agreed with the approval of the Trans Law (of which 42.1% were women, 55.9% were men, and 2.1% did not declare their gender), 14.4% disagreed, and 5.5% preferred not to declare. As for the level of support for the issues related to the law that have been the subject of discussion, we can see that it is lower than the general support for the law, but they are mostly supported by more than half of those surveyed: 61.9% agreed with gender self-determination, 58% agreed with the inclusion of non-binary people in the law, and 67.4% agreed with trans-inclusive feminism (Table 2).

Regarding the level of information, on a scale from 1 (*little information*) to 5 (*much information*), 48.6% of the participants stated that they had "little information" on the subject; 19.3%, "neither much nor very little information"; 14.4%, "some information"; 11.6%, "enough information"; and 6.1%, "a lot of information."

Table 1. Sample of tweets.

Total sample of tweets	Discarded	Validated sample	
5,000	1,244	3,756	
		No hate	*Hate*
		3,394	362

Table 2. Public opinion survey general results.

Question	Yes	No	Did not answer
Do you agree with the approval of the law for the real and effective equality of trans people and for the guarantee of the rights of LGTBI people, known as the "Trans Law''?	80.1%	14.4%	5.5%
Do you agree that transgender people should be able to make a legal gender change without the need for a medical diagnosis (self-determination)?	61.9%	30.9%	7.2%
Do you agree that feminism, in general, should include trans women?	67.4%	13.3%	19.3%
Do you agree that non-binary people should be included in Trans Law?	58%	19.9%	22.1%

The mean is 2.12, which shows a low level of information about the law. We can verify the level of support for the law in each group of respondents with different levels of information. Regarding the participants who stated that they had "little information" on the subject, 72.7% agreed with the approval of the law and 17% disagreed. As for those surveyed who claimed to have "neither much nor little information," 85.7% agreed with the law's approval, while 11.4% disagreed. Considering the respondents who said they had "some information," 92.3% agreed with the approval of the law and 7.7% disagreed. Regarding the respondents who stated that they had "enough information," 90.5% of them agreed and 9.5% disagreed. Of the respondents who stated that they had "a lot of information," 72.7% agreed and 27.3% disagreed. Responding to H1, the higher the level of information declared about the Trans Law, the greater the support shown for the law; there is no statistically significant relationship between the support for the Trans Law and the level of information declared about the law and its content ($\chi^2[1, N = 171] = 4.095$; $p > 0.05$).

Considering age, the mean of those surveyed was 54 years ($M = 54.01$, $SD = 17.44$). The age variable was recoded as an ordinal variable in order to make group comparisons, grouping the observations into three age groups made up of a group of young people and young adults from 16 to 46 years old (30.4% of the participants); another group of adults, from 47 to 60 years old (30.4% of the participants); and a third group made up of older adults, from 60 to 89 years old (39.2% of the participants). The group made up of young people and young adults (16 to 46 years old) had the least favorable opinion: The results show that 74.5% of this group agreed with the approval of the Trans Law, 18.2% disagreed, and 7.3% preferred not to answer. In the next group (47 to 60 years old), 81.8% of adults agreed, 12.7% disagreed, and 5.5% preferred not to answer. The group of older people had the most favorable opinion: 83.1% agreed, 12.7% disagreed, and 4.2% preferred not to answer. This data shows that the youngest age group is the one that disagreed the most with the approval of the law. There is no statistically significant relationship between support for the approval of the Trans Law and age ($\chi^2[1, N = 171] = 56.374$; $p > 0.05$). Therefore, H2, stating that the younger the person, the greater the support they show for the Trans Law proposal, can be rejected. In addition, regarding other issues (legal recognition of gender by self-determination, the inclusion of non-binary people in the Trans Law, and the inclusion of trans women in feminism), we observed that these groups of different ages also behaved similarly. The differences observed when comparing the opinions expressed by the three age groups regarding the approval of the law do not present statistically significant results.

Paying attention now to the declared political ideology, the average of those surveyed is 4.54 ($SD = 2.10$), which implies a slight tendency to the left. The political ideology variable was recoded as an ordinal variable, in order to make group comparisons, extracting three groups by ideology with the following distributions: 38.1% left (from 1 to 4), 35.4% from the center (5 and 6), and 12.2% from the right (from 7 to 10). However, a considerable group of people (14.4%) did not answer the question. The leftist group had the most favorable opinion towards the Trans Law. Of the left-wing participants, 88.4% agreed with the approval of the law, 8.7% disagreed, and 2.9% preferred not to answer. Of the centrist participants, 79.7% agreed, 14.1% disagreed, and 6.3% preferred not to answer. The group least favorable to the Trans Law was the right: 63.6% agreed, 31.8% disagreed, and 4.5% preferred not to answer. This trend extends to the rest of the themes (see Table 3). To respond to H3, which stated that the more of a leftist ideology an individual has, the greater their support for the Trans Law proposal, we observed a statistically significant relationship ($p < 0.05$) between support for the Trans Law and political ideology. The relationship is weakly positive ($r = 0.171$); that is, the greater their support for the Trans Law, the greater the probability that an individual will identify as left-wing ($\chi^2[1, N = 148] = 20{,}100$; $p < 0.05$). Therefore, we can accept our second hypothesis.

We next review H4, which stated that compared to people who do not support the Trans Law, people who express their support for the Trans Law will be more likely to agree with trans people legally changing their gender without the need for a doctor's diagnosis (self-determination). Although self-determination has generated much controversy and has been criticized by civil society groups, 61.9% of all respondents agreed with transgender people legally changing their gender without the need for a medical diagnosis. Considering only the group that supports the Trans Law, 71% supported

Table 3. Support to the Trans Law topics (divided by ideology groups).

Subject	Left-wing (support in %)	Centrist (support in %)	Right-wing (support in %)
Approval of the Trans Law	88.4%	79.7%	63.6%
Gender self-determination (no medical requirements)	73.9%	62.5%	45.5%
Inclusion of non-binary people in the Trans Law	63.8%	65.6%	36.4%
Inclusion of trans women in feminism	81.3%	68.8%	45.5%

self-determination, 23.4% did not support it, and 5.5% did not answer. There is a statistically significant relationship ($p < 0.01$) between support for the Trans Law and support for gender change without a medical diagnosis. The relationship is positive on average ($r = 0.433$); that is, the greater the support for the Trans Law, the greater the probability of support for gender change without a medical diagnosis ($\chi^2[1, N = 163] = 30{,}600$; $p < 0.01$) which allows us to accept this hypothesis.

Most of the people who supported the law also agreed with the inclusion of trans women by feminism at levels even higher than the general population: 67.4% of all the participants said they agreed. Considering only the group that supported the Trans Law, 76.6% supported the inclusion of trans women within feminism, compared to 7.6% who disagreed, and 15.9% who preferred not to answer. Responding to what was stated in H5, that, compared to people who do not support the Trans Law, those who support it will agree that feminism, in general, should incorporate and be inclusive of trans women, there is a statistically significant relationship ($p < 0.01$) between support for the Trans Law and support for feminism to be inclusive of trans women. The relationship is positive on average ($r = 0.500$); that is, the greater the support for the Trans Law, the greater the probability of support for the inclusion of trans women by feminism ($\chi^2[1, N = 141] = 35{,}301$; $p < 0.01$).

To revise our last hypothesis (H6, compared to people who do not support Trans Law, people who express support for Trans Law will show agreement with the incorporation of non-binary people's rights in the law), considering only the group that supports the law, 65.5% also agreed with the inclusion of non-binary people in this legislation, compared to 15.9% who disagreed and 18.6% who did not answer. There is a statistically significant relationship ($p < 0.01$) between support for the Trans Law and support for the inclusion of non-binary people in the law. The relationship is weakly positive ($r = 0.366$); that is, the higher the support for the Trans Law, the greater the probability of support for non-binary people to be included in the law ($\chi^2(1, N = 136) = 18.186$; $p < 0.01$).

4.2. Computational Methods and Content Analysis

Finally, to answer RQ2, on whether greater support shown for the proposed Trans Law results in less hate speech on social networks, we analyzed the messages collected on Twitter. Of the messages obtained in the first process, 75.12% met the validation criteria. Considering the results obtained, added to the review of the qualitative analysis, we observe that most of the tweets collected in the sample have a favorable tone regarding the recently approved law for real and effective equality of trans people and for the guarantee of rights of LGBTI people. Of the validated messages, 90.36% were identified as not hateful and 9.64% as hateful. The percentage of hate messages found in the sample is still high. If we compare it with the general support for the Trans Law that declares citizenship, we can see that there is still work to be done. Although the level of support found is high, the presence of hate speech on the social network Twitter continues to be important. It is convenient to continue preparing new analyses and to analyze a direct relationship between the variables of support and the presence of hate speech.

However, we have identified that of the hateful messages, the majority were directed at trans women, where they were classified as "disguised men" or accused of being dangerous to cisgender women or even of wishing to mutilate and give hormones to underage children. Below are some examples translated into English as accurately as possible:

> Men are NOT women, not even disguised as cabaret dancers nor wearing women underwear….It is imposible that men feel like women.

> About the minors whose hormones and mutilation you are going to legalize with your abject law.

> Only women can gestate. Woman is the human female. Sex is immutable.

> Boys and girls: Eating chocolate and chopped pork sandwich—NO, BAD, INSANE. Boys and girls: Hormones and mutilations from the age of 10 to calm your parents' #homophobia and #lesbophobia—VERY GOOD, HEALTHY. Alberto Garzón and Irene Montero are a plague for #Childhood.

5. Conclusions, Limitations, and Future Studies

Comparing the data obtained by the public opinion survey and the classification of Twitter messages, both

methodologies pointed to a high level of support for the Trans Law; 80% of the respondents agreed with the law's approval and 90% of the messages collected were classified as non-hate. These numbers are consistent with other data regarding the Spanish population's perception of the LGBTI community in general. According to the special report *Discrimination in the European Union* (Eurobarometer, 2019), Spain is considered the third most tolerant European country towards LGBT rights: 91% of its population agrees that homosexual or bisexual people should have the same rights as heterosexual people. The high acceptance rate places the country in third place at the European level, behind Sweden (98%) and the Netherlands (97%).

In general, data from Spain are more LGBTI-friendly than the rest of Europe. Also, according to Eurobarometer (2019), we can see that the Spanish population supports other topics that dialogue with the Trans Law. Such as that, school lessons and material should include information about diversity in terms of sexual orientation (84%), being transgender (81%), and being intersex (80%). In addition, most of the Spanish population is favorable to non-binary people's rights: 63% of Spanish people believe that official documents should have a third option besides male and female. This data is close to that obtained by our public opinion survey, in which 58% of the participants stated that non-binary people should be included in the Trans Law. The rights of the non-binary population are supported by around 60% of the sample.

Although the Spanish population seems tolerant, our study and official data evidence that LGBT phobia is still a problem. Regarding crimes committed against the LGBTI population through the internet and social networks, although our data analysis indicates that 9.7% of the messages analyzed were classified as hate, official data suggests that this subject has increased significantly. In 2019, there were 17 cases of cybercrime motivated by a victim's sexual orientation or gender identity; by 2020, that number rose to 32 cases; by 2021, it had risen to 60 cases (López Gutiérrez, Sánchez Jiménez, et al., 2021), an 87.50% increase. Hate speech on social networks has affected a large part of the Spanish population, over half of whom have received offensive, discriminatory messages or threats on social networks or the internet in the last five years, and almost 60% perceive that this type of crime has increased (López Gutiérrez, Fernández Villazala, et al., 2021).

Our work also points to the need to discuss subjects related to the law in addition to its approval. For example, gender self-determination, pointed out as a highly relevant issue by LGTBI+ organizations, appears in many of the hate comments we studied. Notably, the level of support for the issues raised (gender self-determination, trans-inclusive feminism, and the rights of non-binaries) is lower than the declared support for the law itself. Lower levels of support for specific issues may reflect a lack of information regarding the law. More than half of the responses (67.9%) stated having "little information" or "neither too much nor too little information" about the bill, which may mean that although the subject is on the media agenda, it may be being discussed in a very superficial manner, without delving into the substantive points.

This study had some limitations that are important to highlight. The use of social media in scientific research has certain obstacles, including the type of sampling used. In this study, considering the composition of the social network Twitter, certain sociodemographic groups are not represented. So, using this methodology, we can identify hateful speech, but we cannot study the characteristics of its producers nor compare them to the respondents to identify how the recoded social groups behave on social networks. Meanwhile, the fact that Spain's equality minister Irene Montero (member of the left-wing political party Podemos) was mentioned in some hateful messages steps up the historical trend also verified by our survey results that sexual minorities' rights are traditionally associated with the left and center political ideologies. Considering the difficulty of getting the answers and the high average age, we question the telephone interview as a tool for this specific topic. It is pertinent to use other methodologies to continue exploring public opinion perception, such as a virtual questionnaire, which would allow access to a greater number of people from all autonomous communities and all age ranges.

Finally, the proportion of hateful tweets detected is less than those belonging to the regular class—in our study case, less than 10% of the messages analyzed were hateful content—but its potential harm should not be underestimated. Having compared this data with the official data, we found some new possible hypotheses that deserve further investigation in future research. The first is that this percentage can grow if this methodology is replicated in a year since online hate crimes against the LGBTI population are known to be a growing trend. The second is that the public perceives that hateful content on social networks is even higher. That is why it is interesting to develop new qualitative and quantitative methods to effectively compare the perceptions of the population with the official data and the debate carried out on social networks.

Acknowledgments

The authors express their gratitude to the staff of the master's program in Audiovisual Communication: Research and Innovation of the University of Salamanca for their support in the resources and tools for this research. The authors would like to express their thanks for all the help in the development of the fieldwork to Alejandra Garboza, Alejandro García Ulloa, Diana Mediavilla, Frescia Hidalgo, Germán Enríquez, Isabel Somoza, Jenifer Abad, Jingzi He, Julieta Collado, Linshu Sheng, Lorena Arévalo, Lu Lin, María Rodríguez,

Sergio Sevillano, Mingshu Fang, Suzanne Marion, William González, Williams Antonio González, Weixing Zhu, Xingchen Liu, Xinnan Zhang, Xinxin Kong, and Zhou Hengrui. They also thank Cristina Calvo for her help in the coordination of the survey fieldwork.

Conflict of Interests

The authors declare no conflict of interests.

References

Abascal, S. (2021, July 2). La Ley Trans, una amenaza para mujeres y niños [The Trans Law, a threat to women and children]. *El Mundo*. https://www.elmundo.es/espana/2021/07/01/60ddf552e4d4d8b6788b458a.html

Abuín-Vences, N., Cuesta-Cambra, U., Niño-González, J. I., & Bengochea-González, C. (2022). Análisis del discurso de odio en función de la ideología: Efectos emocionales y cognitivos [Hate speech analysis as a function of ideology: Emotional and cognitive effects]. *Comunicar*, *71*, 37–48. https://doi.org/10.3916/C71-2022-03

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Álvarez-Rodríguez, I. (2022). Apuntes constitucionales sobre la futura Ley Trans [Constitutional notes on the Trans Draft Law]. *Revista General de Derecho Constitucional*, *36*. https://www.iustel.com/v2/revistas/detalle_revista.asp?id_noticia=424778

Arcila-Calderón, C., Amores, J. J., Sánchez-Holgado, P., & Blanco-Herrero, D. (2021). Using shallow and deep learning to automatically detect hate motivated by gender and sexual orientation on Twitter in Spanish. *Multimodal Technologies and Interaction*, *5*(10), 63–75. https://doi.org/10.3390/mti5100063

Arcila-Calderón, C., Blanco-Herrero, D., & Valdez Apolo, M. B. (2020). Rejection and hate speech in Twitter: Content analysis of tweets about migrants and refugees in Spanish. *Revista Española de Investigaciones Sociológicas*, *172*, 21–40. https://doi.org/10.5477/cis/reis.172.21

Arcila-Calderón, C., Sánchez-Holgado, P., Quintana-Moreno, C., Amores, J. J., & Blanco-Herrero, D. (2022). Hate speech and social acceptance of migrants in Europe: Analysis of tweets with geolocation. *Comunicar*, *71*, 21–35. https://doi.org/10.3916/C71-2022-02

Bayo Fernández, I., Mecha, R., & Milán, M. (2018). *La comunidad científica ante los medios: Guía de acción para la divulgación de la ciencia* [The scientific community before the media: Action guide for the dissemination of science]. Complutense University of Madrid. https://www.ucm.es/data/cont/docs/1334-2018-05-04-Gu%C3%ADa%20de%20actuaci%C3%B3n%20WEB.pdf

Bilewicz, M., & Soral, W. (2020). Hate speech epidemic: The dynamic effects of derogatory language on intergroup relations and political radicalization. *Political Psychology*, *41*(S1), 3–33. https://doi.org/10.1111/pops.12670

Calvo, K., & Trujillo, G. (2011). Fighting for love rights: Claims and strategies of the LGBT movement in Spain. *Sexualities*, *14*(5), 562–579. https://doi.org/10.1177/1363460711415330

Carratalá, A. (2020). Audiencias críticas en Twitter frente a coberturas tránsfobas: La identidad de género como nuevo derecho y su tratamiento periodístico [Critical audiences on Twitter against transphobic coverage of gender identity as a new right and its journalistic treatment]. In J. M. Gómez y Méndez, M. C. Turón Padial, & M. Cartes-Barroso (Eds.), *Más sobre periodismo y derechos humanos emergentes* [Journalism and emerging human rights] (pp. 64–78). Equipo de Investigación de Análisis y Técnica de la Información, Universidad de Sevilla.

Centro de Investigaciones Sociológicas. (2021). *Encuesta sobre relaciones sociales y afectivas en tiempos de pandemia de la Covid-19* [Survey on social and affective relations in times of pandemic of Covid-19]. https://datos.cis.es/pdf/Es3325sdMT_A.pdf

Chon, M.-G., & Park, H. (2021). Predicting public support for government actions in a public health crisis: Testing fear, organization-public relationship, and behavioral intention in the framework of the situational theory of problem solving. *Health Communication*, *36*(4), 476–486. https://doi.org/10.1080/10410236.2019.1700439

Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, *118*(9). https://doi.org/10.1073/pnas.2023301118

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Córdoba, C. (2021). La situación actual del Colectivo LGTBI en España: Un análisis legislativo de los derechos reconocidos y la protección de víctimas de discriminación por orientación sexual y/o identidad o expresión de género [The current situation of the LGTBI Collective in Spain: A legislative analysis of the recognized rights and protection of victims of discrimination based on sexual orientation and/or gender identity or expression]. *EHQUIDAD Revista Internacional De Políticas de Bienestar y Trabajo Social*, *2021*(16), 141–164. https://doi.org/10.15257/ehquidad.2021.0017

Corral-Díaz, J. (2021). *Análisis del conflicto en torno a la Ley Trans en la red social Twitter: Actores, jerarquías y posiciones* [Analysis of the conflict around Trans Law in the social network Twitter: Actors, hierarchies, and positions] [Bachelor's dissertation, University of A Coruña]. UDC Repository. https://ruc.udc.es/dspace/handle/2183/28929

Cortes Generales. (1978). Constitución Española

[Spanish constituiton]. *Boletín Oficial del Estado*, *1978*(311). https://www.boe.es/buscar/act.php?id=BOE-A-1978-31229

Esteve-Alguacil, L., & Nonell i Rodríguez, A. (2021). Analysis of the draft act for the real and effective equality of trans people and the guarantee of the rights of LGTBI people. *InDret*, *3*, 267–290.

Eurobarometer. (2019). *Special Eurobarometer 493—Discrimination in the European Union: The social acceptance of LGBTI people in the EU*. European Commission.

European Union Agency for Fundamental Rights. (2020). *A long way to go for LGBTI equality*. https://fra.europa.eu/en/publication/2020/eu-lgbti-survey-results

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

FELGTBI+, Chrysallis, & Fundación Triángulo. (2021). *FELGTB, Triángulo y Chysallis presentan propuestas para que las leyes trans y LGTBI garanticen todos los derechos* [FELGTB, Triángulo, and Chysallis present proposals so that trans and LGTBI laws guarantee all rights]. FELGTBI+. https://felgtb.org/blog/2021/02/11/felgtb-triangulo-y-chysallis-presentan-propuestas-para-que-las-leyes-trans-y-lgtbi-garanticen-todos-los-derechos

Ferré-Pavia, C., & Zaldívar, G. (2022). El feminismo trans excluyente en Twitter: Un monólogo sesgado en #ContraElBorradoDeLasMujeres [Exclusionary trans feminism on Twitter: A biased monologue in #ContraElBorradoDeLasMujeres]. *ICONO 14. Revista Científica de Comunicación y Tecnologías Emergentes*, *20*(2). https://doi.org/10.7195/ri14.v20i2.1865

Imaz, E. (2017). Same-sex parenting, assisted reproduction and gender asymmetry: Reflecting on the differential effects of legislation on gay and lesbian family formation in Spain. *Reproductive Biomedicine & Society Online*, *4*, 5–12. https://doi.org/10.1016/j.rbms.2017.01.002

Instituto de la Juventud. (2020). *Informe juventud en España 20202* [Youth in Spain report 2020]. http://www.injuve.es/sites/default/files/adjuntos/2021/03/informe_juventud_espana_2020.pdf

Jefatura del Estado. (2018). Ley Orgánica 3/2018, de 5 de diciembre, de protección de datos personales y garantía de los derechos digitales [Organic Law 3/2018, of December 5, on personal data protection and guarantee of digital rights]. *Boletín Oficial del Estado*, *2018*(294), 119788–119857. https://www.boe.es/eli/es/lo/2018/12/05/3

Jørgensen, F., Bor, A., Lindholt, M. F., & Petersen, M. B. (2021). Public support for government responses against Covid-19: Assessing levels and predictors in eight Western democracies during 2020. *West European Politics*, *44*(5/6), 1129–1158. https://doi.org/10.1080/01402382.2021.1925821

López Gutiérrez, J., Fernández Villazala, T., Cortinas, C. J. M., San Abelardo Anta, M. Y., Esteban, J. G., Sánchez Jiménez, F., Herrera Sánchez, D., Martínez Moreno, F., García, M. R., Pérez, V. G., Santiago Orozco, A. M., & Gómez Martín, M. A. (2021). *Informe de la encuesta sobre delitos de odio* [Hate crime survey report in Spain 2021]. Ministerio del Interior. https://www.interior.gob.es/opencms/pdf/servicios-al-ciudadano/delitos-de-odio/estadisticas/Informe_encuesta_victimas_delitos-de-odio_version_definitiva.pdf

López Gutiérrez, J., Sánchez Jiménez, F., Fernández Villazala, T., Máñez Cortinas, C. J., Herrera Sánchez, D., Martínez Moreno, F., San Abelardo Anta, M. Y., García, M. R., Pérez, V. G., Santiago Orozco, A. M., Gómez Martín, M. A., Esteban, J. G., Arjona, R. P., Matos, G. M., Amado Hernández, M. P., & Molina, A. M. (2021). *Informe sobre la evolución de los delitos de odio en España 2021* [Report on the evolution of hate crimes in Spain 2021]. Ministerio del Interior. https://www.interior.gob.es/opencms/pdf/servicios-al-ciudadano/delitos-de-odio/estadisticas/INFORME-EVOLUCION-DELITOS-DE-ODIO-VDEF.pdf

Ministry of Equality. (2020). *Proyecto de Ley para la igualdad real y efectiva de las personas trans y para la garantía de los derechos de las personas LGTBI* [Draft law for the real and effective equality of trans people and for the guarantee of the rights of LGTBI people: Regulatory impact analysis report]. https://www.congreso.es/public_oficiales/L14/CONG/BOCG/A/BOCG-14-A-113-1.PDF

Miró-Llinares, F. (2016). Taxonomía de la comunicación violenta y el discurso del odio en internet [Taxonomy of violent communication and hate speech on the internet]. *IDP Revista de Internet, Derecho y Política*, *22*, 82–107. https://doi.org/10.7238/idp.v0i22.2975

Müller, K., & Schwarz, C. (2020). *Fanning the flames of hate: Social media and hate crime*. SSRN. http://dx.doi.org/10.2139/ssrn.3082972

Naser, A., Williner, A., & Sandoval, C. (2021). *Participación ciudadana en los asuntos públicos: Un elemento estratégico para la Agenda 2030 y el gobierno abierto* [Citizen participation in public affairs: A strategic element for the 2030 Agenda and open government] (Report No. LC/TS.2020/184). Comisión Económica para América Latina y el Caribe. https://www.cepal.org/es/publicaciones/46645-participacion-ciudadana-asuntos-publicos-un-elemento-estrategico-la-agenda-2030

Platero-Méndez, R. (2008). Transgenderism and political agenda: A history of (dis)continuities and pathologization. *Política y Sociedad*, *46*(1/2), 107–128. https://revistas.ucm.es/index.php/POSO/article/download/POSO0909130005A/21831

Regehr, K. (2022). In(cel)doctrination: How techno-

logically facilitated misogyny moves violence off screens and on to streets. *New Media and Society*, *24*(1), 138–155. https://doi.org/10.1177/1461444820959019

Rodríguez-Ruiz, B., & Mestre i Mestre, R. (2021, February 8). Ley Trans: Autodeterminación, felicidad y derechos. [Trans law: Self-determination, happiness and rights]. *El Diario*. https://www.eldiario.es/comunitat-valenciana/opinion/ley-trans-autodeterminacion-felicidad-derechos_129_7203213.html

Soral, W., Bilenwicz, M., & Winiewski, M. (2018). Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior*, *44*(2), 136–146. https://doi.org/10.1002/ab.21737

Sponholz, L., & Christofoletti, R. (2019). From preachers to comedians: Ideal types of hate speakers in Brazil. *Global Media and Communication*, *15*(1), 67–84. https://doi.org/10.1177/1742766518818870

Torregrosa, J., Panizo-Lledot, Á., Bello-Orgaz, G., & Camacho, D. (2020). Analyzing the relationship between relevance and extremist discourse in an alt-right network on Twitter. *Social Network Analysis and Mining*, *10*, Article 68. https://doi.org/10.1007/s13278-020-00676-1

About the Authors

Patricia Sánchez-Holgado is an assistant professor and post-doc researcher in the Department of Sociology and Communication at the University of Salamanca (Spain). A member of the Audiovisual Content Observatory (OCA), she holds a PhD in communication from the University of Salamanca. Her lines of research focus on communication, dissemination, and scientific culture; the social perception of science; online hate speech against vulnerable audiences; the adoption and use of communication technologies; and gender and equality studies.

Carlos Arcila-Calderón is an associate professor at the University of Salamanca (Spain), editor-in-chief of the journal *Disertaciones*, principal investigator of the European project "Preventing Hate Against Refugees and Migrants" (funded by the REC program of the European Commission), and local PI of "Enhanced Migration Measures From a Multidimensional Perspective" (HumMingBird, funded by H2020). His research interests focus on ICT adoption, media and social media analysis, computational methods, and big data.

Marcos Gomes-Barbosa is a PhD student in Education in the Knowledge Society at the University of Salamanca. He has a master's degree in Communication for Social Purposes: Strategies and Campaigns (University of Valladolid) and is also a member of the Audiovisual Content Observatory (OCA). His research topics are social networks, hate speech, gender studies, and social movements.

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 198–209
https://doi.org/10.17645/pag.v11i2.6375

Article

"My Way or No Way": Political Polarization and Disagreement Among Immigrant Influencers and Their Followers

Daniela Jaramillo-Dent

Department of Communication and Media Research, Media Change, and Innovation Division, University of Zurich, Switzerland; d.jaramillo@ikmz.uzh.ch

Submitted: 30 October 2022 | Accepted: 16 January 2023 | Published: 17 May 2023

Abstract
This article explores the online discussions between Carlos and Lizzy, two Latin American immigrant influencers in the United States with profiles on TikTok and Instagram. The dataset comprises a 90-minute live debate between them, that took place on September 25th, 2021, broadcast on Instagram and available on one of the profiles, as well as 1200 comments by 933 different viewers, received during the broadcast. The analysis relies on previous research on polarization, deliberation, and disagreement on social media platforms and it provides insights regarding the political and ideological diversity that exists among immigrant influencers and their followers. It focuses on the discursive strategies deployed by these content creators to discuss issues related to immigration reform and activism from two distinct political stances. It also provides a glimpse into topics of interest for the immigrant community as reflected by these content creators and commenting followers. The findings reflect the value of the ongoing relationship between content creators and their followers in the personal support and acceptance Carlos receives. In contrast, Lizzy is largely rejected and attacked, but a few of her arguments resonate with viewers. Comments about the debate itself are mostly negative due to the perceived low quality of the arguments, the mocking attitude of both debaters, and the need for a moderator to control the times. Comments that are critical of the debate often describe expectations of a more civil discussion and pathways to improve future debates.

Keywords
debate; immigration; influencers; Instagram; platforms; political polarization; social media; TikTok; United States of America

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

There are approximately 21 million immigrants of Latin American origin in the United States; of these, close to 7.4 million are undocumented (Millet & Pavilon, 2022). The political climate in recent years has been characterized by the Trump administration's stern policies establishing all forms of immigration as a threat to American society, hardening the United States' approach to immigration (Pierce & Selee, 2017). Meanwhile, the Biden administration has attempted to undo some of these policies, especially accepting refugees and reducing deportations (Krogstad & Gonzalez-Barrera, 2022). Historically, the social construction of Latin American immigrants in the United States has been shaped by perceptions of them as a monolithic group of "disposable strangers" regardless of their immigration status (Oboler, 2021).

This article analyses the live debate between two Latin American immigrant influencers: Izzy Lagüera and Carlos Eduardo Espina on TikTok and Instagram. The 90-minute debate in Spanish took place on September 25, 2021 and resulted from months of disputes between the two creators, who have conflicting political ideologies. On the one hand, Carlos is a Uruguayan immigrant, as well as a pro-immigration and human rights activist and a law student who supports

immigrants by providing important information about laws, requirements, and fundraising initiatives. He has also organized events such as *Un Día Sin Inmigrantes* on February 14, 2022, across 13 cities in the United States. On the other hand, Lizzy Lagüera is a Mexican immigrant who politically supports the republican party, especially Trump. She defends the right of immigrants to support the conservative party and periodically joins forces with a community of republican Latin American immigrants who support conservative ideologies and counter the community of liberal immigrant influencers. Moreover, the comments on the debate video reflect an audience profile as a group of people who are mostly Carlos' followers, speak Spanish, have a migration background, lean toward pro-immigration opinions, are from Latin American countries, and live across the United States according to the locations they mentioned when prompted by Carlos.

Both creators can be considered examples of minority celebrity, which has been conceptualized as:

> Fame and recognition founded on commodifying and representing a usually marginalised and stigmatised demographic of society, built upon the validation and celebration of minoritarian values, with the political agenda of making public and critiquing the systemic and personal challenges experienced by the minority group in everyday life. (Abidin, 2021, p. 600)

In this case, mainstream metrics of influence such as the number of followers are not as relevant, as these creators target minority communities, but the levels of interaction and content creation are significant for the marginalized communities they target. The difference in followers between them at the time of the debate—where the liberal influencer has significantly more followers than the conservative creator—reflects the ideological profile of the Latin American community in the United States, with 60% considering that the Democratic party represents their interests better compared to 34% who lean toward the Republican party (Krogstad et al., 2022).

The present study aims to bring light to how minority celebrities who have a personal stake in certain political issues discuss these issues and the understudied phenomenon of political polarization and deliberation among minorities. It also aims to understand the viewers' reactions to the arguments and perspectives presented in terms of their heterogeneous discussion or polarization possibilities.

The article starts with an overview of political polarization on social media and immigration as a key topic of contention. Then, an explanation of the role of immigrant content creators and influencers on social media follows, contextualizing it within the existing literature on influencers and their followers, while positioning the uniqueness of social media as a space for ideological and political deliberation and polarization. This is followed by a description of the digital ethnography and qualitative multimodal analysis carried out. The findings reflect the value of the ongoing relationship between content creators and their followers in the personal support and acceptance Carlos receives. In contrast, Lizzy is largely rejected and attacked, but a few of her arguments resonate with viewers. Comments about the debate itself are mostly negative due to the perceived low quality of the arguments, the mocking attitude of both debaters, and the need for a moderator to control the times. Finally, comments that are critical of the debate often describe expectations of a more civil discussion and pathways to improve future debates.

2. Immigration, Online Political Polarization, and the Potential of Social Media for Political Deliberation

It is relevant to define the different types of polarization that apply to political deliberations online such as the case explored in this article. In this sense, Jost et al. (2022, p. 562) differentiate between high "ideological polarization" as polarization that takes place when groups move toward extreme views on an issue, and "partisan issue alignment" which refers to situations where groups are divided but agree among themselves on various issues. Moreover, Yarchi et al. (2021) explore three modes of polarization that are relevant to online spaces: "interactorial polarization"—referring to homophily versus heterogeneity in online interactions—"positional polarization," which describes how online polarization is diverse and dependent on the platform itself, and "affective polarization", which refers to intense feelings expressed toward perceived members of the outgroup. Research on social media and polarization has found the phenomenon of echo chambers—understood as the process where people seek out like-minded individuals and information that is compatible with their opinion (Jost et al., 2022)—as key to understanding online opinion formation and disagreement. Research on echo chambers has been conflicting, with authors arguing that exposure to information on digital platforms can increase polarization (Bail et al., 2018), while others assert that selective exposure online has been overestimated (Barberá et al., 2015).

On the other hand, literature on the potential of social media to promote political deliberation has emphasized the importance of robust discussion characterized by civility and diversity to strengthen the quality of deliberation (Papacharissi, 2004). Moreover, Halpern and Gibbs (2013) describe the differences between the two platforms to enable a more equal deliberation and the relevance of sensitive topics in the emergence of impolite messages. Schäfer et al. (2022) explain how individuals are more likely to participate in online discussions if comments include evidence, which makes them feel more knowledgeable on the topic. This is important for the case explored in this article, as this is a topic that is highly personal and sensitive for the audience who have first-hand knowledge of the issue, and thus, it presents a unique context for political conversation.

In this sense, it is important to consider the relevance of immigration as a key issue in the North American political ecosystem. In January 2022 the Pew Research Center found that 67% of Republicans and 35% of Democrats considered immigration a top issue, and almost half of Americans (49%) considered immigration a priority for the year. Moreover, a Gallup poll in 2022 found that 75% of Americans consider immigration to be good for the country, with undocumented immigration being considered by the majority as a threat to national security (Gallup, 2022). Meanwhile, in a March 2021 survey, 39% of Latino adults mentioned their concern that a person close to them could face deportation and 84% of Latin American adults in the United States support a pathway for undocumented immigrants in the country to become documented, compared to 68% of the general United States population (Schaeffer, 2022). The significance of the Latino population in the United States has meant that these second, third, and fourth-generation migrants are considered a powerful political force in the country and one that gives voice to vulnerable immigrants who are unable to vote or even speak about the difficulties of the documentation process.

In this context, the case of immigrants is unique because research suggests that this community is notoriously voiceless in issues that affect them (Chouliaraki & Georgiou, 2019; Georgiou, 2018). Meanwhile, studies on polarization and public opinion have often mentioned immigration as a key political issue (e.g., Bail et al., 2018; Jost et al., 2022) but immigrants seem to be considered only when they become political subjects who can vote (Krogstad et al., 2022). In this sense, this exploratory article fills this gap in the literature by considering the opportunity of digital social media to get a glimpse of the political profile of Latin American immigrant communities online.

3. TikTok and Instagram as Spaces for Political Debate

Although the debate was recorded and uploaded to Instagram, the creators studied in this article gained their following on TikTok, and the debaters are active on both platforms. These platforms are relevant as spaces that have gained great popularity and become increasingly political in recent years. Instagram has become a space of visual (self)representation and very specific aesthetics and templates that generated unique digital cultures (Leaver et al., 2020). In the political realm, specifically related to immigration, studies have focused on the multimodal characteristics of #MigrantCaravan posts on Instagram (Jaramillo-Dent & Pérez-Rodríguez, 2021; Rosa & Soto-Vásquez, 2022).

Although the scholarly exploration of political content on TikTok is still emerging, the political nature of the platform cannot be denied. In recent months, the Ukraine war has played out on the platform as a space for resistance (Specia, 2022). Moreover, in the recent Italian elections, candidates turned to TikTok in an attempt to reach undecided voters (Pianigiani, 2022). Moreover, this platform has been pointed as key in shaping political communication and partisanship through cocreation functionalities such as the duet and overlaid text to express political and partisan content in the United States (Carson, 2021; Medina Serrano et al., 2020). These recent uses of the platform and analyses of political polarization are relevant for the present exploration of TikTok as a space for political discussion and polarization among immigrant influencers.

4. When Influencers Go Political: Personal Attacks and Social Media Wars

The political practices of influencers have been explored from different perspectives. For instance, a recent study on Finnish influencers found that many of them avoid political topics due to fear of the comments and attacks they could receive, this study also found that influencers are more likely to discuss politics when they receive positive reinforcement from their followers and the support they expect to receive from their followers in the case of a crisis (Suuronen et al., 2022). This shows the importance of follower-creator interactions for political engagement and the key role of followers in promoting and continuing to discuss political issues on social media, specifically within the context of social media influencers and their communities of support and/or opposition. Other studies have focused on the relationship between influencers and followers; this is the case of Zhang's (2022, p. 1) exploration of #stopasianhate on Instagram, which highlights followers' five types of responses: (a) educated, (b) feelings-based, (c) supportive, (d) resonating, and (e) seeking reasons.

The impact of influencers' personal characteristics in political debates has also been explored in previous literature. For instance, a study on Czech female influencers engaging in political debates found that they tend to feel that it is harder for women to be respected as political influencers. They also report feeling that the attacks they receive often refer to aspects unrelated to the political topic at hand such as their appearance, their mood, their disposition, or their wit. These creators feel that men do not receive such comments while expressing their political opinions in debates and that this happens across ideological lines (Vochocová, 2018).

5. Methods

To understand this instance of political debate, the analysis followed a qualitative approach and was guided by the followers' comments and their responses to the debate to respond to the three research objectives set out:

> RO1: Identify the multimodal and discursive characteristics of viewers' comments about the debater/argument/ideology they explicitly support.

RO2: Identify topics of interest and arguments within the debate that prompt discussion among viewers.

RO3: Explain how these comments and topics of interest reflect processes of heterogeneity/deliberation or polarization from the viewers toward the debaters and their arguments.

A 24-month digital ethnography was part of a larger study exploring immigrant content creators and influencers on social media residing in the United States and Spain. In this process, the researcher observed the ideological conflict between two Latin American immigrant TikTokers/Instagrammers in the United States as well as the attacks between them and their followers. The ethnographic method followed Pink et al.'s (2016) asynchronous presence approach as the fieldnotes were taken through observations of online content that was broadcast and commented live and then archived by the creators. The analysis was also informed by previous field notes from the extended observation of the community of Latin American creators that enabled the identification of these two immigrant influencers.

The 90-minute video was downloaded and viewed four times to inductively identify and code the topics discussed by the debaters. The entirety of the comments (1200) by 933 unique viewers was extracted in Excel format using the paid version of Export Comments (https://exportcomments.com). The video was coded based on the topics discussed in the debate as described in Table 1.

ATLAS.ti (version 22.2.3) was used to organize and code the comments and video. After translating comments that were originally in Spanish and categorizing them according to eight categories, open coding was carried out to identify emerging themes within each of the opinion-based categories of codes (Charmaz, 2000). The coding process focused on the commenter's perspectives about the issues discussed by the two debaters and their perspective of the debate itself. Figure 1 describes the coding process followed in this study for the comments.

The analysis was carried out between March–June 2022 through an iterative coding process that combined inductively derived codes—from the data, through an interpretative process—and deductive approaches—comparing the data with existing literature about online deliberation, where Zhang (2022) emerged as a relevant study to explore this case. The video and comments were coded separately but at a later stage, the topics that were discussed the most by viewers of the debate were identified. To maintain the commenters' anonymity, comments were translated from Spanish and slightly paraphrased without losing the original tone and intention to avoid identification. The findings are organized based on the three research objectives proposed and presented in the next section.

6. Results

6.1. Multimodal and Discursive Characteristics of Viewers' Comments

The analysis suggests that viewers expressed a closer relationship with and preference toward Carlos and distanced from Lizzy and her stance, as reflected in their language and the emojis used. This was expected considering the ideological profiles of both debaters and the significant difference in popularity between them—Carlos with 3 million versus Lizzy with 28 thousand at the time of the debate, as well as the fact that the debate was broadcast in Carlos' profile. Thus, it is likely that most of the viewers/commenters belonged to Carlos "community." Various discursive markers of interest support this. Comments such as "not even God knows this woman" point to Lizzy's status as an outsider. Meanwhile, Carlos is someone they know and support, as reflected in comments such as "let's all support Carlos so this fool sees

Table 1. Topics discussed in the debate.

Debate topics	Definition
Immigration reform	Points to discussions related to legislation or ideological stance regarding immigration reform, and specific immigration policies.
Political stance	Describes creators' mention of a politician, political party, or political leaning.
Attention economy	Labels instances when creators mention indicators related to the attention economy such as views, likes, monetization, or followers.
Personal attack	Signals attacks that target personal features of the opponent but not the issues discussed.
(Self)representation	Designates instances when creators describe their self-perception or who they are.
Debunk falsehoods	Indicates that the creator is attempting to discredit statements by the opponent by providing alternative information or asserting it is false.

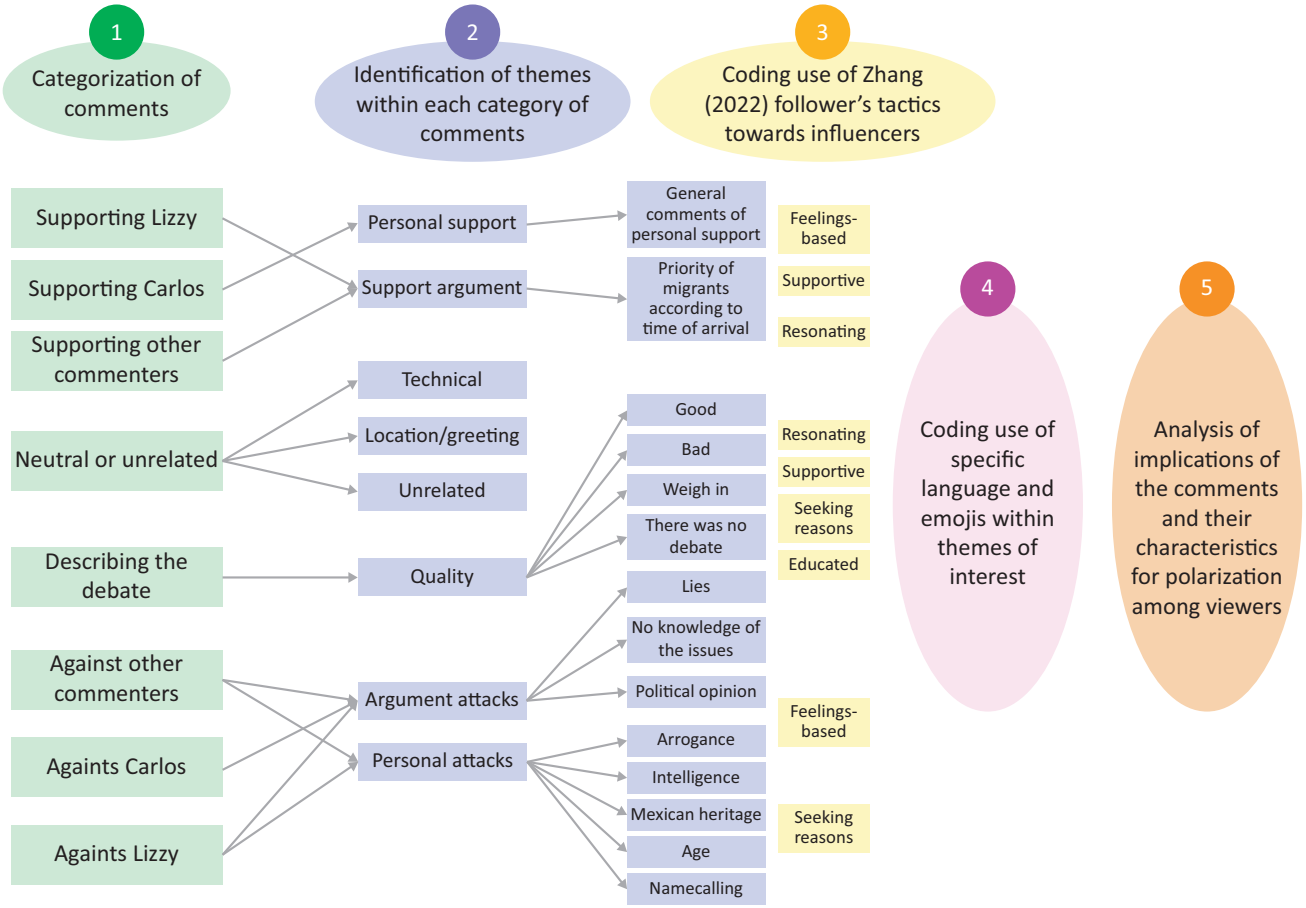


Figure 1. Coding procedure for comments.

that the people are behind him" reflecting his influence among these viewers/followers, which is relevant when considering the selection and preference for certain political information on social media (Anspach, 2017).

Lizzy's status as an outsider is also reflected in the prevalence of comments that speak directly to Carlos, both when speaking against her "Carlos, you wasted your time with this woman" and in instances of support toward her "sorry Carlos, I support you, but you did badly in comparison to her." In contrast, very few commenters speak to Lizzy in the first person. This difference reflects a longstanding relationship with him and a sort of ongoing, imaginary "conversation" between followers and the creator they follow. The comments supporting her arguments while specifying a preference for Carlos are relevant in the study of polarization, as some of these viewers reflect their openness toward alternative perspectives on the issue of immigration even when these perspectives come from a person perceived as an outsider. These instances reflect that, although they are a minority, some viewers can overcome ego and group justification which may lead to more polarized groups (Jost et al., 2022) by crossing the ideological aisle in specific topics that resonate with them.

Multimodally, the use of emojis reflects the viewers' preference for Carlos and their distance from Lizzy and her arguments. The use of emojis throughout the dataset is reflected in Figure 2, an emoji cloud where the size of the emoji illustrates its prevalence and its position in the cloud reflects the debater toward whom the emoji was directed, the figure is based on the cooccurrence analysis feature of ATLAS.ti.

Figure 2 illustrates viewers' acceptance and support for Carlos during the debate. It is noteworthy that although Carlos received most of the supportive messages and emojis, most of them were general and did not reflect a specific argument or statement. Comments such as "excellent work on behalf of the immigrant community" reflect ongoing forms of support for Carlos rather than the ideas he put forward during the debate. This was in stark contrast with the personal attacks Lizzy received, as reflected in Figure 3.

Some of the insults toward her refer to her age, political allegiance, and intelligence. Most of the comments she receives are attempts to minimize her political arguments as trivial and uninformed. This is similar to Vochocová's (2018) findings about female influencers engaging in political discussions and receiving comments unrelated to their political stance. Moreover, her support for Trump and his conservative, anti-immigrant agenda may also contribute to this type of verbal abuse considering the audience and their political leaning. On the other hand, some commenters attempt to distance themselves from Lizzy as a female Mexican immigrant in response

Figure 2. Emojis used by commenters toward both debaters.

to a key moment in the debate—when Lizzy stated: "I speak loudly because I am a Mexican woman"—as her comment reflects a stereotype of ethnic minorities and specifically female Latinas (Ghavami & Peplau, 2013). In response, viewers explained that she doesn't fit Mexican culture: "I am ashamed that she said that she is from Mexico"; "I am a Mexican woman, and I am not as vulgar as her." These commenters reflect different forms of gatekeeping of who is Mexican and what Mexican culture looks like as reflected in Figure 4.

In turn, viewers attempt to make sense of her immigrant background and her position against some forms of immigration with comments such as "got her papers and thinks she is a gringa" and other variations that reflect the perception that some immigrants in the United States discriminate against others once they become

Figure 3. Words used by commenters to refer to Lizzy during the debate. Note: Translated from Spanish by the author.

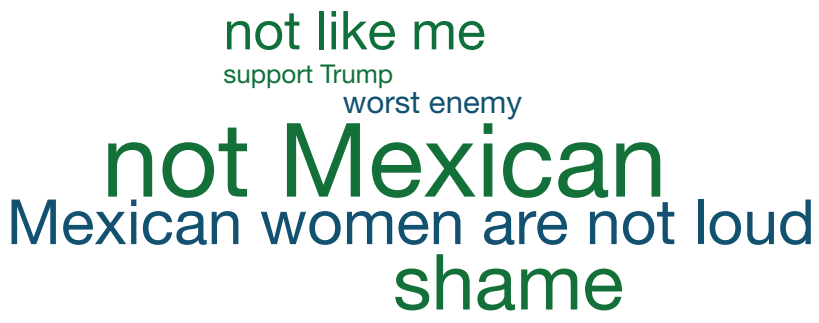


Figure 4. Words used by commenters to refer to Lizzy throughout the debate. Note: Translated from Spanish by the author.

documented. This reflects the horizontal forms of hierarchization, discrimination, and control that exist among immigrants online as described by Jaramillo-Dent et al. (in press).

6.2. Topics of Interest in the Debate According to Commenters

6.2.1. Potential for Deliberation in Support of Arguments While Rejecting the Speaker

As mentioned in previous sections, the audience was biased toward Carlos and his influencer/activist profile. However, one of the arguments that resonated with viewers the most throughout the debate was Lizzy's contention of "false hope" and perceived unfairness in the immigration process. In several instances during the debate, Lizzy identifies false hope as the key problem in Carlos' content related to immigration and the liberal agenda. Lizzy describes the rhetoric of false hope as "unfair to people who already live here, people who came as undocumented 20 or 30 years ago." Her rhetoric is vague and dispersed, as she does not provide any specific examples but only general observations about her perception of the situation at the border. The sensitive nature of the topic and the viewers' personal experience are reflected in their support for this argument which is never accompanied by an acceptance of her but of one of her arguments as part of a shared migratory experience. To Carlos' question "so why do you support a political party that doesn't want to legalize immigrants?" Lizzy also responds with a "false hope" argument, noting that Democrats (and Carlos) make promises that are never realized. Within academic texts, the idea of false hope for immigrants in Democrats' rhetoric has not been explored but recent news articles describing recent policies by Democrats mention the notion of false hope (Rappaport, 2022) that has led to confusion at the border (Kriel, 2021).

As the debate continues, Carlos attempts to discuss the need for immigration reform. Both debaters agree that immigration reform is necessary, but they disagree as to who is responsible and how this type of legislation should be proposed. As the conversation advances, Lizzy offers an argument giving reasons why immigrants should care about the situation at the border when she states:

> Any person whether they are legal, or illegal is affecting the system because the same system processes legal and illegal visas. It doesn't affect me personally because I have my citizenship but people who are home, waiting for those papers, who did it legally.…There is a way to do it, what is happening at the border is not the correct way.

Lizzie repeats this argument several times throughout the debate and in response to different arguments by Carlos. Her stance follows well-established perspectives that argue that immigrants should migrate "the correct way" following the misconception that immigrants can just gain documented status in the United States by applying. The American Immigration Council clarifies this misconception by explaining that many undocumented immigrants have no pathway to becoming legal residents or citizens in the United States and that even those who pay taxes and contribute are not able to gain documented status unless a new immigration reform is passed to provide them with legal status (American Immigration Council, 2019). What is rare is that this debater is an immigrant herself, who has gone through the system and doesn't seem to understand the difficulties faced by many immigrants entering the United States. In this sense, the debate provides a simplistic perspective about a very complex issue involving the lives of millions of people living in the United States and entering the country, who are unable to obtain a legal status due to a series of issues in the country's immigration processes.

On the other hand, Carlos argues for an approach that emphasizes human rights when he asserts:

> I support giving papers to all who are already here, do you support this?…The issue at the border is not as you paint it, they are people who want a better life, just like you and me.…I support the right of people to migrate, and some laws are unfair, so I want to work to help immigrants.

Carlos' perspective relies on the idea that borders and immigration legislation are often unfair in line with arguments for the abolition of borders. These arguments consider borders as symbols of inequality, that systemically constrain and render some individuals inferior to others due to their immigration status. Supporters of this perspective call for more humanitarian immigration processes (Anderson et al., 2009).

This is a key moment in the debate, as viewers also weigh in about the reality of the immigration system, a commenter asserts:

> No Carlitos, she is right about that, that's why the people who are here never see a light, because those who are entering are the priority, all the programs are for those who enter, and for those of us who have 10, 20 years here nothing.

This comment resembles Zhang's (2022) "resonating" reactions by commenters in the #stopasianhate conversation on Instagram. This comment reflects that regardless of her status as an outsider within this community, Lizzy's argument resonates with their experience or their perception of the immigration system.

In the next section, the attempts by both influencers to force the other to accept a political allegiance are explored.

6.2.2. Forcing a Political Allegiance

Throughout the debate, there are several instances where both participants attempt to corner the other within an established political stance. They do this by referring to the opponent in relation to a politician (such as Trump vs. Biden), a political party (Democrat vs. Republican), or a political stance (open borders vs. border control). They do this with statements such as "your president," "your political party," or "the ones you support." Carlos clearly establishes his perspective, as well as his political stance toward Democrat politicians and immigration issues:

> I can guarantee that very few or none of the people within the Democratic party support open borders….I don't consider myself a Democrat. I vote for Democrat candidates because they are the closest to my ideals but there are many Democrats that I don't support.

He reflects a critical and informed political posture and one that mimics many immigrant activists and members of the immigrant community who assert that the Democratic party has once and again betrayed the immigrant community. This stance also enables him to distance himself from some policies that the Democratic party brings forward concerning immigrants—such as criticism from progressive groups who have called on Democrats to "lean in" on issues of immigration, rather than evade them and let Republicans appropriate this issue (Bernal, 2022). This stance also allows Carlos to offer an alternative pathway that evidences his activism and support for immigrants without fully supporting a politician or political party. Conversely, Lizzy responds to this push toward a political ideology differently. She seems comfortable supporting Trump and his policies, stating that she believes that "Donald Trump won the 2020 elections" and defending his immigration policies by asserting:

> So, if finally, the Republicans had reached an agreement, when Trump was in office, they wanted to give some immigrants a pathway to citizenship.

The debate falls into an "all or no one" exchange where both debaters interrupt each other and do not let each other speak. Carlos responds by shifting his attention to the current legislation in congress that has not passed due to the lack of support by the Republican party, Lizzy deviates attention to topics that are unrelated and avoids answering the question.

In balance, Carlos attempts to discuss issues rather than politicians, although he prefers the Democratic party. Lizzy is comfortable openly supporting Trump throughout the debate and defending his administration and his policies. It remains unclear how Lizzy supports the immigrant community as she asserts. Moreover, commenters overwhelmingly reject Lizzy's support for Trump, as can be seen in Figures 3 and 4. As a result, her Trump-related arguments generate a wave of comments and attacks toward her and against the former president's discriminatory migration policies (Pierce & Selee, 2017).

6.2.3. Debunking False Statements

Much of the debate is spent attempting to debunk false statements by the opponent with no concrete evidence to differentiate fake from fact. Although clarifying data and information can be a great way to enrich the credibility and reliability of the issues discussed, both creators refer to the other's videos and their perception of the other's general approach to politics and immigration. This is problematic considering that both these creators have had their accounts temporarily blocked by TikTok, and viewers may not be able to find the videos they are discussing. It becomes a "he said she said" situation rather than a clear and solid debate on the serious issues that affect immigrant communities. As they discuss falsehoods, Carlos points to a few inconsistencies:

> First, when you say you don't use your platform to promote your half-crazed agenda on Trump, well, you surely know you are lying. Second, it is funny that you come to a Live event where thousands of people are watching to say that you support reform for undocumented people when we know that is not true, because there is a proposal right now in congress. In your videos, you have said: "Oh we don't need more illegals."

Lizzy's debunking style focuses on what she perceives as flawed sources of information Carlos uses such as Telemundo:

> You rely on sources such as Telemundo and Univision, hahaha….Your content is based on people's feelings, you have taken advantage of people's feelings and emotions. You have given people misleading information. You are telling them you are with them and supporting them, but you are not doing it.

Both debaters try to debunk the other's arguments based on the credibility of their statements and their sources. Lizzy's perspective that Telemundo and Univision—two of the main news channels in Spanish serving the Latin American communities in the United States—are unreliable sources follows a general trend among Republicans to doubt mainstream news and prefer to find alternative media for information (Allcott & Gentzkow, 2017). In the case of Carlos, he refers to her general political approach and her videos, which deviates from the debate itself and makes it hard to distinguish reliable information during their exchange.

6.2.4. The Attention Economy and Mutual Surveillance

The conversation between them falls into a series of discussions about the value of views/social media metrics and monetization as an inauthentic form of aid towards immigrants. In these conversations, Carlos defends the creation of videos to make visible the work he does for the immigrant community, and Lizzy argues that real help doesn't need as many views. Such an exchange is illustrated in the following:

> Lizzy: You care about views; you have been losing followers and supporters and you really care about that. People are not dumb, and they know what you are doing…you often have 70K views and only 5K likes.
>
> Carlos: I uploaded a video yesterday that has more than 700K views and if it doesn't [have views], I don't really care. Last week I had five videos with more than a million. If I was searching for views I would do it with someone who is worth it; you have like 10K followers, I don't understand what is the problem with creating videos to evidence the work that is being done in the community….I think what hurts you is that nobody cares or identifies with your videos, your only commenters make fun of you.

This exchange reflects an additional dimension for online political discussion and participation, the issue of attracting engagement in the attention economy. The attention economy refers to the shift from an economic system based on material goods to a system that relies on attention as proposed by Goldhaber (1997). This exchange also brings to light the mixed reactions these two creators receive from their commenters and the fact that both creators are surveilling each other's videos to assess their opponent's level of engagement. Lizzy mentions his lack of engagement as evidence of his failure to support immigrants. He responds with his argument that she has fewer followers and most of their comments make fun of her. This reflects the platformed dimension of the debate, where ideological opponents engage in attacks that relate to their success within the attention economy and whether they are relevant to the community. However, viewers' comments reflect a rejection of the discussion about likes and views:

> God, this idiot is focused on likes and views, seriously, she has no gray matter at all 🧠😳
>
> Omg! So much for the debate only for views or likes 👎👎 omg 😒 how boring
>
> We all like Carlos and she only wants fame

Many commenters perceive her motivations as attention-based and her participation as an attempt to gain visibility through Carlos. Many of the viewers lose interest when the two speakers start talking about popularity, which happens at different points in the conversation. This reflects the interest of many commenters and viewers in the issues rather than personal and superficial exchanges.

In the next section, the discussion about the debate itself is analyzed as a reflection on the audience's perception of healthy, valuable deliberation.

6.2.5. The Debate and Its Characteristics as Issues of Contention

In relation to the debate itself, comments mentioning the quality of the debate reflect the perception of worthlessness and a waste of time, hindering the possibility of healthy deliberation from the viewers in line with previous research on online deliberation (Halpern & Gibbs, 2013; Papacharissi, 2004). There are only a few comments that engage with migration-related topics in depth, focusing more on the personal attacks (toward Lizzy) and personal support (toward Carlos). However, there are a few "educated" comments that are critical of the debate itself and provide additional dimensions for consideration:

> It was not a debate, it seemed more like a personal dispute with a lack of legal knowledge. I agree that immigrants should come and that they should be gradually regularized. But there needs to be a process…if you agree to open the doors of your country without question is like opening the doors of your house!
>
> All opinions are valid but when they are imposed, they lose value and even more if they do not have good reasonable arguments.

As shown in these quotes and the previous results regarding the debate, many viewers expressed their desire for a more informed exchange, in line with Zhang's (2022) proposal of "educated" and "seeking reasons" comment typologies that question and provide critical perspectives about immigration while rejecting the debaters' vendetta. This suggests that there is potential and desire for a more deliberative discussion among viewers and that they are aware of the characteristics of a more robust/deliberative discussion about immigration as noted in the examples of their comments.

Comments that described the debate as "good" did not offer any details about the aspects that made the exchange good. Conversely, critical comments identified aspects that would have made it better such as a commenter who stated:

> I really expected more. I would like it to be repeated, having a moderator in charge, and keeping track of time.

Figure 5. The language used in comments to describe the debate. Note: Translated from Spanish by the author.

Such comments reveal the importance placed by viewers on the imbalance in speaking time (Figure 5), which reflects their understanding of equal participation as key to healthy political deliberation online (Halpern & Gibbs, 2013).

6.3. Implications for Deliberation Versus Polarization Among Viewers

The analysis throughout Section 6.2 emphasized the parallel development of the issues discussed in the debate and viewers' comments. This analysis provides isolated but valuable insights about the opportunity of such exchanges online for healthier deliberation among viewers and debaters. Moreover, the comments overwhelmingly reflect a debate that fails to discuss issues in a structured way and a polarized viewership that has intense affective responses to the topics and debaters due to their personal experiences and the sensitive issues/stances discussed. Although it could be argued that the viewers feature forms of "affective polarization" (Jost et al., 2022; Yarchi et al., 2021) toward the debaters and some of the issues mentioned, the unique case of immigrants calls for a nuanced analysis of a community that is personally affected by and have unique knowledge of the issues discussed (Schäfer et al., 2022).

Moreover, the perception of Lizzy as an outsider and Carlos as the preferred debater, as reflected in the comments, is relevant in terms of the possibility for a more heterogeneous deliberation. In this sense, Lizzy's lack of acceptance among viewers limits their openness toward the alternative views on immigration she presents. Thus, the personal contempt they express against her may drive viewers toward a more extreme and polarized political position as described by Jost et al. (2022). Moreover, the personal traits of the viewers, as reflected in their comments, suggest a group of Spanish-speaking Latin American immigrants living in the United States, who express support and allegiance to Carlos. This suggests that the required ideological diversity to achieve a healthier deliberation is not present in this online community, as the ideological majority is well-defined toward the left.

The topics analyzed and mentioned by commenters, including the role of the "attention economy" (Goldhaber, 1997) driving part of the conversation, are relevant to the analysis of social media platforms as unique spaces for deliberation and polarization. In this case, the perception of Lizzy as a less popular creator who engages in this conversation in a quest for increased celebrity deviates from the migration issues at hand and reflects a unique form of minority celebrity (Abidin, 2021) where the debater seems to engage in an extreme attempt to reflect her integration into the United States society and herself as a "good migrant" in line with previous research on the self-representation of immigrants as deserving (Georgiou, 2018; Jaramillo-Dent et al., in press).

7. Conclusions

This study brings light to how niche influencers and their followers discuss issues that are personally relevant to them and their potential for deliberation or polarization. First, the issues and conversations that frame these exchanges reflect the need to further theorize political opinion, deliberation, and polarization among minorities who are directly affected by the issues discussed. In this sense, mainstream studies exploring issue, party, and affective polarization (e.g., Jost et al., 2022; Yarchi et al., 2021) versus effective online deliberation (e.g., Halpern & Gibbs, 2013; Papacharissi, 2004) have traditionally focused on the general population. More research is needed to understand the processes of deliberation and polarization that take place among communities that have a personal interest in the issue. This research may illuminate the unique nature of grassroots political discussion among minorities.

The cultural sensitivity reflected by viewers who gatekeep Mexican culture and identity in comments against Lizzy is also relevant as an aspect that shapes the possibilities for political discussion. This reflects horizontal forms of policing and control among the immigrant community to establish the "correct" ways to be an immigrant (Jaramillo-Dent et al., in press) which in this case extends to ideological and behavioral aspects of her

profile. However, the presence of educated and critical comments by viewers about the quality of the debate sheds light on the interest of the community and their desire for a political discussion that focuses on issues that are unique to immigrants, and their interest in different perspectives on these issues, rather than only receiving information that supports their previous ideas. This uncovers the potential of the study of disenfranchised communities online to understand heterogeneous conversations among groups that share certain political, demographic, and life experiences such as immigrants.

The limitations of this study include the specificity of the case analyzed and the ever-changing nature of the platforms under study. In terms of future lines of research, it is necessary to carry out analyses and explorations of political conversations and opinions among racial, national, and ethnic minorities in an increasingly globalized and diverse world shaped by human mobility.

Conflict of Interests

The author declares no conflict of interests.

References

Abidin, C. (2021). Minahs and minority celebrity: Parody YouTube influencers and minority politics in Singapore. *Celebrity Studies*, *12*(4), 598–617. https://doi.org/10.1080/19392397.2019.1698816

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.%2031.2.211

American Immigration Council. (2019). *Why don't immigrants apply for citizenship?* https://www.americanimmigrationcouncil.org/sites/default/files/research/why_dont_immigrants_apply_for_citizenship_0.pdf

Anderson, B., Sharma, N., & Wright, C. (2009). Why no borders? No borders as practical politics. *Refuge: Canada's Journal on Refugees*, *26*(2), 5–18.

Anspach, N. (2017). The new personal influence: How our Facebook friends influence the news we read. *Political Communication*, *34*(4), 590–606. https://doi.org/10.1080/10584609.2017.1316329

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Fallin Hunzaker, M. B., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221. https://doi.org/10.1073/pnas.1804840115

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, *26*(10), 1531–1542. https://doi.org/10.1177/0956797615594620

Bernal, R. (2022, September 13). Progressive groups call on Democrats to lean in on immigration. *The Hill*. https://thehill.com/latino/3640470-progressive-groups-call-on-democrats-to-lean-in-on-immigration

Carson, D. (2021). *A content analysis of political discourse on TikTok* [Bachelor's thesis, University of Mary Washington]. Eagle Scholar. https://scholar.umw.edu/student_research/415

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 509–535). SAGE.

Chouliaraki, L., & Georgiou, M. (2019). The digital border: Mobility beyond territorial and symbolic divides. *European Journal of Communication*, *34*(6), 594–605. https://doi.org/10.1177/0267323119886147

Gallup. (2022). *Immigration*. https://news.gallup.com/poll/1660/Immigration.aspx

Georgiou, M. (2018). Does the subaltern speak? Migrant voices in digital Europe. *Popular Communication*, *16*(1), 45–57. https://doi.org/10.1080/15405702.2017.1412440

Ghavami, N., & Peplau, L. A. (2013). An intersectional analysis of gender and ethnic stereotypes: Testing three hypotheses. *Psychology of Women Quarterly*, *37*(1), 113–127. https://doi.org/10.1177/0361684312464203

Goldhaber, M. (1997). The attention economy and the net. *First Monday*, *2*(4). https://doi.org/10.5210/fm.v2i4.519

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. https://doi.org/10.1016/j.chb.2012.10.008

Jaramillo-Dent, D., Alencar, A., & Asadchy, Y. (in press). Migrant agency and platformed belongings: The case of TikTok. In K. Leurs & S. Ponzanezi (Eds.), *Doing digital migration studies: Theories and practices and the everyday.* Amsterdam University Press.

Jaramillo-Dent, D., & Pérez-Rodríguez, M. A. (2021). #MigrantCaravan: The border wall and the establishment of otherness on Instagram. *New Media & Society*, *23*(1), 121–141. https://doi.org/10.1177/1461444819894241

Jost, T., Baldassari, D., & Druckman, J. (2022). Cognitive-motivational mechanisms of political polarization in social-communicative contexts. *Nature Reviews Psychology*, *1*, 560–576. https://doi.org/10.1038/s44159-022-00093-5

Kriel, L. (2021, May 16). How inconsistent policies and enforcement have created false hope for migrants at the border. *Texas Public Radio.* https://www.tpr.org/border-immigration/2021-05-16/how-inconsistent-policies-and-enforcement-have-created-false-hope-for-migrants-at-the-border

Krogstad, J. M., Edwards, K., & Lopez, M. H. (2022). *Most Latinos say Democrats care about them and work hard for their vote, far fewer say so of GOP.* Pew

Research Center Race & Ethnicity. https://pewrsr.ch/3TOwBbT

Krogstad, J. M., & Gonzalez-Barrera, A. (2022). *Key facts about U.S. immigration policies and Biden's proposed changes.* Pew Research Center. https://www.pewresearch.org/race-ethnicity/2022/09/29/most-latinos-say-democrats-care-about-them-and-work-hard-for-their-vote-far-fewer-say-so-of-gop

Leaver, T., Highfield, T., & Abidin, C. (2020). *Instagram: Visual social media cultures*. Polity Press.

Medina Serrano, J. C., Papakyriakopoulos, O., & Hegelich, S. (2020). Dancing to the partisan beat: A first analysis of political communication on TikTok. *WebSci '20: 12th ACM Conference on Web Science* (pp. 257–266). Association for Computing Machinery. https://doi.org/10.1145/3394231.3397916

Millet, E., & Pavilon, J. (2022). *Demographic profile of undocumented Hispanic immigrants in the United States*. Center for Migration Studies. https://cmsny.org/wp-content/uploads/2022/04/Hispanic_undocumented.pdf

Oboler, S. (2021). Disposable strangers: Mexican Americans, Latinxs, and the ethnic label "Hispanic" in the twenty-first century. In A. Y. Ramos-Zayas & M. Rúa (Eds.), *Critical dialogues in Latinx studies: A reader* (pp. 67-80). New York University Press.

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, *6*(2), 259–283. https://doi.org/10.1177/14614448040414

Pianigiani, G. (2022, September 25). Hello, fellow kids: Italian politicians turn to TikTok for votes. *The New York Times*. https://www.nytimes.com/2022/09/25/world/europe/tiktok-italy-elections.html

Pierce, S., & Selee, A. (2017). *Immigration under Trump: A review of policy shifts in the year since the election*. Migration Policy Institute. https://www.migrationpolicy.org/sites/default/files/publications/TrumpatOne-final.pdf

Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2016). *Digital ethnography: Principles and practice*. SAGE.

Rappaport, N. (2022, June 20). Democrats' latest DREAM Act should be called "False Hope Act". *The Hill.* https://thehill.com/opinion/immigration/3529769-democrats-latest-dream-act-should-be-called-false-hope-act

Rosa, F. R., & Soto-Vásquez, A. D. (2022). Aesthetics of otherness: Representation of #migrantcaravan and #caravanamigrante on Instagram. *Social Media + Society*, *8*(1). https://doi.org/10.1177/20563051221087623

Schaeffer, K. (2022). *State of the union 2022: How Americans view major national issues.* Pew Research Center. https://www.pewresearch.org/fact-tank/2022/02/25/state-of-the-union-2022-how-americans-view-major-national-issues

Schäfer, S., Müller, P., & Ziegele, M. (2022). The double-edged sword of online deliberation: How evidence-based user comments both decrease and increase discussion participation intentions on social media. *New Media & Society*. Advance online publication. https://doi.org/10.1177/14614448211073059

Specia, M. (2022, March 25). "Like a weapon": Ukrainians use social media to stir resistance. *The New York Times*. https://www.nytimes.com/2022/03/25/world/europe/ukraine-war-social-media.html

Suuronen, A., Reinikainen, H., Borchers, N. S., & Strandberg, K. (2022). When social media influencers go political: An exploratory analysis on the emergence of political topics among Finnish influencers. *Javnost — The Public*, *29*(3), 301–317. https://doi.org/10.1080/13183222.2021.1983367

Vochocová, L. (2018). Witty divas, nice mothers and tough girls in a sexist world: Experiences and strategies of female influencers in online political debates. *Media, Culture & Society*, *40*(4), 535–550. https://doi.org/10.1177/0163443717729211

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2021). Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. *Political Communication*, *38*(1/2), 98–139. https://doi.org/10.1080/10584609.2020.1785067

Zhang, W. (2022). *The engagement between social media influencers and their followers on Instagram: A case study of #stopasianhate* [Master's thesis, Auckland University of Technology]. Tuwhera Open Access Theses & Dissertations. https://bit.ly/3SVOLYI

About the Author

Daniela Jaramillo-Dent is senior research and teaching associate at the Department of Media and Communication Research, University of Zurich. She is part of the Media Change and Innovation Division and currently works on the project "The chilling effects of dataveillance." Her research explores othering, belonging, identity, activism, and minority content creation on visual social media.

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 210–220
https://doi.org/10.17645/pag.v11i2.6397

Article

Southern European Journalists' Perceptions of Discursive Menaces in the Age of (Online) Delegitimization

David Blanco-Herrero [1,*], Sergio Splendore [2], and Martín Oller Alonso [2]

[1] Department of Sociology and Communication, University of Salamanca, Spain
[2] Department of Social and Political Sciences, University of Milan, Italy

* Corresponding author (david.blanco.herrero@usal.es)

Submitted: 30 October 2022 | Accepted: 16 January 2023 | Published: 17 May 2023

Abstract
In a new communication context, factors such as the rise of hate speech, disinformation, or a precarious financial and employment situation in the media have made discursive menaces gain increasing significance. Threats of this kind challenge the legitimacy of institutional news media and professional journalists. This article contributes to the existing literature on the legitimization of journalism and boundary work through a study that seeks to understand the perceptions of Southern European journalists of the threats that they encounter in their work and the factors that help explain them. To this end, a survey of 398 journalists in Spain, Italy, and Greece was conducted to learn what personal or professional factors influenced their views and experiences of discursive and non-discursive menaces. Results show that discursive threats, such as hateful or demeaning speech and public discrediting of one's work, are the most frequent to the safety of journalists, while expressions of physical violence are less common. Younger and more educated journalists tended to perceive themselves as having been victims of discursive menaces more often, although not many significant differences were observed between different groups of journalists. Even though it could show a worrying trend, this finding can also indicate a growing awareness about menaces of this kind.

Keywords
discursive menace; Greece; hate speech; Italy; journalists; legitimacy of journalism; Southern Europe; Spain

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

Spain, Italy, and Greece are three of the most important and influential countries of Southern Europe and the Mediterranean area. They share multiple sociocultural similarities and were some of the most affected countries by the 2008 financial and economic crisis, suffering particularly from high unemployment rates—especially among young people—as well as severe debt problems. In demographic terms, this crisis led to the emigration of large numbers of mostly young highly-educated people, leading to a general trend of older populations and negative net growth rates. However, these countries are also the three main gateways for immigration into the South of Europe, having similar migration patterns—what some have called the "Southern European migratory regime" (Arango, 2012; King, 2000). Furthermore, they share very complex political scenarios. Not only do they face economic and migratory challenges, but are strongly polarized in political terms, and populist and extremist parties have gained a significant presence in the government and parliamentary powers in recent years. Overall, Southern European countries share sociodemographic, cultural, political, and economic features that differentiate them from other European countries (Hall, 2007; Reher, 1998; Rhodes, 2015).

The connections between these three countries are obvious, but they also have many similarities in the

journalistic field. According to the theoretical model developed by Hallin and Mancini (2004), they belong to the Mediterranean media system. This model is defined by weaker professionalization, strong state intervention in the media, and high political parallelism between parties and news outlets. Even though this model has limitations and many changes have taken place since its academic formalization, the main patterns persist. In fact, the similarities are not only in journalistic production—on which this model focuses—but also in consumption, as stated in the Reuters Institute's *Digital News Report 2022* (Newman et al., 2022). In this sense, the three countries are similar in what concerns the growing importance of online news outlets and sharply decreasing print media consumption, whereas social media are frequent sources of news (71% of Greeks, 56% of Spaniards, and 47% of Italians claim to access news on social media). Trust in the news is rather low (35% in Italy, 32% in Spain, and 27% in Greece), following a clearly decreasing trend in recent years, and with significantly lower levels of trust in the news posted on social media. Moreover, the process of hybridization (Chadwick, 2013) and technological saturation (Harambam et al., 2018) are raising similar challenges for news media that often entail homogeneous efforts of adaptation within media systems that already have many features in common. All this makes it advisable, or even necessary, to go beyond a national approach and treat these three Southern European countries as a single entity.

Our research sought to understand the challenges and threats faced by journalists in Southern Europe because the current communication environment has given rise to a continuous discursive challenge to the legitimacy of institutional news media and professional journalists (Hanitzsch & Vos, 2017; Koliska et al., 2020). Many actors within, or on the borders of, the journalistic and political communication fields derive authority by delegitimizing news media (Van Dalen, 2021). This study makes an unprecedented effort to understand what perceptions journalists have about those discursive and non-discursive menaces to their work and what factors might explain those perceptions. Mediterranean journalism is particularly interesting because of its traditionally low level of professionalization and high level of political parallelism, but also because the serious economic crisis—not only the one in 2008, whose effects are not completely over, but also the ones that have followed, including the ones produced by the pandemic and the war in Ukraine—is especially affecting journalism in Southern Europe. The analysis that follows therefore aims to understand what factors make journalists feel that they are not legitimated and protected in their independent production of information.

1.1. Professional Journalism in Southern Europe

In recent years, journalism has undergone a profound transformation caused by new technologies, in particular, the irruption of social media (Carlson, 2018; Hermida, 2013) and the overall process of digitalization (Steensen & Westlund, 2021). This new paradigm has had a strong impact also at a professional level, generating new professional roles and backgrounds (Salaverría, 2016; Splendore, 2017; Splendore & Iannelli, 2022).

This new context has raised new challenges, and precariousness, for instance, is one of the most important ones. The financial weakness and the sales decrease of big media companies have led to redundancies, making the jobs of professionals more insecure (Blanco-Herrero et al., 2020; Figueras et al., 2012; Spyridou et al., 2013). This insecurity increases due to the lack of willingness to pay for journalism in the previously mentioned countries—around 12% are willing to pay for journalism, according to the *Digital News Report 2022* (Newman et al., 2022)—which makes the financial situation of many news outlets more difficult because they have to compete with free (sometimes partisan or low-quality) sources of information. These economic and labor aspects are also among the causes cited in some studies (Bakir & McStay, 2018; Blanco-Herrero & Arcila-Calderón, 2019) for the onset of another of the main challenges that journalism currently faces, namely disinformation, given that less time to compare sources or verify information often leads to a careless journalistic praxis. Additionally, a precarious employment situation also relates to a lack of formation or specialization, the need to cover multiple topics, or the need for immediacy, which are factors that affect the coverage of news by journalists (Solves-Almela & Arcos-Urrutia, 2021).

Disinformation is affecting the credibility of media as a whole, while also becoming a problem for democracies and societies in general (McNair, 2017). Closely connected with disinformation (Grambo, 2019) and the current situation, hate speech has also become a growing concern in the new communication scenario. Following the Council of Europe (1997) or the European Commission Against Racism and Intolerance (2016), we can define hate speech as any message whose objective is to spread discrimination, rejection, exclusion, humiliation, harassment, loss of prestige, and stigmatization of certain groups or people for belonging to a social group (identified by their national origin, ethnicity, color, religion, sexual orientation, etc.). The traditional targets are minorities and vulnerable groups, such as migrants (Arcila-Calderón et al., 2020). Journalists have also been targeted (Charitidis et al., 2020; Obermaier et al., 2018), not only when they belong to a specific stigmatized social group but also because of their exposure as public figures (Koliska et al., 2020; Van Dalen, 2021). Greek journalists identify hate speech as a recurrent problem in participatory spaces where they have to manage user-generated content (Saridou et al., 2019). At the same time, even if less commonly, hate speech has also been spread from within news outlets themselves (Sindoni, 2018), adding another dimension—though not a predominant one—to the phenomenon.

All these new challenges take place in a context in which the rise of social and digital media and, in general, online communication, have not necessarily created new problems but have increased the magnitude of existing ones. For example, anonymity or the use of pseudonyms (Anderson, 2007; Kim & Lowrey, 2015) offer users a sense of impunity and security when spreading content or performing actions that can be considered discursive menaces, given the difficulty of identifying and prosecuting them. Current threats like online harassment cannot be separated from the reality that is being depicted here (Relly, 2021).

Overall, the current scenario combines journalists made more vulnerable and insecure by precariousness with new tools and communication patterns that make discursive menaces more common. However, not only discursive menaces and threats arising from the new communication scenario are matters of concern. The worldwide score of the *World Press Freedom Index* (Reporters Without Borders, 2022), decreased in 2022, and the particular cases of the countries studied worsened as well. Spain dropped from position 29 to 32, Italy from 41 to 58, and Greece from 70 to 108.

Given all these factors, it is important to further explore the perceptions of Spanish, Italian, and Greek journalists, who represent the main Southern European countries, about the threats—both discursive and non-discursive—they face, and to analyze what factors help explain them. This is why we posed our first research question as such:

> RQ1: What perceptions do Spanish, Italian, and Greek journalists have about the potential discursive and non-discursive menaces they face? What factors can help explain these perceptions?

Moreover, to furnish a more detailed analysis, it is necessary to go beyond discursive and non-discursive menaces and seek to understand each of the potential threats independently. This explains our second research question:

> RQ2: What perceptions do Spanish, Italian, and Greek journalists have about the specific threats they potentially face? What factors can help explain these perceptions?

2. Method

2.1. Sample and Procedure

To answer these questions we surveyed journalists in the three countries. The survey was distributed by local researchers in each country using the Qualtrics software. The distribution started in June 2021 in Spain, in July 2021 in Greece, and in August 2021 in Italy. It was completed by September of the same year. The questionnaire was originally drafted in English and translated into each of the corresponding languages. The different amounts of time needed to complete the translation and ensure the quality of the survey explain why it was possible to start distribution earlier in Spain and Greece; in Italy, the delay is explained by the fact that our questionnaire was distributed together with a larger one as part of a broader project.

Different strategies were used to distribute the questionnaire. In Greece and Spain, the distribution was done through professional associations of journalists, trade unions, and similar organizations, and to media of all types to be shared in their networks. Moreover, individual journalists were contacted using a snowball strategy. E-mail was used for this purpose or, when not possible, social media. When no answer was received, the team sent up to two reminders with an interval of around one month between them. The response rate was very low in Greece and Spain—less than 10%, not including the journalists contacted by other journalists or organizations in a snowball procedure. It was equally low in Italy, where the distribution was made within the Worlds of Journalism Study (WJS) fieldwork, but the collection of data followed a more systematic and representative process specifically designed for this project. The difference between Greece and Spain, on the one hand, and Italy, on the other, is that, in the latter, a complete list of journalists exists, and the researchers were able to send each of them an e-mail.

The final sample comprised a total of 398 journalists. Using a confidence level of 95% and a 5% margin of error, and considering an estimated population of around 50,000 active and full-employed journalists in total in the three countries, the sample size should have been at least 381 people, which means that the sample had an acceptable sample size. However, even though the total sample size was adequate, the sample size in each country was not large enough. Given the difficulties of mapping a profession like journalism, a convenience sample was used in each country. Previous relevant studies had also used convenience samples of around 100 journalists per country (Hanitzsch et al., 2011; Pellegata & Splendore, 2017). Furthermore, exploratory analyses were conducted to evaluate whether the distribution of the sample was normal and descriptive information on the demographic variables was observed to evaluate the demographic representativeness of the sample.

2.2. Measures

The questionnaire used for the research was designed ad-hoc based on the one employed by the WJS project, modifying it to fit the goals of the project. WJS is the most significant international attempt to determine the state of journalism and journalists across the world. Among its many topics, safety issues and threats to the profession are of great importance. The questionnaire was validated by a group of experts in the field before being launched. Although it included a larger set of variables, the ones

used for the analysis reported in this article are shown in Table 1. Together with the dependent variables on the perceived threats, the variables studied focused on various factors: three sociodemographic, two educational, and five professional. These measures were taken into consideration because the different sociodemographic, educational, and employment conditions of journalists may influence their exposure to different threats, mostly due to their more or less precarious situation, but also to their different levels of awareness of different issues.

2.3. Analyses

After the construction and validation of the discursive menace and non-discursive menace indexes, various tests were conducted. First, we carried out, with an exploratory purpose, a descriptive analysis that also tested the normal distribution of the responses. Then, we first used the two scales specifically constructed for the analysis, studying how significant discursive menaces were perceived and their potential differences from other forms of threat to journalists, so that RQ1 could be answered. Afterward, tests were conducted in a more general way with all 10 potential threats to answer RQ2. For each of these cases, inferential statistics tests were conducted, trying to understand the factors that could help explain the potential differences: correlations for continuous variables (age, political ideology, and experience as a journalist), student's T-tests for two independent samples (gender, specialized education in the field,

Table 1. Variables used in the analysis.

Variable name	Description
Age	Continuous variable measured in years.
Gender	Although the option of "non-binary" was possible, all answers were either "male" or "female."
Political ideology	Continuous variable between 0 (*far left*) and 10 (*far right*).
Education level	Educational level reached by the journalist (without secondary education; secondary education; some unfinished university studies; undergraduate degree or equivalent; master's or equivalent; doctorate).
Specialized education in the field	Whether or not the journalist had completed studies in the field of journalism or communication.
Level of responsibility	Decision-making capacity of the journalist (no decision-making; operative decision-making; strategic decision-making).
Level of specialization	Whether the journalist worked on a specific newsbeat or as a generalist journalist.
Experience as a journalist	Continuous variable measured in years.
Funding of the news outlet	How the news outlet for which the journalist worked was financed (private company; public service/state-owned; different structures, such as community platforms, non-for-profit media, etc.).
Scope of the news outlet	Scope of the news outlet for which the journalist worked (local; regional; national; transnational).
Perceived experience of threats (dependent variable)	Using a Likert scale from 1 (*never*) to 5 (*many times*), the respondents were asked whether, in the last five years, they had experienced different types of risky situations in the context of their work: demeaning or hateful speech; public discrediting of their work; stalking; other threats or intimidation; surveillance; hacking or blocking of social media accounts or websites; arrests, detentions or imprisonment; legal actions against them because of their work; sexual assault or sexual harassment; and other physical attacks.
Perceived experience of discursive and non-discursive menaces (dependent variable)	To test the potential differences between discursive menaces and other types of threats, two scales were constructed, following an exploratory factorial analysis. Discursive menaces comprised four items (demeaning or hateful speech, public discrediting of the work, stalking, and other threats or intimidation; Cronbach's alpha = 0.829). Non-discursive menaces comprised six items (surveillance, hacking, or blocking of social media accounts or websites, arrest, detention or imprisonment, legal actions because of the work, sexual assault or sexual harassment, and other physical attacks; Cronbach's alpha = 0.638).

and specialization in a specific newsbeat), and one-way ANOVA tests for multiple independent samples (country, education level, level of responsibility, funding of the news outlet, and scope of the news outlet).

3. Results

Before proceeding with the presentation of the results, let us look at the composition of the sample. 51.8% of the surveyed journalists were male, 36.9% were female, and 11.3% did not state their gender. The mean age was 49.39 years (*SD* = 12.47) and the median was 49 years. The majority of the journalists worked in Italy (55.8%), followed by Spain (28.9%), and Greece (15.3%).

The above values show the adequacy of the sample. Journalism still tends to be a rather masculine profession, while the average age is not far from that observed in other studies, such as the Worlds of Journalism Study. Regarding the country of origin, there is a lack of balance, given that the Italian answers were collected together with a larger project, which made it easier to achieve a larger sample. The values for Spain and, especially, Greece, are lower. Another factor that justifies this particular result is that the population of journalists is the largest in Italy, followed by Spain and, finally, Greece.

The journalists surveyed tended to show a rather progressive ideology (*M* = 3.82; *SD* = 2.268), to have higher education qualifications (80% of them had a university degree), and 68.5% held specialized education in the field of communication or journalism. Regarding their careers, 40% had no management role, 37.1% had some operative decision capacity, and 22.9% had strategic decision capacity. Moreover, they had long experience in the field (*M* = 21.67; *SD* = 11.429) and they worked more frequently on general topics (60.7%) than on a specific newsbeat (39.3%). They mostly worked for private news outlets (68%) than for national (42.2%) or regional (21.7%) ones.

3.1. Journalists' Perceived Experience of Discursive and Non-Discursive Menaces

First, it is important to underline that we are focusing on the journalists' perceptions. We do not check—and methodologically there is no way to do so—if those threats were real. This is important because the same experience might be perceived differently by different journalists. However, the main effect of threats is their capacity to be perceived as such by the journalist, potentially affecting their conduct. In other words, if the journalist perceived that they had been a victim of any of the mentioned threats, their conduct might have been affected, no matter whether that threat can be officially considered as such (Hanitzsch & Vos, 2017).

Using the two scales created for this study, it is possible to compare the perceived experience of discursive threats and other kinds of threats. First of all, we can see that discursive menaces were obviously perceived more often (*M* = 1.95; *SD* = 0.878) than non-discursive ones (*M* = 1.33; *SD* = 0.449). Furthermore, 80.3% of all journalists claimed that they had experienced some form of discursive threat to some extent. This percentage fell to 58% regarding non-discursive threats.

No significant differences were perceived between male and female journalists. However, a significant and negative correlation was detected between age and having experienced discursive threats ($R(326) = -0.208$, $p < 0.001$), which means that the younger the journalist, the greater the likelihood that they had experienced some form of discursive threat.

Regarding ideology, there was a significant correlation between having experienced non-discursive menaces and political ideology ($R(312) = 0.12$, $p < 0.05$). Given that the higher values of the ideology scale corresponded to the right, this positive correlation means that more conservative journalists claimed that they had experienced such threats more often. No significant correlation was observed regarding the discursive menace.

The highest educational level obtained by the journalist seemed to be significant, both for the perception of having experienced discursive menaces ($F(5, 19.304) = 5.597$, $p < 0.01$) and for non-discursive menaces ($F(5, 18.502) = 3.298$, $p < 0.05$). The post-hoc tests showed that those journalists with some undergraduate studies were those experiencing the lowest perception of threats in both cases. These differences were only significant when the respondents with less educational level were compared with those who had completed undergraduate studies and those with master's degrees. No other relevant trend was observed.

The differences related to the completion of studies in the field of communication or journalism were also significant according to the student's T-tests conducted for the independent sample. Thus, the journalists with specialized studies (*M* = 2.07; *SD* = 0.899) had experienced more discursive menaces than those without ($M = 1.72$; $SD = 0.830$; $t(329) = -3.375$, $p < 0.001$, $d = 0.41$). The same was observed regarding non-discursive menaces ($t(303) = -2.703$, $p < 0.01$, $d = 0.34$), since those respondents with specialized studies (*M* = 1.38; *SD* = 0.433) perceived more threats of this kind than did those without specialized studies (*M* = 1.22; *SD* = 0.493).

The years of experience working as a journalist, the decision-making capacity, and the fact they worked on a specific newsbeat or as a generalist journalist, did not play a significant role in the perception of having experienced these two forms of threat. Neither the type of funding of the news outlet for which the journalist worked nor its scope, showed significant effects either.

3.2. Journalists' Perceived Experience of Risky Situations

Each of the 10 analyzed threats can be observed independently to furnish more detailed information. Being the victim of hate or demeaning speech was the most commonly experienced threat (*M* = 2.36; *SD* = 1.206), followed by

public discrediting of one's work (M = 2.30; SD = 1.211). More violent situations, such as arrest or imprisonment (M = 1.05; SD = 0.328), other physical attacks (M = 1.15; SD = 0.483), or sexual assault or harassment (M = 1.17; SD = 0.617), were very rare. Table 2 summarizes all the average values of the 10 situations studied.

Comparing the perceived experiences based on gender using student's T-tests shows that men (M = 1.46; SD = 0.867) experienced legal actions more often than women (M = 1.24; SD = 0.604; t(328.948) = 2.655, p < 0.01, d = 0.29). However, women (M = 1.32; SD = 0.825) had experienced sexual assault or harassment significantly more often than men (M = 1.06; SD = 0.354; t(168.169) = −3.480, p < 0.01, d = 0.41).

Considering the influence of age on these experiences, we see a significant and negative correlation with having experienced hate speech (R(329) = −0.168, p < 0.01), public discrediting of the journalist's work (R(329) = −0.206, p < 0.001), and other threats or intimidations (R(330) = −0.164, p < 0.01). These negative correlations mean that younger journalists claimed to have experienced these threats more often than older ones.

Focusing now on ideology, there is a significant correlation of this variable with having experienced hacking or the blocking of social media accounts or websites (R(313) = 0.146, p < 0.01), legal actions (R(314) = 0.187, p < 0.01), or other threats or intimidations (R(316) = 0.111, p < 0.05). These positive correlations mean that the more conservative a journalist is, the more they claim to have experienced these threats, and the opposite for more progressist journalists.

The educational level reached does not show significant differences or any trend that could be taken into account regarding the experience of risky situations. Only the three journalists without a secondary education seem to have experienced more of these situations, but the reduced size of the sample in this category prevents any statistically relevant comparison.

Respondents with specialized education in the field of journalism or communication claimed to have experienced more hate or demeaning speech (M = 2.47; SD = 1.228) than did those without education of this kind (M = 2.13; SD = 1.199; t(332) = −2.325, p < 0.05, d = 0.28). There are also significant differences regarding the experience of public discrediting (t(211.698) = −3.533, p < 0.01, d = 0.41), which was higher among journalists who had completed specialized studies in the field (M = 2.45; SD = 1.236) than among those who had not (M = 1.96; SD = 1.137). Journalists with specialized studies had experienced more surveillance (M = 1.98; SD = 1.202) than those without such studies (M = 1.56; SD = 1.118; t(312) = −2.904, p < 0.01, d = 0.41). A similar pattern is apparent regarding the hacking and blocking of social media accounts and websites. Those with specialized education experienced more of this kind of threat (M = 1.49; SD = 0.939) than those without (M = 1.27; SD = 0.769; t(236.975) = −2.217, p < 0.05, d = 0.26). Harassment was more frequently experienced among respondents with studies in the field (M = 1.55; SD = 0.947) than among those without (M = 1.27; SD = 0.766; t(2422.837) = −2.873, p < 0.01, d = 0.33). Finally, the experience of other threats or intimidations was more frequent among those respondents who had completed studies in the field of journalism or communication (M = 1.80; SD = 0.992) than among those who had not (M = 1.50; SD = 0.921; t(333) = −2.545, p < 0.05, d = 0.31).

No significant influence on the experience of risky situations was observed regarding years of experience as a journalist, the level of responsibility, and work on a specific newsbeat or general topics.

There were some differences based on the way the news outlet for which the journalist worked was funded. Experience of hateful or demeaning speech was significantly more common among journalists working for private news outlets (M = 2.50; SD = 1.185) than it was for those working in other types of media different from public or private ones (M = 2.02; SD = 1.094; F(2, 343) = 3.553, p < 0.05). The same is apparent when we consider the experience of public discrediting, which

Table 2. Experience of different risky situations during the past five years by Spanish, Italian, and Greek journalists.

Situation	*M*	*SD*	Percentage of the journalists who had experienced the threat to some extent
Demeaning or hateful speech	2.36	1.206	66.8%
Public discrediting of work	2.30	1.211	65.0%
Stalking	1.46	0.891	76.9%
Other threats or intimidation	1.69	0.964	42.0%
Surveillance	1.89	1.189	43.6%
Hacking or blocking of social media accounts or websites	1.42	0.873	24.8%
Arrest, detention, or imprisonment	1.05	0.328	02.9%
Legal actions against them because of their work	1.36	0.752	23.4%
Sexual assault or sexual harassment	1.17	0.617	09.7%
Other physical attacks	1.15	0.483	11.3%

is more common among journalists working for private media (M = 2.41; SD = 1.242) than among those working for other types of media (M = 2.00; SD = 1.071); $F(2, 102.414) = 3.412$, $p < 0.05$). These differences are also present regarding the experience of sexual assault or harassment, which was more frequent for journalists working in private media (M = 1.18; SD = 0.651) than in other types (M = 1.05; SD = 0.282; $F(2, 109.841) = 4.283$, $p < 0.05$). The other risky situations did not show significant differences based on this variable.

The experience of public discrediting of the journalist's work exhibited significant differences according to the scope of the news outlet ($F(3, 352) = 4.749$, $p < 0.01$), being higher among journalists working for regional (M = 2.61; SD = 1.172) or local (M = 2.51; SD = 1.150) media than among those working for transnational media (M = 1.89; SD = 1.201). There were also differences regarding the experience of the hacking or blocking of social media or websites ($F(3, 148.620) = 3.641$, $p < 0.05$), which were more common for journalists working in regional media (M = 1.64; SD = 0.916) than for those working in national ones (M = 1.28; SD = 0.774). No other situation showed significant differences for this variable.

4. Discussion and Conclusions

After the murders of the Maltese journalist Daphne Caruana Galizia in 2017 and the Slovakian journalist Ján Kuciak in 2018, the concern about the safety of European journalists grew (Coelho & Alves Rodrigues, 2020). This article has shown that Spanish, Italian, and Greek journalists do not perceive to be affected so much by violent threats as by a discursive form of threat, that has gained ground due to multiple factors. Although violent crimes are those that capture the most attention among media, scholars, and public opinion, we maintain that discursive threats—on the authority and legitimization of both journalism and journalists—are equally dramatic. This kind of attack is a constant feature of the environment in which journalists have to work—an environment where their legitimacy is not adequately recognized. Such attacks—hateful or demeaning speech, public discrediting, and threats perpetrated by any social actor—have gained presence, becoming the most common threat to the Mediterranean journalists' safety. In recent years, many studies have analyzed the challenges to the authority of journalism (Carlson, 2017; Van Dalen, 2021). The data used here give consistency to those concerns in a context in which the effects of digitization and economic crisis are considered even more influential.

Among the various factors that we analyzed, perhaps the most important is the fact that younger journalists perceive these discursive threats to a greater extent. Concerns about the digitalization of journalism (Waisbord, 2020) appear to be even more challenging when the younger generation is the one perceiving itself as working in a context that does not consider safe and in which journalists feel that their work is delegitimized. It is also interesting that it was precisely the respondents with higher education levels who more strongly perceived these discursive and non-discursive attempts to delegitimize their work, which is indicative of the fact that this influence is perceived by those who should know the importance of autonomy. Journalists who had completed specialized studies in the fields of journalism and communication claimed to have experienced more situations that posed a threat to their safety. This seems partly counter-intuitive because having specialized training should lead to higher and better-paid positions, and not make a journalist more vulnerable. However, it might make a journalist more aware of those risks, so they can better recognize them. This result significantly contributes to the boundary work literature (see Carlson & Lewis, 2015), because it makes it possible to identify a specific group that is struggling for its legitimacy not against the machine (Belair-Gagnon & Holton, 2018) or the audience (Robinson, 2010), but to affirm its professional right and duty to tell the truth about current affairs.

The rest of the factors analyzed concerned only very specific types of threats and the patterns were not significant when talking about discursive and non-discursive threats. The goal of studying those threats more in detail was to observe these cases in particular. We now briefly review and interpret them further.

Gender differences were not particularly common. Though it might be considered surprising that female journalists were not significantly more commonly victims than male ones in situations such as hate speech, as previous studies have suggested (Antunovic, 2019; Sarikakis et al., 2021). The clearest threat experienced by women is sexual assault—significantly more female journalists suffer from it—which corresponds with the literature findings on the subject (Harris et al., 2016). More surprising is the rather low level of reported sexual assault. Even though our study does not seek to explain these differences, one possible explanation might be the comparison with other (mostly discursive) threats. In absolute terms, sexual assault may be a significant and worrying issue, but the number of occasions on which it happens may be smaller than those of other threats. Therefore, even if some studies have shown that sexual assaults are rather common, they are still reported less frequently than other threats faced by journalists.

Another factor that showed significant effects was the political ideology of the journalists. About this, it was found that more conservative journalists experienced more threats (mostly non-discursive ones) even though the effect sizes were rather small. This may also be considered surprising, given that much of the existing literature focuses on the threats posed by far-right movements (Figenschou & Ihlebæk, 2019). However, the risk might come from both sides, as previous work on affective polarization has shown (Iyengar et al., 2019). Polarization is not likely to be the only factor explaining possible differences in the perceptions of conservative

and progressive journalists. Another potential explanation is the fact that at the time of the survey, the Italian and Spanish governments were respectively center and left leaning politically, and this may have contributed to the conservative journalists' perceptions about the menaces against them. The ideological proximity model usually explains these dynamics (Curini, 2022). However, future research might need to further explore its connection with the topics studied in this article.

Regarding the funding and scope of the news outlets for which the journalists worked, the differences were not generalized, and the effect sizes were also rather small. But it seems that journalists working for private and regional media were the ones claiming to have experienced the most threats. It could be argued that public-funded and national or transnational platforms are more stable and secure, thus being able to support and protect their journalists, who perceive fewer threats, whereas those working for private and regional (and maybe even local) platforms may be less protected or may have more precarious situations. Other factors, such as public exposure (in theory, greater in platforms with larger scope) or perceived risk of the threat (in local outlets being closer to the threat may increase the perception of risk) could also play a role, but they seem less explanatory of these findings. In any case, more research needs to be conducted in this regard.

The years of experience as a journalist, the level of responsibility, and the fact that they work on a specific newsbeat or on general topics (all of them being individual professional factors) played no significant role. The first two could have been expected to show some differences, given that experience and level of responsibility tend to give journalists a higher status and a greater capacity to deal with potential threats. However, this does not seem to affect their perceptions of these threats. In regards to the third factor, no comparisons were made between beats. Future studies will be able to further explore the potential existence of differences here.

It is important to stress that this study does not measure the existence of threats, but rather the perceptions of the journalists about them. Consequently, these perceptions may differ according to the aforementioned factors, even though the experiences may be similar. In general terms, discursive threats exhibit greater variability among different groups. Thus, the experience of hateful speech, the public discrediting of one's work, or other threats and intimidations present significant variations depending on some of the factors studied. Other threats, like physical attacks, stalking, or sexual assault, present fewer or almost no significant differences. One potential and preliminary explanation might be that the perceptions about having been a victim of hateful speech or similar discursive menaces differ from one person to another, whilst having been a victim of legal actions or physical violence might be more homogeneously perceived.

Overall, the main finding is that younger and more educated journalists claim to have experienced more threats, especially discursive threats. Among other reasons, this could be explained by the fact that new generations have gained awareness of the risks posed by discursive threats and are more sensitive to discursive challenges. This is the first step for journalists to combat increasingly worrying threats—such as hate speech—and defend their legitimacy against intruders.

Finally, a limitation of this study is that the analysis reported did not seek representativity, for which a larger and more probabilistic sample would have been needed. The purpose of the analysis was exploratory with the goal to test the influence of some independent variables on the dependent ones, for which no representative samples were needed. Furthermore, a snowball distribution procedure was followed in Greece and Spain and for that reason, only those journalists with stronger opinions about the topic were likely the ones who participated in the survey. Even though this strategy might be adequate for exploratory studies like this one, future research with larger and representative samples, especially regarding Spanish and Greek journalists, will be able to furnish more consistent observations. Lastly, although no results seem to hint that way, it is not possible to completely remove the potential effect of the Covid-19 pandemic, given that the study was conducted during the third semester of 2021. Future replications will be able to test potential longitudinal changes.

Acknowledgments

This research has been conducted with financial support from the Ministry of Universities of Spain under the FPU18/01455 grant within the FPU program, and from the European Commission, within the Rights, Equality and Citizenship Program (REC-RRAC-RACI-AG-2019), which finances the project Preventing Hate Against Refugees and Migrants (PHARM), with the grant agreement no. 875217.

Conflict of Interests

The authors declare no conflict of interests.

References

Anderson, P. (2007). *What is web 2.0? Ideas, technologies and implications for education*. JISC Technology & Standards Watch. http://21stcenturywalton.pbworks.com/f/What%20is%20Web%202.0.pdf

Antunovic, D. (2019). "We wouldn't say it to their faces": Online harassment, women sports journalists, and feminism. *Feminist Media Studies*, *19*(3), 428–442. https://doi.org/10.1080/14680777.2018.1446454

Arango, J. (2012). Early starters and latecomers: Comparing countries of immigration and immigration regimes in Europe. In M. Okólski (Ed.), *European*

immigrations: Trends, structures and policy implications (pp. 45–64). Amsterdam University Press.

Arcila-Calderón, C., Blanco-Herrero, D., & Valdez-Apolo, M. B. (2020). Rejection and hate speech in Twitter: Content analysis of tweets about migrants and refugees in Spanish. *Revista Española de Investigaciones Sociológicas*, *172*, 19–36. https://doi.org/10.5477/cis/reis.172.21

Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions. *Digital Journalism*, *6*(2), 154–175 https://doi.org/10.1080/21670811.2017.1345645

Belair-Gagnon, V., & Holton, A. E. (2018). Boundary work, interloper media, and analytics in newsrooms: An analysis of the roles of web analytics companies in news production. *Digital Journalism*, *6*(4), 492–508. https://doi.org/10.1080/21670811.2018.1445001

Blanco-Herrero, D., & Arcila-Calderón, C. (2019). Deontología y noticias falsas: Estudio de las percepciones de periodistas españoles [Deontology and fake news: A study of the perceptions of Spanish journalists]. *El profesional de la información*, *28*(3), Article e280308. https://doi.org/10.3145/epi.2019.may.08

Blanco-Herrero, D., Oller Alonso, M., & Arcila Calderón, C. (2020). Las condiciones laborales de los periodistas iberoamericanos. Diferencias temporales y geográficas en Brasil, México, Chile, España y Portugal [Working conditions of Ibero-American journalists. Temporary and geographical differences in Brazil, Mexico, Chile, Spain, and Portugal]. *Comunicación y Sociedad*, Article e7636. https://doi.org/10.32870/cys.v2020.7636

Carlson, M. (2017). *Journalistic authority: Legitimating news in the digital era*. Columbia University Press.

Carlson, M. (2018). Facebook in the news. *Digital Journalism*, *6*(1), 4–20.

Carlson, M., & Lewis, S. C. (2015). *Boundaries of journalism: Professionalism, practices and participation*. Routledge.

Chadwick, A. (2013). *The hybrid media system: Politics and power*. Oxford University Press.

Charitidis, P., Doropoulos, S., Vologiannidis, S., Papastergiou, I., & Karakeva, S. (2020). Towards countering hate speech against journalists on social media. *Online Social Networks and Media*, *17*, Article 100071. https://doi.org/10.1016/j.osnem.2020.100071

Coelho, P., & Alves Rodrigues, I. (2020). Rebuilding investigative journalism. Collaborative journalism: Sharing information, sharing risk. *Observatorio (OBS*)*, *14*(4), 135–157. https://doi.org/10.15847/obsOBS14420201748

Council of Europe. (1997). *Recommendation No. R (97)20 of the Committee of Ministers to member states on "hate speech."* https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=0900001680505d5b

Curini, L. (2022). What you seek is who you are: An applied spatial model of newspapers' ideological slant. *The International Journal of Press/Politics*, *27*(1), 96–119. https://doi.org/10.1177/1940161220973174

European Commission Against Racism and Intolerance. (2016). *General policy recommendation No. 15 on combatting hate speech*. https://rm.coe.int/ecri-general-policy-recommendation-n-15-on-combating-hate-speech-adopt/16808b7904

Figenschou, T. U., & Ihlebæk, K. A. (2019). Challenging journalistic authority: Media criticism in far-right alternative media. *Journalism Studies*, *20*(9), 1221–1237. https://doi.org/10.1080/1461670X.2018.1500868

Figueras, M., Mauri, M., Alsius, S., & Salgado, F. (2012). La precariedad te hace dócil. Problemas que afectan a la profesión periodística [Insecurity makes you docile. Problems that affect the journalist profession]. *El profesional de la información*, *21*(1), 70–75. https://doi.org/10.3145/epi.2012.ene.09

Grambo, K. (2019). Fake news and racial, ethnic, and religious minorities: A precarious quest for truth. *Journal of Constitutional Law*, *21*(5), 1299–1345.

Hall, P. A. (2007). The evolution of varieties of capitalism in Europe. In B. Hancké, M. Rhodes, & M. Thatcher (Eds.), *Beyond varieties of capitalism* (pp. 39–88). Oxford University Press.

Hallin, D., & Mancini, P. (2004). *Comparing media systems: Three models of media and politics*. Cambridge University Press.

Hanitzsch, T., Hanusch, F., Mellado, C., Anikina, M., Berganza, R., Cangoz, I., Coman, M., Hamada, B., Hernández, M. E., Karadjov, C. D., Moreira, S. V., Mwesige, P. G., Plaisance, P. L., Reich, Z., Seethaler, J., Skewes, E. A., Noor, D. V., & Yuen, E. K. W. (2011). Mapping journalism cultures across nations: A comparative study of 18 countries. *Journalism Studies*, *12*(3), 273–293. https://doi.org/10.1080/1461670X.2010.512502

Hanitzsch, T., & Vos, T. P. (2017). Journalistic roles and the struggle over institutional identity: The discursive constitution of journalism. *Communication Theory*, *27*(2), 115–135. https://doi.org/10.1111/comt.12112

Harambam, J., Helberger, N., & Van Hoboken, J. (2018). Democratizing algorithmic news recommenders: How to materialize voice in a technologically saturated media ecosystem. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*, Article 20180088. https://doi.org/10.1098/rsta.2018.0088

Harris, J., Mosdell, N., & Griffiths, J. (2016). Gender, risk and journalism. *Journalism Practice*, *10*(7), 902–916. https://doi.org/10.1080/17512786.2016.1166449

Hermida, A. (2013). #Journalism: Reconfiguring journalism research about Twitter one tweet at a time. *Digital Journalism*, *1*(3), 295–313. https://doi.org/10.1080/21670811.2013.808456

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N.,

& Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Kim, Y., & Lowrey, W. (2015). Who are citizen journalists in the social media environment? *Digital Journalism*, *3*(2), 298–314. https://doi.org/10.1080/21670811.2014.930245

King, R. (2000). Southern Europe in the changing global map of migration. In R. King, G. Lazaridis, & C. Tsardanidis (Eds.), *Eldorado or fortress? Migration in Southern Europe* (pp. 3–26). Palgrave Macmillan.

Koliska, M., Chadha, K., & Burns, A. (2020). Talking back: Journalists defending attacks against their profession in the Trump era. *Journalism Studies*, *21*(11), 1496–1513. https://doi.org/10.1080/1461670X.2020.1763190

McNair, B. (2017). *Fake news: Falsehood, fabrication and fantasy in journalism*. Routledge.

Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., & Nielsen, R. K. (2022). *Digital news report 2022*. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf

Obermaier, M., Hofbauer, M., & Reinemann, C. (2018). Journalists as targets of hate speech: How German journalists perceive the consequences for themselves and how they cope with it. *Studies in Communication and Media*, *7*(4), 499–524. https://doi.org/10.5771/2192-4007-2018-4-499

Pellegata, A., & Splendore, S. (2017). Media and corruption: The other way round—Exploring macro determinants of journalists' perceptions of the accountability instruments and governmental pressures. *International Journal of Public Opinion Research*, *30*(4), 561–582. https://doi.org/10.1093/ijpor/edx008

Reher, D. (1998). Family ties in Western Europe: Persistent contrasts. *Population and Development Review*, *24*(2), 203–234. https://doi.org/10.2307/2807972

Relly, J. E. (2021). Online harassment of journalists as a consequence of populism, mis/disinformation, and impunity. In H. Tumber & S. Waisbord (Eds.), *The Routledge companion to media disinformation and populism* (pp. 178–187). Routledge.

Reporters Without Borders. (2022). *2022 World Press Freedom Index*. https://rsf.org/en/index

Rhodes, M. (2015). Southern Europe in social science: A "Southern European model?" In M. Baumaster & R. Sala (Eds.), *Southern Europe? Italy, Spain, Portugal, and Greece from the 1950s until the present day* (pp. 51–76). Campus.

Robinson, S. (2010). Traditionalists vs. Convergers: Textual privilege, boundary work, and the journalist—Audience relationship in the commenting policies of online news sites. *Convergence: The International Journal of Research into New Media Technologies*, *16*(1), 125–143. https://doi.org/10.1177/1354856509347719

Salaverría, R. (2016). Redefinir al comunicador [Redefining the communicator]. *El profesional de la información*, *25*(2), 163–167. https://doi.org/10.3145/epi.2016.mar.02

Saridou, T., Panagiotidis, K., Tsipas, N., & Veglis, A. (2019). Towards a semantic-oriented model of participatory journalism management: Perceptions of user-generated content. *Journal of Education, Innovation and Communication*, *1*, 27–37. https://doi.org/10.34097/jeicom_SI_Dec2019-2

Sarikakis, K., Kassa, B. E., Fenz, N., Goldschmitt, S., Kasser, J., & Nowotarski, L. (2021). "My haters and I": Personal and political responses to hate speech against female journalists in Austria. *Feminist Media Studies*. Advance online publication. https://doi.org/10.1080/14680777.2021.1979068

Sindoni, M. G. (2018). Direct hate speech vs. indirect fear speech. A multimodal critical discourse analysis of *The Sun*'s editorial "1 in 5 Brit Muslims' sympathy for jihadis." *Lingue e Linguaggi*, *28*, 267–292. https://doi.org/10.1285/i22390359v28p267

Solves-Almela, J., & Arcos-Urrutia, J. M. (2021). ¿Ha cambiado la cobertura periodística de las migraciones en España? Una visión de los periodistas especializados [Has journalistic coverage of migrations in Spain changed?]. *Estudios sobre el Mensaje Periodístico*, *27*(1), 257–268. https://doi.org/10.5209/esmp.71342

Splendore, S. (2017). The dominance of institutional sources and the establishment of non-elite ones: The case of Italian online local journalism. *Journalism*, *21*(7), 990–1006. https://doi.org/10.1177/1464884917722896

Splendore, S., & Iannelli, L. (2022). Non-elitist truth? The epistemologies of Italian journalists in the hybrid media system. *Social Media + Society*, *8*(3). Advance online publication. https://doi.org/10.1177/20563051221118378

Spyridou, L.-P., Matsiola, M., Veglis, A., Kalliris, G., & Dimoulas, C. (2013). Journalism in a state of flux: Journalists as agents of technology innovation and emerging news practices. *International Communication Gazette*, *75*(1), 76–98. https://doi.org/10.1177/1748048512461763

Steensen, S., & Westlund, O. (2021). *What is digital journalism studies?* Routledge.

Van Dalen, A. (2021). Rethinking journalist–politician relations in the age of populism: How outsider politicians delegitimize mainstream journalists. *Journalism*, *22*(11), 2711–2728. https://doi.org/10.1177/1464884919887822

Waisbord, S. (2020). Mob censorship: Online harassment of US journalists in times of digital hate and populism. *Digital Journalism*, *8*(8), 1030–1046. https://doi.org/10.1080/21670811.2020.1818111

About the Authors

David Blanco-Herrero is a PhD student and predoctoral researcher at the Department of Sociology and Communication of the University of Salamanca. He holds a bachelor's in journalism from the Distance University of Madrid (2016) and in business management from the Universidad de León (2014), and a master's in audiovisual communication from the University of Salamanca (2018). He is editor of the *Anuario Electrónico de Estudios en Comunicación Social "Disertaciones*." He has participated in 12 research projects and has published 27 articles in peer-reviewed academic journals.

Sergio Splendore (PhD) is an associate professor at the University of Milan. His fields of expertise are journalism and political communications. His works are about digital journalism, epistemology, and media trust. On those topics, he has published articles in peer-reviewed journals such as *Journalism*, *Journalism Studies*, *International Journal of Press/Politics*, *New Media and Society*, and many others.

Martín Oller Alonso has a PhD in journalism from Rey Juan Carlos University (2012) and is involved in several international research projects, such as the World of Journalism Study, Journalistic Cultures, and Pre-Professional Journalistic Culture. He has published 25 scientific books and collaborated on almost 60 more, produced more than 70 articles, and participated in nearly 90 conferences. He currently pursues his scientific activity in Latin America, while conducting a postdoctoral residency at the University of Milan as part of the MediaDelCom project.

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 221–234
https://doi.org/10.17645/pag.v11i2.6419

Article

Polarization in Media Discourses on Europeanization in Spain

Ana Pérez-Escoda [1,*], Sonia Boulos [2], María-José Establés [3], and Lucía García-Carretero [4]

[1] Faculty of Communication and Arts, Antonio de Nebrija University, Spain
[2] Faculty of Law and International Relations, Antonio de Nebrija University, Spain
[3] Faculty of Information Sciences, Complutense University of Madrid, Spain
[4] Faculty of Economics and Business Administration, Madrid Open University, Spain

* Corresponding author (aperezes@nebrija.es)

Submitted: 31 October 2022 | Accepted: 12 January 2023 | Published: 17 May 2023

Abstract
Political and media polarization has had a detrimental impact on democratic principles and democratic processes on a global scale. In Europe, such polarization has eroded the trust in national and European institutions and has challenged the basic values that stand at the heart of the European integration project. The aim of this study is to analyze Spanish media discourses on Europeanization, with an attempt to identify key areas in which polarizing narratives related to Europeanization are more prevalent. To conduct our study, six national media outlets were selected based on four criteria: media format, ownership, ideology, and consumption. A final sample of 540 news items collected between July 2021 to March 2022 was selected for analysis. Using a qualitative methodological approach, the study was carried out in two stages. In the first phase, we conducted a content analysis to identify the main topics discussed in relation to the European Union and the actors represented in them. This led to the identification of polarizing narratives and discourses emerging in the context of the discussed topics. In the second phase, we used critical discourse analysis to analyze polarizing discourses.

Keywords
democracy; Europeanization; European Union; media discourses; polarization; political communication; Spanish media

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

Democratic stagnation has been on the rise since the beginning of the 21st century (Carothers & O'Donohue, 2019). The crumbling support for traditional parties, the rise of illiberal political leaders exhibiting autocratic traits, the increasing influence of authoritarian powers, and the unprecedented levels of political polarization are becoming a major challenge for new and established democracies alike (Arbatli & Rosenberg, 2021; Carothers & O'Donohue, 2019; Kaufman & Haggard, 2019). In Europe, old patterns of consensual competition between the center-right and the center-left are being challenged by political polarization and populism, endangering, hence, traditional political institutions and norms (Carothers & O'Donohue, 2019). This is reflected in the recent rise of anti-establishment parties in Europe, and in the increasing distance between political parties and the erosion of a common ground between their respective voters (Casal-Bértoa & Rama, 2021). Attempts to conceptualize political polarization pose a complex question on how to draw a line between positive polarization that enriches democratic processes and harmful polarization that hinders the legitimacy of the entire system and the values it represents (Carothers & O'Donohue, 2019; Pausch, 2021). Kaufman and Haggard (2019, p. 418) conceptualize polarization as "a cumulative process through which cross-cutting cleavages are submerged into a single, re-enforcing dimension that pits 'us' versus 'them' on a range of issues." According to DiMaggio et al.

(1996, p. 693), polarization is both a state and a process. Polarization as a state refers to "the extent to which opinions on an issue are opposed in relation to some theoretical maximum." Polarization as process refers to "the increase in such opposition over time" (DiMaggio et al., 1996, p. 693). Polarization is dangerous because it hinders group formation at the center of opinion distribution and fosters the formation of groups with irreconcilable policy preference (DiMaggio et al., 1996). Acute political polarization leads voters "trade off democratic principles for partisan interests," making highly polarized societies more vulnerable to democratic backsliding (Svolik, 2019, p. 23). Pausch (2021) highlights four characteristics of polarization: discrepancy of opinions, group formation that divides the world into "us versus them," purism and the rejection of reconciliatory positions, and the existence of a political struggle to promote polarized positions.

The European Union (EU) has never been a polarizing political project in Spain. Historically, the public opinion in the country reflected a high degree of pro-Europeanness, as well as a strong consensus among Spanish elites regarding European affairs (Vázquez et al., 2014). Moreover, empirical studies, such as the one conducted by Sojka and Vázquez (2014), demonstrate that the identification of Spanish elites with the European project (above 90% identify with the EU) is stronger than the identification of Spanish public with it (60%). Among Spanish elites, political elites show the highest levels of identification with the EU, followed by media and union elites (Sojka & Vázquez, 2014). This general agreement among the elites and the citizens on Europeanization has resulted in clear support for the idea of shared governability between the EU and its member states (Sojka & Vázquez, 2014).

Having this context in mind, the aim of this study is to analyze Spanish media discourses on the EU in an attempt to identify polarizing narratives related to Europeanization, and key areas in which such narratives arise. It also attempts to identify the actors behind the emergence of polarizing narratives and the communicative strategies they deploy to promote them.

2. Europeanization and Polarization in Spain

Following Dutceac and Bossetta (2019, p. 1054), we use the term Europeanization to refer to "any process whereby a feature of the domestic (whether it be an identity, a policy, or a discourse) takes on a European dimension." Europeanization in Spain has been marked by a strong political pro-European consensus from the moment Spain first applied for membership in the European Community (EC) until today (Avilés, 2004; Real-Dato & Sojka, 2020; Ruiz & Egea, 2011). After the death of Franco, "Europeanism" become a national project promoted by the leading political figures of the transition period. The emerging political consensus identified freedom and democracy with modernization and Europeanization (Ruiz & Egea, 2011). Joining the EC was largely perceived as an opportunity for Spain to overcome years of isolation during Franco's dictatorship. In the early years of accession, the Spanish public opinion believed that the greatest benefit derived from EC membership consists in Spain's increased role in world affairs (Powell, 2007). While marginal, some Euroscepticism existed in the first decades following Spain's accession to the EC. For example, the United Left (IU) criticized the neo-liberal orientation of European institutions and their democratic deficit (Avilés, 2004; Ruiz & Egea, 2011). The IU's position could be categorized as "soft" Euroscepticism. The latter is defined as the existence of:

> Not a principled objection to the European integration or EU membership but where concerns on one (or a number) of policy areas lead to the expression of a qualified opposition to the EU, or where there is a sense that "national interest" is currently at odds with the EU's trajectory. (Taggart & Szczerbiak, 2008, p. 8)

Today, as Arregui (2022, p. 1) points out, the EU "has become the most important frame of reference for [Spanish] national actors." The early stages of Spain's membership in the EC were marked by European and local debates that lead to the adoption of the Single European Act in 1987, culminating in the establishment of the single market. Spain had to prove that it is a trustworthy partner, capable of developing a "Europeanist" approach that transcends narrow national interests (Powell, 2007). This meant that the government of Felipe Gonzalez had to support initiatives that purported to deepen the EC without demanding too much in return in a period of internal economic crisis. The EC's decision to grant regional, structural, and development funds to enable less developed members to adapt to the single market helped the Spanish government to promote the idea that "deepening" the EC is fully compatible with the promotion of national interests (Ruiz & Egea, 2011). Fear of peripheralization was evident as early as 1989 in governmental circles due to the economic vulnerability of Spain. As a result, Spain vehemently rejected the idea of "multi-speed Europe," which entailed that the different European countries would integrate at different levels and pace, depending on their internal political situation (Powell, 2007).

The unanimous support for joining the EC did not reflect a unified vision of Europe. The two major Spanish parties—the Spanish Socialist Worker's Party (PSOE) and the Popular Party (PP)—had different visions on how Europe should be constructed. For the PP, it was a project that should be based on cooperation between states, rejecting, hence, the idea of a federal model of the EU. On the other hand, the socialists believed in the construction of a more social Europe, in the adoption of the European Social Charter, and in the territorial cohesion

of the EU (Avilés, 2004; Powell, 2007). The debates on the adoption of a European constitutional treaty demonstrated best the gap that existed then between the PP and PSOE. While the PP government, led by Jose Maria Aznar, was in favor of a European constitutional treaty, it insisted that Spain should retain the same power within the EU institutions that it had under the Nice Treaty. This position was not shared by the PSOE (Ruiz & Egea, 2011). The PP was also very critical of the European policies of Felipe Gonzalez, accusing him of failing to defend Spain's national interests in the EC. The PP's neoliberals argued that Spaniards had become accustomed to a wide range of public subsidies, under the auspices of González's European policy, which were incompatible with genuine socio-economic modernization (Powell, 2007). However, when the PP came to power, Aznar's own rhetoric resonated with that of the PSOE. To justify the implementation of internal reforms, required for incorporating the single currency, the government argued that it was in Spain's interest to make more Europe and not less Europe (Roch, 2019). Both parties had also shown a similar approach to the eurozone crisis, and both implemented drastic austerity measures under the auspices of EU institutions. Both parties supported the 2011 amendment of Article 135 of the Spanish Constitution, which prioritized the payment of the Spanish debt over social spending (Roch, 2019). This suggests that the ideological gap between the two parties on the construction of the European project is not very significant.

2.1. Europeanization and End of Bipartidism in Spain

Spain was hit hard by the eurozone crisis and by the austerity measures mandated by the EU. This led to the emergence of anti-austerity public mobilizations that voiced criticism of European institutions. In 2011, the 15-M movement emerged as a social movement against austerity measures (Feenstra et al., 2017), and it set the stage for the emergence of new political parties ending decades of bipartisanism in Spain (Pavan, 2017). The institutionalization of these social movements led to the establishment of Podemos (in English, "We Can") in 2014, a political party that presented itself as an alternative to the traditional left (della Porta et al., 2017). While Podemos challenged the neoliberal policies promoted by national and EU institutions and elites, it did not reject the European project altogether. Instead, it attempted to redefine the meaning of Europeanization, as reflected in the words of one of its most prominent founders, Pablo Iglesias (Iglesias, 2015, p. 27): "The strategy we have followed is to articulate a discourse on the recovery of sovereignty, on social rights, even human rights, in a European framework." A 2014 manifesto, initiated by the founders of the party, alludes to the "crisis of legitimacy of the EU," with the EU mandated austerity measures being labelled as a "financial coup d'état" against Southern European countries (Podemos, 2014, p. 1). Podemos described the amendment of Article 135 of the Spanish constitution as "the surrender of sovereignty" (Fernandez, 2014). The party portrayed Europe as a fractured project between the North and the South (Roch, 2019), with Southern Europeans wanting to recover "the dignity and the future" of their peoples (Fernandez, 2014). While emphasizing that the party "loves Europe" if it means "freedom, equality, and fraternity," Podemos accused the Spanish political elites of wanting to convert Spain into "a German colony." It also accused European elites of corruption, linking the president of the European Central Bank, Mario Draghi, to Goldman Sachs, and comparing Jean-Claude Juncker to Al Capone (Fernandez, 2014).

Some mainstream media outlets attempted to associate Podemos with far-right positions towards the EU. This was met with an increasing effort by Podemos to present itself as a pro-European party, fighting for a new social democracy in Europe (della Porta et al., 2017). In the 2015 elections, the party's manifesto lacked an explicit criticism of the EU. Reference to the EU included mainly a call to advance the European integration in key policy areas, such as single fiscal policy, eurobonds, common EU budget, common defense, and security policy (Real-Dato & Sojka, 2020). Likewise, in the 2016 elections, the EU was barely present in the manifesto of Unidas Podemos (United We Can), a coalition formed by Podemos and IU. The manifesto included a positive stance towards European integration, calling for the democratization of the eurozone economic governance, the flexibilization of the Stability and Growth Pact, and the creation of a common fiscal policy and a European investment plan (Real-Dato & Sojka, 2020).

Another newly established far-right party that introduced Eurosceptic discourses was VOX. The party became the third parliamentary party in the November 2019 general elections (Real-Dato & Sojka, 2020). The party was formed in 2013 by members of the PP who were disappointed with the policies of the government, which they perceived as too moderate (Real-Dato & Sojka, 2020). VOX capitalized on the migration crisis and the political crisis triggered by the separatist process in Catalonia to rally political support. Both crises had bearing on the public perception of the EU (Real-Dato & Sojka, 2020). Like Podemos, VOX did not advocate a straightforward rejection of the European integration project. In the 2014 European elections, VOX defined itself as deeply Europeanist (VOX, 2014). Failing to gain seats, VOX intensified its nationalist and nativist positions in subsequent national, regional, and European elections. This entailed the adoption of a Eurosceptic stance calling for the respect of the supremacy of national law over EU law and promoting an inter-governmental model of the EU (Lerín, 2022; Real-Dato & Sojka, 2020). In its 2019 manifesto for the European elections, VOX reasserted its "Europeanist vocation," while concurrently rejecting "the goal of turning the Union into a disguised federal state," which in the party's view has been achieved "fraudulently in recent years by reducing the [number

of] issues that require unanimity within the Council, conferring greater power on the Commission" (VOX, 2019, p. 5). VOX also called for repealing or adopting a "radical amendment" of Article 7 of the Lisbon Treaty, in order to prevent the EU from interfering in the democratic decisions of member states based on the "respect for democratic values" clause (VOX, 2019, p. 8). VOX accused the EU and its institutions of political bias against certain member states due to the influence of leftist ideologies. Referring to the successionist movement threatening the territorial integrity of Spain, the manifesto accused the EU of being "stingy in defending the unity of Spain as the basis and foundation of Europe," describing the current state of integration as "European project of elites far from reality" (VOX, 2019, p. 6).

However, the Eurosceptic stance of both Podemos and VOX lacks consistency when analyzed over time. This suggests that their discourses on the EU are usually shaped by their electoral strategies (Real-Dato & Sojka, 2020).

2.2. Media and Europeanization in Spain

The Spanish media is characterized by a high level of political parallelism (Teruel-Rodríguez, 2016). This term refers to "a pattern or relationship where the structure of the political parties is somewhat reflected by the media organizations" (de Albuquerque, 2018). In such contexts, polarization requires the confluence of political and media actors (Teruel-Rodríguez, 2016). In the absence of anti-European agenda in the Spanish political landscape, it is not surprising, then, that media elites in Spain show high levels of identification with the EU (Sojka & Vázquez, 2014). However, Spanish media outlets remain highly focused on national political systems. The coverage of the EU is given a lower priority (Sotelo, 2009). A study by Berganza (2009) suggested that more attention was given to the EU by the Spanish media when the new socialist government prioritized Spain-EU partnership over the United States–Spain bilateral relations, prioritized by the previous government.

Another feature of the coverage of the EU by European media in general is its dependence on official sources. As a result, the media tend to reproduce the messages of experts, especially in economic matters (Arrese & Vara-Miguel, 2016). Rivas-de-Roca and García-Gordillo (2023) point out that excessive dependence on official sources could create a journalism of statements, that mostly consists of the transferal of the interests of the political class to citizens. This could enhance the existing distrust among citizens towards media, as pointed by Pérez-Escoda (2022).

A comparative study by Menéndez (2010) suggested that even when a media outlet is pro-European it can still portray the EU negatively. According to him, there were more negative stories about the EU in pro-European Spanish media (and in France and the UK) compared to positive or neutral stories. This could be explained by the general tendency of the media to highlight negative or conflicting news to attract readers. The prioritization of negative frames, such as the recurrent use of the term "crisis," could eventually diminish the publics' identification with the European project (Rivas-de-Roca & García-Gordillo, 2023).

3. Objective and Methodology

This study analyzes Spanish media discourses on the EU, and it aims to identify polarizing narratives related to Europeanization and key areas in which such narratives arise. It also attempts to identify the actors behind the emergence of polarizing narratives. The study adopts a qualitative methodology using content analysis in the first phase of analysis, and critical discourse analysis (CDA) in the second one. This allowed the research team to engage in a deeper reflection on the analyzed news items in line with the framing theory (Porismita, 2011). The sample was obtained from six different Spanish nationwide media outlets, selected based on four criteria (see Table 1): (a) ownership (public/private), in order to represent both models existing in the Spanish media landscape; (b) format (traditional/digital) to guarantee the representation of the digital media market; (c) editorial line (conservative/liberal) by selecting media outlets from the most conservative to the most liberal editorial lines; and (d) type of media (TV stations and newspapers), for representing the media with the largest readership and viewership in Spain.

Table 1. Media outlets description according to the selection criteria.

Media outlets	Ownership	Format	Editorial line	Type	Readership/Viewership*
*ABC***	Private	Traditional	Conservative	Newspaper	Fourth position
Antena 3	Private	Traditional	Soft conservative	TV station	First position
El Confidencial	Private	Digital	Conservative	Newspaper	First position***
elDiario.es	Private	Digital	Liberal	Newspaper	First position***
El País	Private	Traditional	Liberal	Newspaper	First position
RTVE	Public	Traditional	Soft liberal	TV station	Second position

Notes: * According to the 2021 KREAB Report (Cabirta et al., 2022); ** the oldest conservative daily newspaper in Spain; *** first positions as rigorous digital media outlets.

3.1. Sample and Row Material

One of the major challenges in obtaining the final sample was deciding on how to access news items and subsequently make a final selection. The research team opted for the Twitter accounts of all selected media to carry out a massive data download. Spanish media use their Twitter accounts for their agenda-setting and to reach their audiences directly via an alternative channel (Casero-Ripollés et al., 2022). Using the scrapping technique, which is the most adequate for managing a huge amount of data, all data from 12 Twitter accounts belonging to the selected media outlets were downloaded and stored during a period of nine months (July 2021 to March 2022). The selection of the chronological period was stablished in the context of the European project Mediatized EU. The accounts selected were: *El País* (@el_pais; @elpais_espana; @elpais_inter), *ABC* (@abc_mundo; @abc_es), *El Confidencial* (@ECInter; @elconfidencial), e*lDiario.es* (@eldiarioes), Antena 3 (@antena3int; @A3Noticias), and RTVE (@rtvenoticias; @telediario_tve). The scrapping technique was applied by using the web browser from NVivo software called NCapture that allowed a quick and easy capture of all tweets posted in the Twitter accounts of the selected media outlets.

The filtering process was carried out by using keywords previously agreed upon by the researchers and considering the research objectives. Table 2 shows the monthly filtered news items per media during the selected period. A final sample of 543 news items from July 2021 to March 2022 were selected for analysis.

3.2. Methodological Stages: Content Analysis and Critical Discourse Analysis

Once the final sample was obtained, all data stored in Excel sheets was downloaded to NVivo SQR Software in order to proceed with a dual-step analysis. The first stage consisted of content analysis, which according to Thayer et al. (2007), is more appropriate for communication research since it reveals hidden connections between concepts and relationships among ideas that might seem disconnected at a first glance. As in previous works related to the project (García-Carretero et al., 2022; Pedrero-Esteban et al., 2021), analytical standards were agreed upon in various working sessions. The codes developed and modified by a researcher were taken into account by the rest of team using the same operational definitions for each category to maximize coding neutrality. A code book was designed and followed by all researchers in order to ensure a rigorous analysis. The coding process related to the content analysis included two different steps:

1. Coding all news items based on attributes, defined as a set of characteristics describing the units of analysis of our research. The attributes were designed in a deductive way. The team relied on theoretical references to define the attributes of analysis, such as media, authorship, approach, stance toward the EU (pro/neutral/anti), journalistic genre, and semiotic elements used in each piece of information.
2. Assigning theme-based nodes for an in-depth content analysis. These nodes are understood as categories or containers generated while examining the selected data. As pointed by Guix (2008, p. 28), in the content analysis "we should establish the categories in an inductive way, i.e., by creating the various labels under which we will order our information as they emerge from the exploitation of our sources." In this study, two different categories were created: main actors involved in the analyzed items and dominant topics.

The content analysis provided us with quantitative data that allowed a second stage of analysis using CDA for an in-depth study of the identified polarizing narratives and discourses. As stated by Van-Dijk (2016, p. 204) this is not a methodological approach *per se* but "an analytical practice that can be found in all discourse studies areas…the CDA is defined as the study of discourses within an attitude." This practice allows researchers to analyze social problems, discursive power relations, and the relationship between the text analyzed and society,

Table 2. Sample description per media and month.

		News items selection per month									
Media outlets	No. tweets	Jul	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar	Total
ABC	27,708	7	6	5	11	15	8	7	14	9	82
Antena 3	18,325	5	1	5	9	9	8	6	7	11	61
El Confidencial	21,804	9	0	9	12	23	4	7	18	12	94
elDiario.es	24,902	8	7	11	17	28	14	10	18	17	130
El País	35,592	10	6	3	8	9	7	7	13	17	80
RTVE	34,613	8	7	7	13	11	12	10	17	11	96
Total	162,944	47	27	40	70	95	53	47	87	77	543

offering an important interpretative frame. As Fairclough (2003, p. 163) described it previously: "Explanation is a matter of seeing a discourse as part of processes of social struggle, within a matrix of relations of power." The adoption of this approach helps decoding meanings and connections implied in the analyzed news items. According to Berelson (1971), such analysis lies in the description of objective, systematic, and quantifiable manifest content. Lately, the development of critical discourse studies has legitimized this methodology, which has become a valid and rigorous approach to study the ways in which discourses and ideologies are disseminated across different channels of communication (Machin, 2013; Van Leeuwen, 2013).

Although CDA is a flexible analytical approach, it is essential that it is carried out in a standardized way. As López-Noguero (2002) points out, this standardization allows us to detect polarization in the discourses. To this end, the research team used the Memo functionality from NVivo SQR Software based on the creation of collaborative working spaces. Memos are defined as:

> A type of document that enable the team to record the ideas, insights, interpretations or growing understanding of the material in your project. They provide a way to keep your analysis separate from (but linked to) the material you are analysing. (NVivo, 2022)

Specifically, the research team used "memos for a code." This process was enabled by the previous content analysis phase. Once the memos were written, the team carried out a critical analysis of the discourses that contain a polarizing effect according to the "us vs. them" rhetoric and other communicative strategies (Van-Dijk, 2016). The analysis also considers Foucault's (1991) limits and forms of the sayable, conservation, reactivation, memory, and appropriation.

4. Results

4.1. Content Analysis: Topics and Actors in Spanish Media Outlets Discourses

The detection of main topics in media discourses was one of the main objectives of the first stage of analysis. In this regard, the following topics were detected in the news items discussing the EU. Each Spanish media outlet can show a higher or lower percentage of coverage of certain topics depending on its own media agenda. Figure 1 displays these topics per media outlet.

Our content analysis revealed that one area in which polarizing narratives are more prevalent is EU politics. Our codification also introduced (soft) Euroscepticism as a sub-category. Table 3 shows the percentage of both categories, EU Politics and Euroscepticism represented

Figure 1. Main topics detected in the content analysis stage per media in percentages. Notes: For missing categories: Social relations–social problems (human rights, LGBTI, etc.); global politics (UK–Gibraltar, etc.); defense and cybersecurity issues; disinformation and fake news; Spain–Internal politics–national, regional, and local issues; EU countries–Internal politics–national, regional, and local issues.

Table 3. Presence percentage of topics EU Politics and Euroscepticism in Spanish media.

	Antena 3	*ABC*	*El Confidencial*	*El País*	*elDiario.es*	RTVE
EU Politics	3.14%	15.06%	20.05%	13.77%	41.56%	6.42%
Euroscepticism	7.41%	8.59%	20.21%	6.06%	54.97%	2.75%
		Right-wing			Left-wing	
EU Politics		38.25%			61.75%	
Euroscepticism		36.21%			63.78%	

in the Spanish media discourses. As shown in this data, *elDiario.es* is the Spanish media outlet that scored highest in both categories: EU Politics (41,56%) and Euroscepticism (54,97%). In comparison, the two TV stations scored lowest in these two categories. If we analyze this data from a political perspective, we realize that media outlets on the left give more importance in their agenda setting to EU politics (61.75%), and account for 63.78% of news items giving visibility to Euroscepticism, compared to right wing media with an average of 35.25% for EU Politics, and 36.21% for Euroscepticism.

Analyzing in depth (the CDA method) the discourses that emerge from the two above-mentioned categories, the selected media show a clear pro-European tendency in their discourses on EU Politics (62.90%) and illustrate the anti-EU discourses stemming from Eurosceptic parties within the EU (71.53%; see Figure 2).

Likewise, this research has detected the sources and actors most relied upon by the Spanish media when explaining the topics listed above. In this sense, representatives of EU institutions as well as representatives of the EU member states, including those of the Spanish government, are the main actors detected (see Figure 3).

4.2. Critical Content Analysis: Three Poles of Polarization in Spanish Media Outlets

4.2.1. North–South Divide

Fear of marginalization and the attempt to challenge power asymmetries between the North and the South within the EU is still vivid in media discourses on the EU, even when the EU itself is praised. Discourses on the Next Generation funds is a good example of it. The Next Generation funds are presented as a window of opportunity for the future of the Union, as reflected in statements such as "Europe in this last year has been a space for opportunities" (Gallardo, 2021a). The adoption of the funds is described as a "historic" decision (Gallardo, 2021a), and "without a doubt, a gigantic step" (Gallardo, 2021a). These funds are viewed as an opportunity for a deeper and more just economic integration, as stated in an item published by *elDiario.es*: "In this crisis, unlike in previous crises, the economic health of Southern Europe is to be prioritized" (Gil, 2021a). Even Von der Leyen's state of the union address, covered by the Spanish media, is based on this narrative: "A just recovery lies ahead, [one] that avoids social fracture and prepares us for future crises" (Gallardo, 2021d). This speech is described by a TV host as "moving," emphasizing that "we have heard the word soul on several occasions," and that many of its headlines "are called to improve people's lives" (Gallardo, 2021d).

However, the praise of the European funds is paralleled with a discourse that attempts to discredit previous fiscal measures advocated by Germany within the EU. The said measures exemplified the power asymmetries that exist between the North and the South in Europe. To challenge this asymmetry, different actors in the media resort to intertextuality, i.e., constructing the significance of the (historic) recovery funds by comparing them to the previous (notorious) austerity measures. This is reflected in the statement of a progressive member of the European Parliament (Garrido) on RTVE:

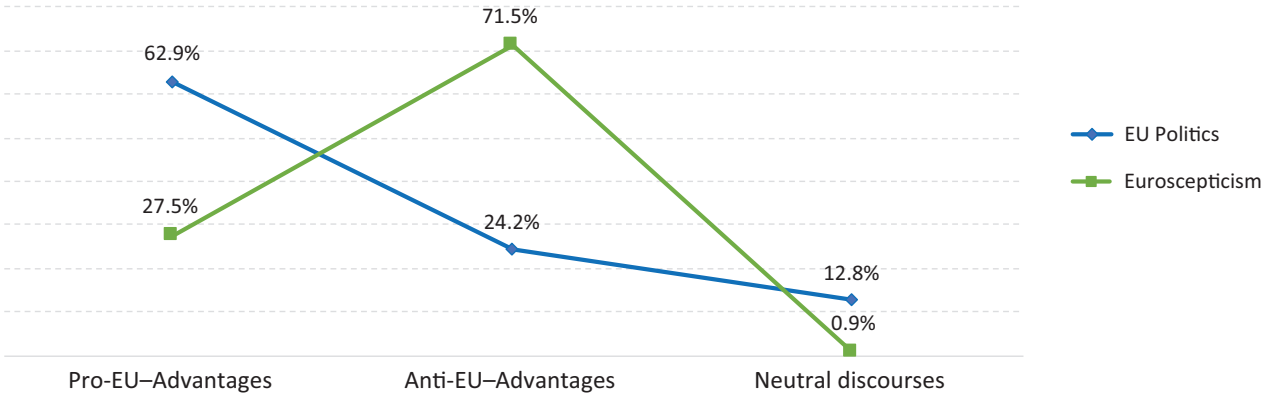


Figure 2. Pro, anti, and neutral discourses detected in the CDA stage within the topics EU Politics and Euroscepticism.

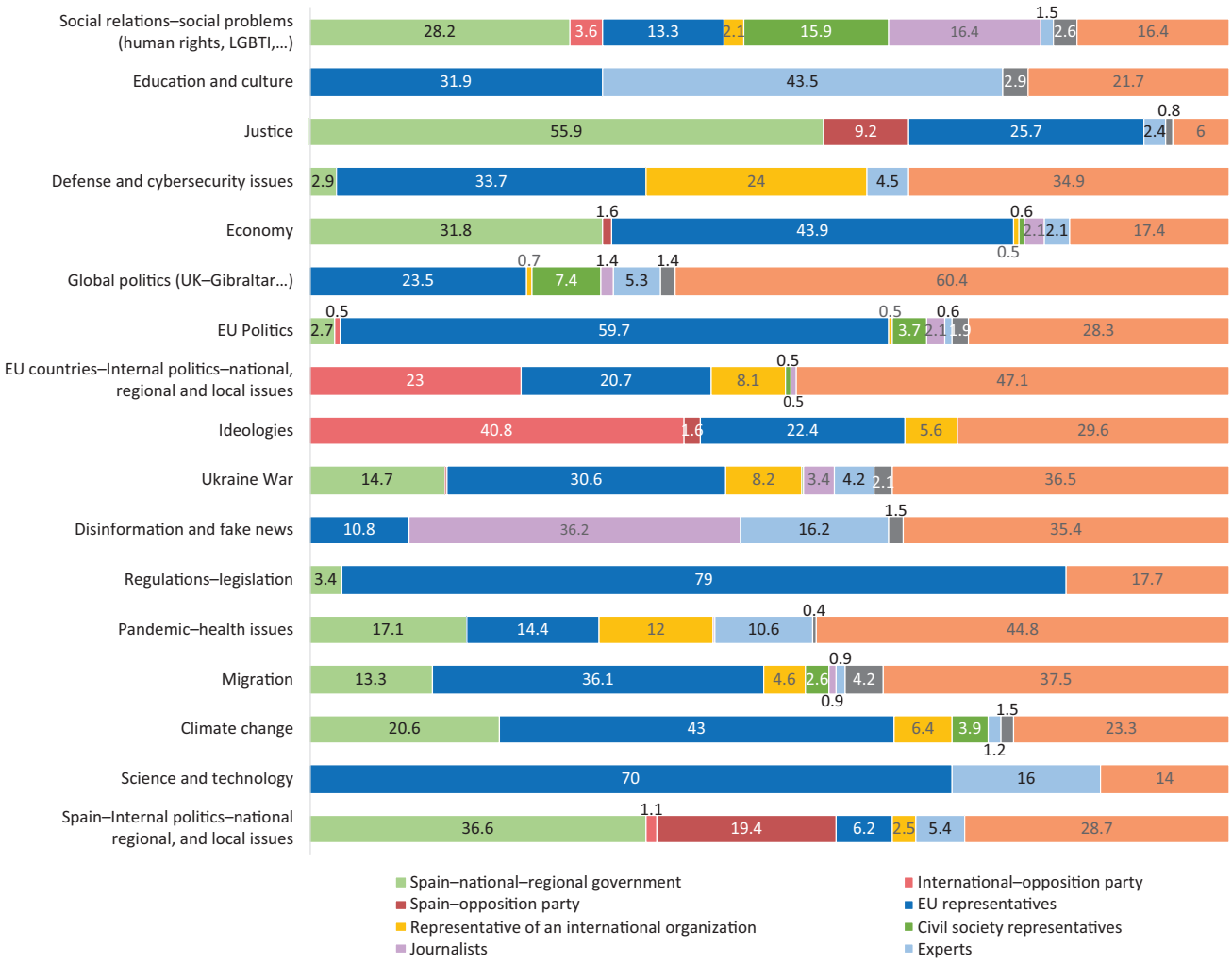


Figure 3. Main actors detected in the content analysis phase.

> The European Union has responded in a very different way than it did when the terrible financial crisis hit 10 years ago. [In] the great recession there was austerity, austerity, austerity. And now the disastrous [outcomes] of austerity are seen. Now the opposite has been done—investment—and the proof of that are the 750 billion euros that are going to be injected into the European economy for the recovery, [in accordance with] the famous recovery funds. (Gallardo, 2021d)

Previous measures, such as the Stability and Growth Pact, are criticized for being rigid and outdated with frugal states being portrayed as an obstacle, as reflected in an excerpt from *elDiario.es*: "But the governments of northern Europe, fiscally conservative, are reluctant to make efforts to substantially relax fiscal rules, compared to the positions of the southern countries" (Gil, 2021c). Another item in *elDiario.es* titled "Paradigm Change in Brussels?" highlights that "gone are the days of the Troika's diktat and of the men in black to access European money" (Gil, 2022a). The same piece continues as follows:

> The European Commission tries to escape from the long shadow of the Troika while enduring 800,000 million euros worth of debts to finance the recovery of the European Union that is struggling to get out of the coronavirus pandemic thanks to the vaccines negotiated and bought by Brussels for 27. Looking back in the mirror of a decade ago, Germany's resounding *nein* on eurobonds has turned into Brussels enduring the largest joint debt in its history. (Gil, 2022a)

Using the Foucauldian concept of memory (Foucault, 1991), which refers to the particular way in which we look at old discourses, such quotes suggest that Eurosceptic discourses on the austerity measures, adopted more than a decade ago, are engraved in public memory albeit dormant. The new measures adopted by the EU only reaffirmed the validity of such critical old discourses that symbolize a highly polarizing moment in the history of the EU, leading to the emergence of new critical approaches to Europeanization such as the ones voiced by Podemos against the austerity measures.

4.2.2. East–West Divide

The EU's decision to withhold funds from Hungary and Poland due to human rights and rule of law concerns gives rise to the discourse that human rights and the rule of law are central components of the European integration process and of Europe's identity. Therefore, the EU must act firmly to protect them. Conditioning the reception of the recovery funds with respecting these values is endorsed by media discourses across the board. These values become the site of contestation between the EU, fully supported by Spain, and states that argue that their sovereignty trumps European normative standards. This polarization is presented by the Spanish media as a conflict over the identity of Europe. Europeans (or true Europeans) are portrayed as strong supporters of these values. For example, in its item "The Europeans want the funds to go only to countries that respect the rule of law," the ABC covered the Eurobarometer on the state of the union and reported that the EU citizens view the rule of law and democratic values as key values that should be respected by member states. Furthermore, European citizens support withholding funds from governments that fail to respect European values ("Los europeos quieren," 2021). Statements by Spain's political elites show full support of the EU withholding funds from Hungary and Poland. An example of this explicit support could be found in the following statement by IU:

> The discriminatory law against the LGTBQ community passed by the Hungarian government requires a firm reaction, and I am glad to have listened today to Mr. Michel and Mrs. Von der Leyen's…clear position in defense of the rule of law. (RTVE Noticias, 2021)

Likewise, statements by the political forces from the right display endorsement of the EU position:

> The European PP defends the rule of law among the 27 member states, and this is an immovable principle for us. That is why we were one of the signatories to the call for European funds to be conditional on compliance with the rule of law by all parties. (Gallardo, 2021c)

The only party that objects the sanction imposed on Hungary and Poland is VOX. This is reflected, *inter alia*, in the statement of Jorge Buxadé, vice-president of the party: "We have not created the EU to subject the governments that have been legitimately appointed by their people to the path that it decides" (Gallardo, 2021c). However, even the conservative media criticized VOX's positions on such issues. The piece "The cavern of Santiago Abascal" is one example of that. The item criticized Abascal's anti-European stance ("Abascal doesn't like the euro or Europe") and his position on key human rights issues:

> [VOX] poisons society with the sinister choreography of xenophobia, machismo, nationalism, anti-Europeanism, homophobia, obscurantism, and resistance to the evidence of climate change. Nothing better than Abascal's messianism to excite the instincts and stimulate the emotions….He would like Spain to be the Hungary of Orbán, the Poland of Law and Justice. (Amón, 2021)

These discourses also create the dichotomy of "us vs. them," with "us" being true Europeanists who believe in human rights values, such as the case of Spain, and "them" the ultra-nationalists who want to access "all the benefits of being a member of the European Union, including resilience and recovery funds, without fulfilling any of the obligations and duties, starting with respect for European law that all the rest of us comply with" (Gallardo, 2021e). Supporting LGBTQ rights becomes a key distinguishing criterion between "us" vs. "them." This is exemplified in headlines, such as "Brussels squeezes Hungary and Poland with the revision of European funds for their authoritarian and homophobic drift" (Gil, 2021b); likewise, the Hungarian prime minister, Viktor Orbán, is described as "ultra-conservative," for linking homosexuality with pederasty: "The Hungarian government is absorbed in its homophobic drift" (RTVE Noticias, 2021). One item on RTVE even compares Orbán's censorship of books on LGBTQ rights to policies of Nazi Germany:

> In Hungary, they want books to be marked with a label. It is very similar to what the Nazis did with degenerate art, etc., and it is a super harmful thing to know that these are not just stories for LGBT families. They are stories for all audiences that simply show an LGBT reality. (Gallardo, 2021b)

4.2.3. Internal Divide

While the Spanish political landscape remains highly pro-European, Europeanization and EU policies can be used as a leverage to feed internal political polarization. In other words, it is not the EU itself that is a subject of contestation, it is the way EU related policies are implemented or promoted by rival political parties. Regarding this internal polarization, two main issues were detected: macro-farms and discourses on the renewal of General Council of the Judiciary (Consejo General del Poder Judicial [CGPJ]). Both issues represent polarized discourses across the right vs. left division, and between the same political camp (left vs. radical left).

In January 2022, a controversy emerged in the Spanish media over a statement by Spain's Minister for Consumer Affairs, Alberto Garzón, on macro-farms in an interview to *The Guardian*. In it, Garzón stated that meat produced in macro-farms was of worse quality than meat produced in extensive livestock production systems, in addition to being more harmful to the environment. This statement was interpreted by his political rivals

as downgrading the quality of Spanish meat. The fact that European Commissioner for Agriculture Janusz Wojciechowski supported Garzón´s position added fuel to the polarizing effect of the statement. This issue is of relevance as it not only divides the government and the political opposition, but also the governmental coalition itself. Some of the most significant headlines were: "Garzón reaffirms himself after Moncloa's disapproval: 'What I said is impeccable'" or "Evil Garzón and green nuclear," "Unidas Podemos questions Sanchez's 'loyalty' for 'giving fuel to a right-wing hoax' against Garzón," or "Garzón says the controversy over his statements is a 'use of disinformation' to wear him down." The controversy reaches the EU, and its officials end up defending the Spanish minister on what had happened:

> The minister for consumer affairs never said in the interview that Spanish meat was of poor quality or of worse quality than that of the rest of the European partners. What he did say was that the production of macro-farms—applicable anywhere in the world—is of poorer quality than that of extensive production. (Gil, 2022b)

Another issue where EU standards were utilized by the opposition to criticize the government on a highly polarized issue in Spain is the renewal of the CGPJ. Beyond being a highly polarized political issue in Spain, the renewal of the CGPJ had been expressly requested by the EU: "Brussels urges the renewal of the Spanish judiciary and constitutional bodies" (Gil, 2021a). This is used by the opposition as an argument against the government, as reflected in the following statement:

> The European Commission says very clearly that there is a lack of independence of the figure of the attorney general in Spain….Europe has said it very clearly, and that is what the Partido Popular is in favor of. We are defending the same as the European Commission. (Gallardo, 2021c)

However, some media outlets criticize the polarizing discourses propagated by the opposition with the publication of EU report on the rule of law in Spain, reminding the former of its own role in the CGPJ crisis: "In effect, the PP did agree on the renewal of RTVE, but has refused to do the same with the judicial governing body or the Ombudsman, for example" (Gil, 2021a).

5. Conclusions

Our study on polarization in media discourses related to Europeanization reveals that even in contexts where the EU is strongly supported by political and media elites, polarization can still emerge. In such contexts, the aim of polarization is not to discredit the EU and its institutions, but to bolster certain perceptions and visions on the construction of the EU and its identity. Polarizing discourses can also be utilized to challenge power asymmetries between member states or to reinforce them. In both cases, the aim is to improve the position of the state within the EU. This type of inter-state polarization was detected by our study in relation to fiscal measures adopted by the EU, and in relation to human rights and the rule of law issues. In relation to fiscal policies, the North–South divide is contested in media discourses by all Spanish actors. Resort to intertextuality to praise the Next Generation funds by discrediting old austerity measures is a strategy that aims to challenge the hegemony of Northern States. In relation to the promotion of human rights and the rule of law, the construction of the East–West divide also serves as a tool to promote Spain's vision of the EU and its identity and values. It also benefits Spain by positioning it in the group of states that genuinely represent European values. A second type of polarization is internal one, but it does not affect the EU itself; instead, it uses the EU and its policies as a leverage to discredit political rivals.

The polarizing narratives detected by our study reflect the Eurosceptic discourses that surfaced with the emergence of Podemos and VOX. In relation to Podemos, it seems that discourses on the austerity measures and the German influence are still vivid in the public memory and were recently validated, as reflected in the discourses on the Next Generation funds. As for VOX, its position on sanctioning Poland and Hungary for their human rights and rule of law record remains consistent with its European electoral program that calls for the primacy of national laws over EU law. Its position is also consistent with its criticism of LGBTQ rights.

The novelty brought by our study is conditioned by two factors. Firstly, by the end of bipartidism in Spain and the emergence of two opposite and radical political parties (Podemos and VOX). Both parties have adopted an incoherent stance towards the EU, amounting at times to Euroscepticism. It remains to be seen whether their future stance towards the EU will be more consistent and whether they are able to increase their role and influence in the political life of the country. Secondly, our study is conditioned by a specific post-Covid context and the beginning of a war in Ukraine (as the study covers the period between July 2021 and March 2022). This period was marked by two major crises that resulted in a higher visibility of the EU in the European media landscape and a more robust collective action at the EU level.

Although the concept of polarization is a complex one (DiMaggio et al., 1996; Kaufman & Haggard, 2019), an increasing number of authors point out to this phenomenon as the end of consensus (Magre et al., 2021; Pausch, 2021; Rodríguez-Virgili et al., 2022; Svolik, 2019). The analysis in this article yields interesting conclusions for academics, media outlets, and policy makers, which can be summarized as follows. First, while the Next Generation funds have improved the image of the EU and reflected a radically different response to the economic crisis triggered by the Covid-19 pandemic compared the

EU's response to the economic crisis of 2008, this by itself is insufficient to erase past experiences of peripheralization. This is evident in the continuous comparison between the (notorious) austerity measures, adopted during the eurozone crisis, and the new (historic) funds that are perceived as genuinely beneficial to all the members of the Union. This suggests that any future attempt to reintroduce rigid fiscal policies would be met with resistance and could lead to the surge of new wave of Euroscepticism, critical of the hegemony of the North and power asymmetries within the EU.

Second, polarization along the East–West axis could undermine the protection of human rights and the rule of law. When the contestation of European human rights and the rule of law standards results in the construction of a clear dichotomy between "us" vs. "them," the space for dialogue and consensus is eroded. It is worth noting that while the "us vs. them" dichotomy on this issue in Spain remains on the inter-state level, the fact that VOX objected the sanctions imposed on Poland and Hungary poses the risk that this polarization could become internal one.

As a third conclusion, Rodríguez-Virgili et al. (2022, p. 97) argue that after studying 40 years of polarization in Spain it seems "that we are at the moment of maximum polarization since the approval of the 1978 Constitution." It is not surprising then, that even if Europeanization is not a polarized political project in Spain, the EU can be used as a leverage to discredit political rivals. This was evident in discourses on the controversy over macro-farms and also in discourses on the failure to renew the CGPJ. Beyond fueling internal polarization, this could risk turning the European integration into a polarized political project.

While our analysis is based on a limited timeframe (July 2021 to March 2022), it offers evidence of a growing and worrying phenomenon, whose impact is seen not only in the political or the media sphere, but it is reinforced by technological disruption and the influence of social networks.

Acknowledgments

This research study is part of the international project financed by the Horizon 2020 program "Mediatized Discourses on Europeanization and Their Representations in Public Perceptions" (Mediatized EU Grant Agreement No.101004534, H2020–SC6–Transformations–2020), funded by the European Commission from January 1, 2021, to December 31, 2024.

Conflict of Interests

The authors declare no conflict of interests.

References

Amón, R. (2021, October 11). La caverna de Santiago Abascal [The cave of Santiago Abascal]. *El Confidencial*. https://bit.ly/3oybAEr

Arbatli, E., & Rosenberg, D. (2021). United we stand, divided we rule: How political polarization erodes democracy. *Democratization*, *28*(2), 285–307. https://doi.org/10.1080/13510347.2020.1818068

Arregui, J. (2022). *Europeanisation of public policies in Spain: Opportunities and challenges*. McGraw Hill.

Arrese, Á., & Vara-Miguel, A. (2016). A comparative study of metaphors in press reporting of the euro crisis. *Discourse & Society*, *27*(2), 133–155. https://doi.org/10.1177/0957926515611552

Avilés, J. (2004). España y la integración europea: Partidos y opinión pública (1977–2004) [Spain and European integration: Parties and public opinion (1977–2004)]. *Espacio Tiempo y Forma: Serie V, Historia Contemporánea*, *2004*(16). https://doi.org/10.5944/etfv.16.2004.3086

Berelson, B. (1971). *Content analysis in communication research*. Hafner Publisher.

Berganza, R. (2009). Framing the European Union and building the media agenda: The 2004 European parliamentary elections in the Spanish daily press. *Journal of Political Marketing*, *8*(1), 59–69. https://doi.org/10.1080/15377850802605973

Carothers, T., & O'Donohue, A. (Eds.). (2019). *Democracies divided: The global challenge of political polarization*. Brookings Institution Press.

Casal-Bértoa, F., & Rama, J. (2021). Polarization: What do we know and what can we do about it? *Frontiers in Political Science*, *3*, Article 687695. https://doi.org/10.3389/fpos.2021.687695

Casero-Ripollés, A., Alonso-Muñoz, L., & Marcos-García, S. (2022). The influence of political actors in the digital public debate on Twitter about the negotiations for the formation of the government in Spain. *American Behavioral Scientist*, *66*(3), 307–322. https://doi.org/10.1177/00027642211003159

Cabirta, A., San José, A., & Zamorano, P. (2022). *KREAB Report: Los medios de comunicación en 2021* [KREAB Report: Media outlets in 2021]. KREAB. https://kreab.com/espana/wp-content/uploads/sites/3/2022/02/2022-informe-kreab-medios.pdf

de Albuquerque, A. (2018). Political parallelism. In J. F. Nussbaum (Ed.), *Oxford research encyclopedia of communication*. https://doi.org/10.1093/acrefore/9780190228613.013.860

della Porta, D., Kouki, H., & Fernández, J. (2017). Left's love and hate for Europe: Syriza, Podemos and critical visions of Europe during the crisis. In M. Caiani & S. Guerra (Eds.), *Euroscepticism, democracy and the media* (pp. 219–240). Palgrave Macmillan.

DiMaggio, P., Evans, J., & Bryson, B. (1996). Have American's social attitudes become more polarized? *American Journal of Sociology*, *102*(3), 690–755. https://www.jstor.org/stable/2782461

Dutceac, A., & Bossetta, M. (2019). Can Euroscepticism contribute to a European public sphere? The Europeanization of media discourses on Euroscepticism

across six countries. *Journal of Common Market Studies*, *57*(5), 1051–1070. https://doi.org/10.1111/jcms.12871

Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. Routledge.

Feenstra, R. A., Tormey, S., Casero-Ripollés, A., & Keane, J. (2017). *Refiguring democracy: The Spanish political laboratory*. Routledge.

Fernandez, A. (2014, May 22). *Acto de Podemos en Sevilla—Intervención de Pablo Iglesias* [Podemos evento in Seville – Pablo Iglesias' speech] [Video]. YouTube. https://www.youtube.com/watch?v=Uw7_GrvxBPk

Foucault, M. (1991). Politics and the study of discourse. In G. Burchell, C. Gordon, & P. Miller (Eds.), *The Foucault effect: Studies in governmentality* (pp. 53–72). The University of Chicago Press.

García-Carretero, L., Establés, M.-J., Pérez-Escoda, A., & Boulos, S. (2022). The representation of the European Union in the Spanish media: Dominant themes and actors on the Europeanisation process. *Estudios Sobre el Mensaje Periodístico*, *28*(4), 803–815. https://doi.org/10.5209/esmp. 82833

Gil, A. (2021a, July 20). Bruselas urge a la renovación de la cúpula judicial y de los órganos constitucionales españoles [Brussels urges renewal of Spanish judiciary and constitutional institutions]. *elDiario.es*. https://bit.ly/3zdt51z

Gil, A. (2021b, July 30). Bruselas aprieta a Hungría y Polonia con los fondos europeos por su deriva autoritaria y homófoba [Brussels squeezes Hungary and Poland with EU funds for their authoritarian and homophobic drift]. *elDiario.es*. https://bit.ly/3Gsxamg

Gil, A. (2021c, August 2). Los países del Norte de Europa agitan los fantasmas de la austeridad mientras el Sur busca nuevas reglas económicas [Northern European countries stir up the ghosts of austerity as the South seeks new economic rules]. *elDiario.es*. https://bit.ly/3zddkaX

Gil, A. (2022a, January 2). Fondos europeos, deuda común y compra centralizada de vacunas: ¿Cambio de paradigma en Bruselas? [European funds, common debt, and centralised vaccine procurement: A paradigm shift in Brussels?] *elDiario.es*. https://bit.ly/3cLSKHb

Gil, A. (2022b, January 17). Planas critica en Bruselas el debate sobre las macrogranjas: "No es creíble, está lejos de la realidad" [Planas in Brussels criticises the debate on macro-farms: "It's not credible, it's far from reality"]. *elDiario.es*. http://bit.ly/3QTHUPe

Guix, J. O. (2008). El análisis de contenidos: ¿Qué nos están diciendo? [Content analysis: What is it telling us?]. *Revista de Calidad Asistencial*, *23*(1), 26–30. https://doi.org/10.1016/S1134-282X(08)70464-0

Iglesias, P. (2015). Spain on edge. *New Left Review*, *93*, 23–42. https://newleftreview.org/issues/ii93/articles/pablo-iglesias-spain-on-edge

Kaufman, R. R., & Haggard, S. (2019). Democratic decline in the United States: What can we learn from middle-income backsliding? *Perspectives on Politics*, *17*(2), 417–432. https://doi.org/10.1017/S1537592718003377

Lerín, D. (2022, September 7). *Changes in VOX's political strategy after its electoral breakthrough: Populism and the search for the working-class vote* [Paper presentation]. XIV Congreso Español de Ciencia Política y de la Administración, Girona, Spain.

López-Noguero, F. (2002). El análisis de contenido como método de investigación [Content analysis as a research method]. *XXI Revista de Educación*, *4*, 167–179. http://hdl.handle.net/10272/1912

Los europeos quieren que los fondos vayan sólo a países que respeten el estado de derecho [Europeans want funds to go only to countries that respect the rule of law]. (2021, September 22). *ABC Internacional*. https://bit.ly/3b9F6NI

Machin, D. (2013). What is multimodal critical discourse studies? *Critical Discourse Studies*, *10*(4), 347–355. https://doi.org/10.1080/17405904.2013.813770

Magre, J., Medir, L., & Pano, E. (2021). Spain: Is ideology back in populist discourse? In C. Bobba & N. Hubé (Eds.), *Populism and the politicization of the Covid-19 crisis in Europe* (pp. 31–44). Palgrave Macmillan.

Menéndez, V. M. (2010). Media representation of the European Union: Comparing newspaper coverage in France, Spain and the United Kingdom. *International Journal of Communication*, *4*(18), 398–415. https://digitalcommons.butler.edu/cgi/viewcontent.cgi?article=1458&context=facsch_papers

NVivo. (2022). *Manual coding techniques*. https://help-nv.qsrinternational.com/20/win/Content/coding/coding-techniques.htm

Pausch, M. (2021). The future of polarisation in Europe: Relative cosmopolitanism and democracy. *European Journal of Futures Research*, *9*(1), Article 12. https://doi.org/10.1186/s40309-021-00183-2

Pavan, E. (2017). The integrative power of online collective action networks beyond protest: Exploring social media use in process of institutionalization. *Social Movement Studies*, *16*(4), 433–446. https://doi.org/10.1080/14742837.2016.1268956

Pedrero-Esteban, L.-M., Pérez-Escoda, A., & Establés, M.-J. (2021). Brexit's impact on Spanish media discourse: News coverage and audience reaction on Twitter. *Profesional de la Información*, *30*(6). https://doi.org/10.3145/epi.2021.nov.10

Pérez-Escoda, A. (2022). Infodemic and fake news turning shift for media: Distrust among university students. *Information*, *13*(11), Article 523. https://doi.org/10.3390/info13110523

Podemos. (2014). *Mover ficha: Convertir la indignación en cambio político* [Making a move: Turning outrage into political change]. https://bit.ly/Moverficha

Porismita, B. (2011). Conceptual issues in framing theory: A systematic examination of a decade's literature. *Journal of Communication*, *61*(2), 246–263. https://

doi.org/10.1111/j.1460-2466.2011.01539.x
Powell, C. (2007). Spanish membership of the European Union revisited.*South European Society and Politics*, *8*(1/2), 147–168. https://doi.org/10.1080/13608740808539647
Real-Dato, J., & Sojka, A. (2020). The rise of (faulty) Euroscepticism? The impact of a decade of crises in Spain. *South European Society and Politics*. Advance online publication. https://doi.org/10.1080/13608746.2020.1771876
Rivas-de-Roca, R., & García-Gordillo, M. (2023). Local information and the European Union: Analysis of media and institutional communication management. *Vivat Academia*, *156*, 1–23. https://doi.org/10.15178/va.2023.156.e1439
Roch, J. (2019). Unpacking EU contestation: Europeanization and critique in Germany and Spain. *Culture*,*4*(2), 13–36. https://doi.org/10.5771/2566-7742-2019-2-13
Rodríguez-Virgili, J., Portilla-Manjón, I., & Sierra-Iso, A. (2022). Forty years of ideological polarisation in Spain. *Revista Empresa y Humanismo*, *XXV*(2), 75–105. https://doi.org/10.15581/015.XXV.2.75-103
Gallardo, J. C. (Executive Producer). (2021a, July 16). *Europa 2021* [Europe 2021] [News program]. RTVE. https://bit.ly/3b8Ozoi
Gallardo, J. C. (Executive Producer). (2021b, September 14). *Europa 2021* [Europe 2021] [News program]. RTVE. https://bit.ly/3PHE4ra
Gallardo, J. C. (Executive Producer). (2021c, September 15). *Europa 2021* [Europe 2021] [Program type]. RTVE. https://bit.ly/3zzFq1r
Gallardo, J. C. (Executive Producer). (2021d, September 17). *Europa 2021* [Europe 2021] [Program type]. RTVE. https://bit.ly/3PHJbYh
Gallardo, J. C. (Executive Producer). (2021e, October 22). *Europa 2021* [Europe 2021] [Program type]. RTVE. https://bit.ly/3zBBRI2
RTVE Noticias. (2021, July 22). *Hungría: Orbán convoca un referendum sobre su polémica ley LGTBI* [Hungary: Orbán calls referendum on his controversial LGTBI law] [Video]. YouTube. https://bit.ly/3PWKvGw
Ruiz, A. M., & Egea, A. (2011). Spain: Euroscepticism in a pro-European country? *South European Society and Politics*, *15*(1), 121–147.
Sojka, A., & Vázquez, R. (2014). European identity in Southern Europe: A comparative analysis of elite perceptions and public opinion in Spain and Portugal.*Revista Española de Ciencia Política*, *2014*(36), 89–114.
Sotelo, J. (2009). The role of the media in European identity formation. In J. J. Fernández Sanz, C. Sanz Establés, & Á. L. Rubio Moraga (Eds.), *Prensa y periodismo especializado IV* (pp. 139–148). Editores del Henares.
Svolik, M. W. (2019). Polarization versus democracy. *Journal of Democracy*, *30*(3), 20–32.
Taggart, P., & Szczerbiak, A. (2008). Introduction: Opposing Europe? The politics of Euroscepticism in Europe. In C. Taggart & A. Szczerbiak (Eds.), *The comparative party politics of Euroscepticism* (pp. 1–15). Oxford University Press.
Teruel-Rodríguez, L. (2016). El impacto de la crisis política y económica sobre la polarización de los medios españoles [The impact of political and economic crisis on the polarization of the Spanish media]. *Historia y Comunicación Social*, *21*(1), 203–220. https://doi.org/10.5209/rev_HICS.2016.v21.n1.52692
Thayer, A., Evans, M., McBride, A., Queen, M., & Spyridakis, J. (2007). Content analysis as a best practice in technical communication research. *Journal of Technical Writing and Communication*, *37*(3), 267–279. https://doi.org/10.2190/TW.37.3.c
Van-Dijk, T. A. (2016). Análisis crítico del discurso [Critical discourse analysis]. *Revista Austral de Ciencias Sociales*, *2016*(30), 203–222. https://doi.org/10.4206/rev.austral.cienc.soc.2016.n30-10
Van Leeuwen, T. (2013). Critical analysis of multimodal discourse. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 4002–4006). Blackwell.
Vázquez, R., Delgado, S., & Sojka, A. (2014). Spain. In N. Conti (Ed.), *Party attitudes towards the EU in the member states: Parties for Europe, parties against Europe* (pp. 115–130). Routledge.
VOX. (2014). *Programa electoral para las elecciones europeas* [Electoral program for the European elections].
VOX. (2019). *Programa electoral para las elecciones europeas de 2019* [Electoral program for the European elections 2019]. https://bit.ly/3X6rEwi

About the Authors

Ana Pérez-Escoda is an associate professor of digital communication at the Faculty of Communication and Arts and coordinator of the PhD program Innovation in Digital Communication and Media from the Antonio de Nebrija University (Spain). Her expertise is related to digital communication, digital and media literacy, digital skills, social media, and innovation. She is also a member of the INNOMEDIA research group and an active research member in several competitive research projects, national and international. Currently, she is part of the ongoing project Mediatized EU: Mediatized Discourses on Europeanization and Their Representations in Public Perceptions, funded by the EU H2020 Research and Innovation Program. Publications: https://scholar.google.es/citations?user=vo89GUMAAAAJ&hl=es

Sonia Boulos holds a doctorate degree in juridical science from the University of Notre Dame (US). She is a professor of international law at the Faculty of Law and International Relations at Nebrija University. Her research interests focus on international protection of human rights. Currently, she is the principal investigator of Nebrija's research team in the European research project Mediatized EU: Mediatized Discourses on Europeanization and Their Representations in Public Perceptions, funded by the EU H2020 Research and Innovation program.

María-José Establés holds a PhD in communication (Pompeu Fabra University). She is a Juan de la Cierva postdoctoral fellow at Complutense University of Madrid (Spain) funded by the Spanish Ministry of Science and Innovation. Her main areas of research are television series and fan communities with a special interest on sociopolitical issues. Her work has been published in various indexed journals such as *Learning, Media & Technology*, *El Profesional de la Información*, *Estudios Sobre el Mensaje Periodístico*, and *Historia y Comunicación Social*, as well as in prestigious publishing houses such as Routledge, Thomson Reuters Aranzadi, and Tirant lo Blanch.

Lucía García-Carretero is a postdoc researcher in the research project Mediatized EU, funded by the Horizon 2020 program of the European Commission. She is part INNOMEDIA research group at Nebrija University. She holds an international PhD in communication from the Pompeu Fabra University and has published in various indexed journals such as *Media and Communication*, *Profesional de la Información*, and *Estudios Sobre el Mensaje Periodístico*. Her main research lines are the analysis of online political communication, social media, and digital networks. ORCID: 0000–0002-1414-3921

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 235–248
https://doi.org/10.17645/pag.v11i2.6391

Article

Twitting Against the Enemy: Populist Radical Right Parties Discourse Against the (Political) "Other"

Laura Cervi [1,*], Santiago Tejedor [1], and Mónica Gracia Villar [2]

[1] Department of Journalism and Communication Sciences, Autonomous University of Barcelona, Spain
[2] Faculty of Social Sciences and Humanities, Universidad Europea del Atlántico, Spain

* Corresponding author (laura.cervi@uab.cat)

Submitted: 30 October 2022 | Accepted: 6 March 2023 | Published: 17 May 2023

Abstract
There is a common agreement in considering populism as a Manichean worldview that oversimplifies and polarizes political options reducing them to a symbolical struggle between an "us" and a "them." "Us" is embodied by "the people," equated with "good," and "them" is identified by political "Others," often embodied by "the elites" who are depicted as inherently "evil." Naturally, the nature and composition of the people and the elite vary according to both ideology and political opportunities. This article examines the discursive construction of political opponents in two populist radical right parties: Lega in Italy and Vox in Spain. Based on the analysis of a selection of tweets by the two party leaders, Santiago Abascal and Matteo Salvini, this study applies clause-based semantic text analysis to detect the main discursive representations of political opponents. The article concludes that Salvini focuses all the attention on the left, while Abascal, although predominantly identifying the left as the main enemy, also targets pro-independence parties. The discursive construction of the "enemy" is based on two main strategies: demonization, the framing of opponents as "enemies of the people" who, along with dangerous "Others" such as immigrants, conspire against the "people" and are blamed for everything that is "wrong" in society; secondly, character assassination of individual politicians through personal attacks, which aim to undermine their reputation and deflect attention from the real issues towards their personal traits and actions.

Keywords
character assassination; demonization; Italy; Lega; political discourse; populism; populist radical right; social media; Spain; Twitter; Vox

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

Populist radical right (PRR) parties are on the rise in Western Europe (Zulianello, 2020). Their emergence and electoral success have generated considerable scholarly research (Ostiguy et al., 2020) that has dug into their distinctive elements (Hawkins et al., 2012).

In particular, there is wide agreement on considering populism as a Manichean worldview that oversimplifies and polarizes political options reducing them to a symbolical struggle between an "us" and a "them" (Mudde & Kaltwasser, 2017). Accordingly, the populist construction of both "the people" and its "Others" has been the focus of comprehensive literature (Betz, 2017). PRR parties' xenophobic and anti-immigration stances have been deeply scrutinized (Cervi et al., 2020).

Nonetheless, studies have so far devoted little attention to showing how these parties discursively address their political opponents (Van Kessel & Castelein, 2016). Although this aspect has been tangentially touched on by previous works (Capdevila et al., 2022; Cervi & Carrillo-Andrade, 2019; Marcos-Marne et al., 2021),

mainly focusing on the anti-elitist aspect of populism (March, 2017), very few studies have dedicated exclusive attention to the construction of political opponents (Berti & Loner, 2021).

Therefore, this article aims to make an empirical contribution to the current debate about PPR parties disclosing how political opponents are constructed in the discourse of contemporary PRR parties. To do so, we study the discourse of two PRR parties in Southern Europe: Lega in Italy and Vox in Spain. After providing an overview of PRR parties and their discourse and the role social media plays in their communication, the article applies clause-based semantic text analysis (CBSTA) to a dataset of tweets by the two party leaders, Matteo Salvini and Santiago Abascal, concluding that both politicians identify the main political opponent as "the left" who are stigmatized through the use of two main rhetorical devices: demonization and character assassination.

2. Populism

While populism is one of the trendiest research topics in contemporary literature (Ekström et al., 2018), it is also one of the most contested concepts in the field of political science (Kefford et al., 2022).

Most literature single out the existence of three main conceptual approaches to studying populism (Kaltwasser et al., 2017): the ideational approach, the political-strategic approach, and the sociocultural or communicative/performative approach.

Arguably, the dominant approach today is the ideational approach, defining populism as a "thin-centered" ideology that considers society to be ultimately divided into two antagonistic and homogenous groups—"the pure people" and "the corrupt elite"—and that politics should be an expression of the *volonté générale* (Mudde, 2004, p. 543). This conception suggests that populism does not offer a complete worldview and fails to exhibit the degree of consistency, depth, and scope of other fully developed, "thick" ideologies such as socialism and liberalism.

This approach entails the main benefit of disclosing why populist parties are so varied and flexible regarding their programs, organization, and leadership and provides the possibility to connect the supply and demand sides of populism (Mudde & Kaltwasser, 2012, p. 10).

The political-strategic approach, mainly represented by Weyland (2001, p. 14), states that populism can be defined as a "political strategy through which personalistic leaders seek or exercise government power based on direct, un-mediated, un-institutionalized support from large numbers of mostly un-organized followers."

The third strand gathers different approaches which, taken together, provide the conceptual tools for broadly making sense of populism as a particular mode of political communication (Moffitt, 2016).

Mostly following Laclau's (2005) seminal work on political logic, but also Wodak's (2015) historical approach to critical discourse analysis, discursive scholars have suggested shifting the focus of analysis away from ideologies to concentrate on how discourses are constructed. Populism is, therefore, viewed as a discourse that seeks to confront "the people" against "the elite," and studies examine the ways in which this construction takes place and how these signifiers play out to simplify the world's complexities (Stavrakakis et al., 2018) and trigger emotional reactions (Krzyżanowski & Ekström, 2022).

Other authors have centered their analysis on the non-verbal and stylistic aspects of the phenomenon. Ostiguy (2009, 2017), the main proponent of the sociocultural or performative approach, for instance, argues that populism should be seen as the "flaunting of the low" in politics, paying attention to language, body language, gestures, and ways of dressing. In line with this approach, Moffitt (2016) defined populism as a distinct "political style," a particular repertoire of mediated performance that includes the appeal to "the people" versus "the elite," "bad manners" as well as the performance of crisis.

We contend, together with Ekström et al. (2018) and Kefford et al. (2022), among others, that the above-mentioned approaches, especially the ideational and the discursive-performative approach, are not mutually exclusive; thus, there is "room for synergic and cumulative work" (Olivas, 2021, p. 834).

First, as Ostiguy (2017, p. 74) points out, there are clear connections between the "believe in" and the "act as," that is to say, between the ideological and the communicative/performative aspects of populism. In other words, although diverse definitions may differ on which requirements or sets of criteria to use, all these approaches coincide on a conceptual core of basic attributes associated with populism and its manifestations, such as the Manichean interpretation of politics, anti-elitism, and an idealized conception of the people.

Second, and most importantly, granting that populist ideas, as with any other ideas, need to be communicated to reach the audience and achieve the communicator's goals, disclosing the communicative tools used for spreading them should be just central (De Vreese et al., 2018) or at least a needed complement to the scrutiny of populist ideas. This growing recognition of the centrality of discourse has led many proponents of the ideational approach to use the term "discourse" and "ideology" interchangeably (Hawkins & Kaltwasser, 2017, p. 514).

Ergo, aligning with De Vreese et al. (2018), we combine Mudde's (2004) ideology-centered and Hawkins' (2010) discourse-centered understanding of populism, considering populism as a discursive manifestation of a thin-centered ideology. Accordingly, in our analysis, we will not only focus on disclosing the "set of basic assumptions about the world" contained in the populist message but on "the language that unwittingly expresses them" (Hawkins et al., 2012, p. 3). Conceiving populism as an ideology articulated discursively by political actors and, as such, an expression of political communication not

only allows bridging political science and communication studies' literature but also grants the chance to deepen the understanding of populism, expanding the analysis to take into consideration broader communicative and performative repertoires.

3. Populist Radical Right Parties and Their Discourse

According to Mudde's (2007) influential definition, three main features characterize the PRR party family: populism, nativism, and authoritarianism.

Populism has been defined in the previous section. However, it is important to stress that, as previously explained, due to the thinness of populist ideology, populist actors combine populism with one or more other ideologies, so-called "host ideologies" (Mudde & Kaltwasser, 2017, p. 21). While left-wing populists often incorporate some form of socialism into their ideology, nationalism, particularly in its exclusionary, nativist form, is the most common addition for right-wing populists.

Nativism is the belief that states should be inhabited exclusively by members of the native group (the nation) and that non-native elements (persons and ideas) threaten homogeneous nation-states (Mudde & Kaltwasser, 2017). Under a nativist notion of citizenship, populists' characteristic appeal to the people becomes an appeal to "our people," the "pure" people (Betz, 2017). Accordingly, PRR parties distinguish "the people" from the "Others," aliens who do not belong to "us" and are consequently considered enemies, accused of conspiring—together or with the direct or indirect help of the elite—against the people (Mudde & Kaltwasser, 2017). Exactly as for the elites, the selection of whom to identify as "Other" depends on the contextual discursive opportunities (Koopmans & Muis, 2009); however, a vast strand of literature has shown that in recent decades most right-wing populists in Western Europe have capitalized on the growing concerns on immigration, activating people's grievances against immigrants (Cervi & Tejedor, 2021; Betz, 2017) and ethnic minorities, (Cervi & Tejedor, 2020).

Authoritarianism (the belief in the value of obeying and valuing authority, granted that it is their own) requires the government to have a significant moral weight in citizens' freedoms and rights (Hooghe & Marks, 2018), favoring strict order and severe punishment for violations (Mudde, 2007). This characteristic perfectly matches another key feature of populism: anti-pluralism (Galston, 2017). Pluralism emphasizes the inevitability and desirability of differences in society, calling for institutions that protect minority rights and differences in the pursuit of a majority will. Thus, those who adhere to pluralism are normally inclined to think of popular sovereignty as a dynamic and open-ended process. Populism, on its side, treats it as the fixed, reified, and unified will of the people (Espejo, 2011), craving moral clarity. Therefore, whereas pluralism sees political relations as essentially those of cooperation and dialogue, populism sees a naturally antagonistic world, rejecting any form of difference (Mudde, 2007).

The label "radical," finally, refers to both the outspoken position at the far end of the political spectrum on issues related to immigration and ethnic diversity and the disruption of political norms, emerging from the rejection of pluralism.

This disruption becomes particularly evident observing populist leaders' discourses, positioned clearly on the "low" end of Ostiguy's (2009) low-high continuum, triggered through the continuous appeal to common sense and enacted using "bad manners" (Winberg, 2017). The frequent use of vulgar language, aimed at pulling away from the lexicon of mainstream politicians (Berti & Loner, 2021), in particular, has been considered a key element in fostering the perception of authenticity (Forchtner & Kølvraa, 2017), strengthening a sense of closeness to "the people" (Ostiguy et al., 2020), and generating hostility towards the elites.

Last but not least, as noted by Meny and Surel (2002, p. 17), "populism advocates the power of the people, yet it relies on the seduction by a charismatic leader." A recurrent populist trope, in fact, also entails the representation of a salvific leader who is at the same time a "man of the street," that is to say, "one of us" and the savior, the "champion of the people" (Bracciale & Martella, 2017).

4. Populist Radical Right Parties and Social Media

The effectiveness of PRR parties' communication would be incomprehensible without considering the impact of new technologies. Social media disintermediation provides direct communication with citizens, allowing populists to circumvent the journalistic gatekeepers that are often hostile to them (Groshek & Engelbert, 2012). In this way, they can capitalize, not only on the generalized mistrust in traditional politics but also the mistrust towards mainstream media (Fawzi, 2019), presenting themselves as authentic, thus closer to the "people."

In addition, social media's attention economy makes them the arena par excellence of emotionality (Hameleers et al., 2017). This feature inherently runs counter to the key traits of establishment politics, perfectly matching with populists' discursive dynamics (Gerbaudo, 2018), which spread a fragmented ideology (Engesser et al., 2016) characterized by emotional elements (Hameleers et al., 2017) and a simplified dichotomous vision of the world. Especially in right-wing populism (Hameleers, 2020), the use of anger and fear exacerbates the distance and antagonism between the people and the elites or between the people and such outgroups as migrants (Cervi, 2020a).

As such, populist messages entail a high potential for virality that allows them to acquire news value (Wodak, 2015); in other words, "the more provocative the message, the more traditional media might be compelled to turn it into news" (Berti & Loner, 2021, p. 5).

5. Methods

5.1. Selection of the Cases

Understanding the comparative method as a method for identifying and explaining similarities and differences between cases using common concepts, we compare *per genus et differentiam* (Sartori, 2005) two parties that belong to the same party family and share common features but operate within divergent political contexts.

Both Vox and Lega exemplify radical right-wing populism (Zulianello, 2020) and can be considered critical case studies in their social media communication. Lega, in particular Salvini's personal massive use of social media (Cervi, 2020b), has been identified as the core element of both the construction of his political and mediated persona (Bobba, 2018) and his electoral success (Diamanti & Pregliasco, 2019). The effective use of social networks has also been found to be key to Vox's mobilization strategy (Barrio et al., 2021, p. 246) and success (Capdevila et al., 2022).

In addition, the two countries share a similar fascist past. However, while Italy has been widely recognized as "the promised land of populism" (Tarchi, 2015, p. 1), Spain is witnessing a new—and, for many analysts, unexpected—rise of this form of politics (Marcos-Marne et al., 2021).

Last but not least, during the analyzed time frame, Vox was in the opposition and Spain was ruled by a left-wing coalition (Marcos-Marne et al., 2021), while Lega, despite disagreeing on many points, formally supported the technocratic government led by Mario Draghi (Garzia & Karremans, 2021). Therefore, we can expect that the two parties might show similar discursive strategies but also some contextual differences concerning both their political and discursive opportunities (Koopmans & Muis, 2009).

5.2. Data Collection and Annotation

Scientific literature has shown that Twitter is largely used by political elites for agenda-building purposes (Parmelee, 2013), especially by populist actors (Jacobs & Spierings, 2018). Accordingly, our empirical analysis is based on a Twitter dataset.

Recognizing the centrality of leaders in PRR parties (Meny & Surel, 2002), the official Twitter accounts of the two party leaders were selected for the analysis. Tweets published by the party leaders, Matteo Salvini and Santiago Abascal, were downloaded using the Twitter API from January to August 2022, excluding retweets. The total number of tweets collected was 1,901 for Salvini and 1,698 for Abascal. After downloading all the tweets, only those mentioning political opponents were selected.

Understanding that in the center of politics lies the competition for political power—intended as the ability to shape and control the content and direction of public policy (Stoppino, 2001)—by political opponents, we understand all those groups that compete in the respective electoral arena. Accordingly, political opponents were defined as political parties/groups officially recognized as being part of the electoral process that can support candidates for elections on a regular basis (Sartori, 2005). In addition, individual politicians competing in the same arena were also considered political opponents, acknowledging the growing personalization of politics (Garzia, 2011). Therefore, to be included in the sample, tweets had to mention Italian or Spanish political parties or individual politicians.

Due to the relatively small *n*, tweets were analyzed manually. First, we isolated all the tweets containing parties' names and personal names of politicians operating in each country.

In a second round, acknowledging that discursive practices happen within specific sociocultural contexts that require a deep understanding of both the textual and contextual facts (Ekström et al., 2018) and in affordance to the driven context that might embed latent messages, the rest of the tweets were analyzed through content analysis, in order to guarantee that all the tweets referring to parties and politicians using other wordings (nicknames, metaphors, indirect reference to current news, etc.) would be properly included in the sample.

The methodology implemented was created by the Populism Team to compile the Global Populism Database (Hawkins et al., 2019): Each tweet was double-coded in its original language by two authors who did not share their work with each other until it had been completed. Discrepancies were subjected to a reconciliation session to adjust criteria. The final Cohen's kappa inter-rater agreement was 0.97, showing nearly perfect agreement among the coders.

5.3. Data Analysis

After analyzing whom the two leaders identify as their political opponents, our main aim is to disclose the construction of actors. Accordingly, we consider Twitter's texts as narrative texts that tell a story made of actors. Narratives are the core mechanism of constructing reality at the sociocognitive level: According to Mayer (2014, pp. 66–71), by "translating experience into the code of story—with plot, and character, and meaning," it allows the incomprehensible to be transformed into something meaningful.

Accordingly, Franzosi's (2010) model of CBSTA was implemented. This model starts from the premise that any story, in any language, can be analyzed, taking into account the structural categories subject–verb–object (Aslanidis, 2018). Concretely, thus, CBSTA consists of extracting triplets formed by the elementary syntactic components of language: subject–verb–object. Triplets allow one to deconstruct and reconstruct a narrative into clusters (Popping & Roberts, 2014), allowing one to code not only the signifiers but their structure in a statement, unveiling the actions of political subjects, the objects

of those actions along with their positive and negative affect, and the combination between these elements (Aslanidis, 2018).

CBSTA, therefore, allows both quantitative and qualitative analysis: The quantitative dataset, composed of the retrieved triplets, shows word co-occurrences and can be analyzed by qualitatively observing the attributes of the actors and their actions, along with epithets and adjectives (Franzosi, 2010).

Textual analysis is widely recognized as one of the best techniques developed to measure the rhetoric of politicians (Hawkins et al., 2012); CBSTA, in particular, as suggested by Aslanidis (2018), and proven by different studies (Cervi, 2020b; Cervi et al., 2021), appears to be a particularly fitting instrument for measuring populist discourse.

Only written text was considered: All multimedia content (videos, images) was excluded from the sample. For each main actor, we selected the characteristic semantic triplets to establish the lexical universes built around each of the aforementioned actors. Consequently, we qualitatively observed adjectives, verbs, and objects to establish relationships between actors and consequently draw the frameworks of reference.

6. Results

6.1. Selection of the Opponents

The total number of tweets collected for Salvini is 1,901, of which more than half (956) contain mention of political opponents. As for Abascal, the result is even more overwhelming: 56% of the tweets contain mention of the political opponents.

Thus, it is possible to state that, coherently with the populist dichotomous vision of the world, most of the tweets are dedicated to attacking the adversary.

It is also interesting to note that there is very little difference in the percentage, showing that although Lega in the analyzed timeframe formally supported the government, while Vox was in opposition, their behavior does not seem to change.

As per the selection of whom to target, Figure 1 shows that the majority of the references refer to individuals. In the case of Salvini, out of 956 tweets containing mention of political opponents, 587 contain personalized references; Abascal calls out individual politicians 800 times out of a total of 951 tweets.

In other words, both leaders personify their "enemies" (Garzia, 2011) by choosing specific individuals as targets. While Salvini mostly refers to the politicians by their names, Abascal tends to use nicknames or references to their political position (The Minister, etc.). In addition, it is possible to observe how official party names are less frequent, being the "Other" category most recurrent. In this category, we have collected all the mentions of political parties that do not contain their official names; rather, they are roughly or derogatorily identified by their ideological positioning: The Communist, *Los Progres* (ironic epithet to mock leftwing leaning individuals), etc.

Unsurprisingly, Salvini's tweets entirely refer to the left (100%). The only party mentioned is Partito Democratico, and the most recurrent definition is "the Left," defined by other adjectives that will be analyzed in the next session. This is due to the Italian party system being divided into two main blocks, the center-left led by Partito Democratico and the center-right, the coalition to which Lega belongs (Zulianello, 2020).

On the other hand, the Spanish political system is more complex since its multiparty system (Gray, 2020) is crossed by the traditional right–left axis and multiple territorial axes. Accordingly, Vox confronts both the left, mainly represented by the Socialist Party and

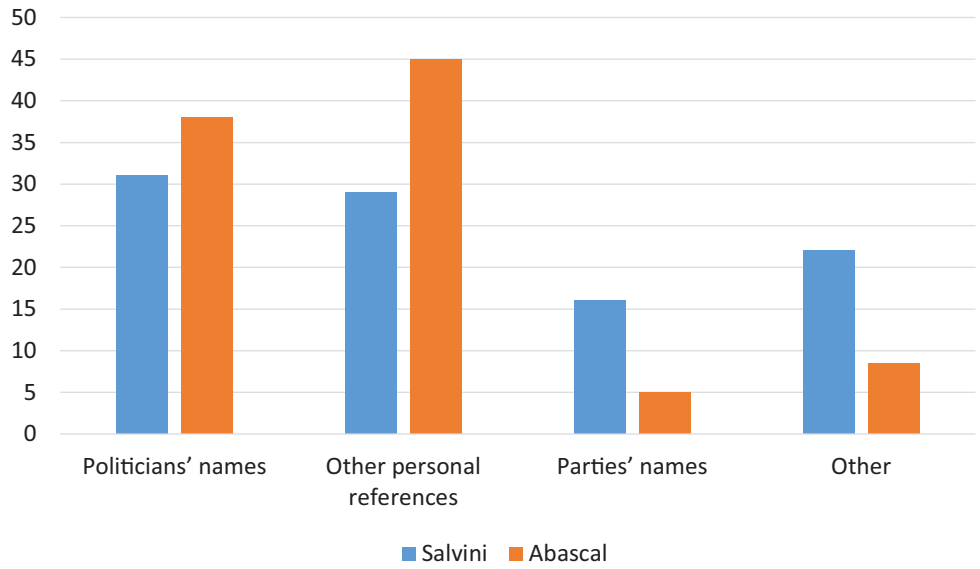


Figure 1. Distribution of the retrieved tweets.

Unidas Podemos (governing in coalition by the time of this research), and the center-right party, Partido Popular in the traditional axis, together with regionalists/separatist parties on the territorial axis (Fernández-García & Valencia Sáiz, 2022). However, Abascal refers to the Socialist Party and its partner in government, Unidas Podemos, in 82% of the tweets that mention political opponents, clearly defining them as the main opponent. In the rest of the tweets, he targets independentist parties (11%), and only 3,7% of the tweets mention the Popular Party. Both these actors are more frequently framed as "allies" of the left rather than subjects of the tweet. A residual 0,3% is dedicated to other parties (mainly regionalist parties from other areas of Spain). As per the general references, exactly as in Salvini's case, the most recurrent definition is "the Left," followed by mentions of Catalan independence parties.

6.2. Construction of the "Enemy": Demonization

As previously mentioned, both Salvini and Abascal mainly refer to political opponents as "the Left."

Table 1 displays the semantic triplets defining the left, illustrating the construction of the enemy. Salvini mainly refers to the left as "these" but also as "communists" and "them," stressing out the typical populist dynamic of the "us vs. them." Similarly, Abascal identifies the political opponent with "the Left," but he also describes the opponent as autarchy, dictatorship, or "the Government" since, as previously mentioned, the Spanish government is made up of the Socialist Party and Unidas Podemos.

In both cases, the opponents show no positive quality and are represented as incapable, shameless, and guilty. In the case of Salvini, the accent is placed on their distance from "normal people" about whom they have no knowledge. The Left is accordingly identified as "radical chic." The term, coined in the 1970s by American journalist Tom Wolfe to satirize composer Leonard Bernstein for hosting a fundraising party for the Black Panthers, aims at lampooning upper-class individuals who endorse leftist radicalism merely to garner prestige rather than to affirm genuine political convictions (Colantone et al., 2022).

To stress this distance from the "real world," they are also portrayed as lacking common sense, living on another planet ("live on planet Mars"), or with words openly referring to madness, specifically *fuori* (from "*fuori di testa*," literally out of their mind, a colloquial word to say that someone has lost his head). In addition, "the Left" is portrayed as only interested in "keeping their chairs" (*poltrone*), that is to say, holding their position of power without really caring for the will or interests of the people.

Abascal, on his side, also underlines the elitist aspect of the ruling class and their distance from "normal people" by using the expression *progres*. However, he openly refers to the left-wing parties ruling the country as criminals identifying with a dictatorship. Figure 2, for instance, displays how the Spanish government is identified as a criminal "gang."

Table 1. Semantic triplets defining "the Left."

Variable	Salvini	Abascal
Subject	The Left	The Left
Definitions	The Left, them, these, the communists	The Left, the autarchy, the dictatorship, the government
Positive adjectives	—	—
Negative adjectives	Inept, useless, crazy, irresponsible, shameless, guilty, *fuori*, incompetent, radical chic, *buonisti*, live on planet Mars	Criminal, dictatorial, shameless, irresponsible, guilty, *progre*
Positive actions	Like, defend, love, show tenderness	Empathize, defend
Object	Immigrants, illegals, *poltrone*	Immigrants
Negative actions	Have no clue, do not care	Do not care, are unable, hate, betray, attack
Object	The people, normal people, Italians	Spaniards, citizens

Figure 2. Santiago Abascal's Tweet. Note: "Spain needs a government which serves the people, not a gang who serves personal, global, or independentist interests."

Both politicians, therefore, implement the strategy of demonizing their political opponents. The concept of demonization resonates with what Sabatier et al. (1987) identified as the "devil shift," the tendency for people to exaggerate the power and maliciousness of political opponents. In its original conception, the devil shift appears unintentional, but the narrative policy framework (Katz, 2018) has widely shown how it corresponds to an intentional strategy to build a villain figure.

Recognizing that human beings make sense of themselves by defining themselves and rhetorically constructing binaries implies that the role administered to the "Other" sets meaningful boundaries (Thurlow, 2010, p. 227). Demonization can be defined as a process through which a source promotes "a symbolic construction of reality created under the conceptual simplification protagonist–antagonist" (Civila et al., 2020, p. 2), in which the ideas of the sender are exposed as correct and justified, while the demonized group is accused of going against the common interest, dissociating them of an equitable moral nature to the "us"(Romero-Rodríguez et al., 2015). Accordingly, the demonization of the opponent provides the opportunity to represent the people as victims (Maronitis, 2021).

Previous studies (Michener et al., 2021) have shown that the perceived remoteness of those in high-status positions in the eye of the lower social strata fosters what Lamont (2018) calls the "recognition gap," causing feelings of alienation, exclusion, discontent, and resentment. The political opponents are accordingly represented as a danger to the people (those who suffer from their incapacity, lack of interest, etc.) and coherently, with PRR parties' nativist ideology, only represented as showing compassion or positive attitudes only toward the immigrants.

In addition, by stressing that "the Left" is persecuting, obsessed by, or insulting them, both leaders can also present themselves as victims, creating a further bond with "the people." Victimhood here should be understood as a performative action taking place on a public stage in which actors create and project performances of their life experiences, anxieties, and motives tailored to audiences (Maronitis, 2021).

Thus, acknowledging that the hero–villain narratives are "ethically constitutive" stories which "have special capacities to inspire senses of normative worth" (Smith, 2003, p. 59), the protagonist, that is to say, the hero, generally attracts empathy, affinity, and positive feelings, because the qualities assigned to the character resonate emotionally with the audience (Homolar, 2022).

In other words, highlighting that "the Left" accuse or insult them, both Salvini and Abascal can straighten their position of "champion of the people" (Bracciale & Martella, 2017). Figure 3 displays a perfect example of this dynamic: According to Salvini, a "desperate left exploit a murder to accuse Salvini, his party, and therefore Italians of being racist."

As previously anticipated, coherently with the different settings of the Spanish political system, Abascal also identifies independence movements, especially in Catalonia, as political opponents. Although statistically, the mentions of these parties are not as meaningful as the references to the left, it is worth analyzing them since, for Vox, the unity of Spain is as central as their positioning on the left–right continuum (Fernández-García & Valencia Sáiz, 2022; Marcos-Marne et al., 2021).

Parties asking for Catalonian independence are demonized following the same discursive pattern previously illustrated (see Table 2). In particular, they are defined as *separatistas* (those who want to break away from Spain) or openly as "those who want to break Spain" and called out as a mafia or terrorists and, as such, framed as a danger to the homeland. Interestingly, they are never called by their official name but only with derogatory epitomes, somehow discursively underlining their illegitimacy.

On the other hand, the Popular Party, Spain's main center-right party, is not only less frequently mentioned, but also its discursive treatment is completely different. Table 2 illustrates how, while independentists are demonized, the Popular Party is treated like a legitimate political competitor and framed as weak or incapable while also being granted positive qualities when it agrees with or supports Vox's political stances.

This difference, on the one hand, stresses the party's intrinsic anti-pluralism (Galston, 2017), showing an open rejection of worldviews that differ from theirs. On the other hand, it highlights an opportunist change of tone when dealing with a party with which Vox has formed—and might form—different alliances and coalitions in regional and local government (Barrio et al., 2021).

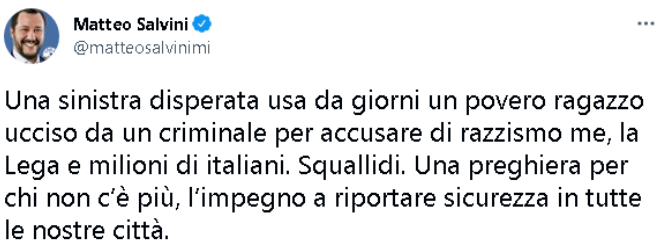


Figure 3. Matteo Salvini's Tweet. Note: "A desperate left uses a poor guy murdered to accuse me, Lega, and millions of Italian of racism. Shameless."

Table 2. Semantic triplets defining the Catalonian independentist parties and the Popular Party.

Variable	Catalonian independentist parties	Popular Party
Definitions	*Separatistas*, those who want to break Spain	"*El PP*"
Positive adjectives	—	Show good intentions
Negative adjectives	Criminal, dangerous, threat, enemies of Spain, mafia, terrorist	Weak, unable, irresponsible, guilty
Positive actions	—	Have understood
Object	—	VOX, Spaniards
Negative actions	Hate, want to break, destroy	Cannot or are not able to, lack the capacity
Object	Spain	Spaniards, citizens

6.3. Character Assassination

As displayed in Table 3, Abascal personalized attacks unsurprisingly mainly focus on Pedro Sanchez, Spanish prime minister and leader of the PSOE. In the case of Salvini, the most recurrent names are not members of the government but politicians from the left, particularly, Calenda and Letta.

The attacks against Sanchez, Calenda, and Letta, respectively, do not consist of argument-based political criticisms but, through aggressive tones, irony, mockery, and insult (Schwarzenegger & Wagner, 2018), they focus on individual traits and behaviors. Their aim, in other words, is not to engage in a political debate with the opponent but rather to undermine the opponent's reputation (Berti & Loner, 2021).

These political figures are targeted as individuals using their alleged personal flaws (such as ignorance or stupidity) or characteristics (being posh, aloof, or distant) and mocked in their personal style (such as wearing or not wearing a tie, as illustrated in Figure 4).

These kinds of attacks can be defined as "character assassination," that is to say, "a deliberate and sustained effort to damage the reputation or credibility of an individual'' (Samoilenko et al., 2016, p. 115) that works similarly to argumentum ad hominem (Wodak, 2015).

Table 3. Semantic triplets of the main political enemies.

Variable	Salvini	Abascal
Subject	Calenda, Letta	Pedro Sanchez
Definitions	These, Enrico (for Letta), Renzi's friend (for Calenda)	Swindler, hustler, dictator (*el autócrata*)
Positive adjectives	—	—
Negative adjectives	Ignorant, *Bocciati*, goes to the beach with a tie (for Calenda), posh, stupid, obsessed (by Salvini)	Dangerous, ignorant, useless, stupid, dictator, shameless, far (from reality and from the people), do not wear a tie
Positive actions	Like, defend, love, show tenderness	—
Object	Immigrants, illegals	—
Negative actions	Have no clue, do not care, are dangerous, hate, attack, has an obsession	Ruin, damage, betray, hate
Objects	Home, Italy, Italians, people, us, me	People, working class, Spain, Spaniards, our homeland

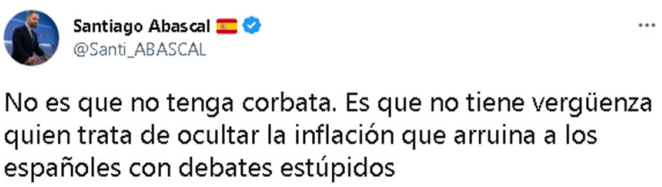


Figure 4. Santiago Abascal's Tweet. Note: "Not only does he not wear a tie. The problem is that he is shameless."

Character assassination directed toward individuals is often not focused on political and professional actions but rather on the target's personality and behavior and can include insults, defamation, and irony (Samoilenko et al., 2016).

Interestingly, following this dynamic, the most recurrent characters after the previously mentioned are two women who seem to have been chosen based more on their symbolic meaning than on their position. As illustrated in Table 4, Salvini mainly points out Laura Boldrini, president of the Chamber of Deputies, between 2013 and 2018, as having no institutional mandate in 2022, and Abascal focuses on Irene Montero, Minister of Equality since 2020.

The harsh attacks against Laura Boldrini are a leitmotif of Salvinian rhetoric (Cervi, 2020b). The reason why Boldrini seems to be the perfect personification of "the enemy" is that she is represented as the archetypical example of the above-mentioned "radical chic." Such individuals, politicians, and intellectuals of the left are believed to "earn a lot of money," "have no clue about the real problems of people," and actually "not care about the people," being in favor of immigrants. Boldrini is accused of being a *buonista*, meaning a "do-gooder," a neologism for those who carry out unnecessary acts of kindness which transforms a positive attribute, goodness, into an insult.

The background of this word, from a historical perspective, comes from the term "pietism," used by the Fascist regime, after 1938, against those who positioned themselves in favor of Jews who were being harassed by racial laws. Here, too, a virtue (piety or compassion) became distorted into a vice, a source of weakness (Cervi, 2020b). Accordingly, she is made fun of by portraying her as out of her mind and obsessed with Salvini, calling her "this" (seminal to "she," someone who does not deserve to be called by their name or title), and by identifying her by reference to her physical appearance as "tiger eyes" (see Figure 5).

Similarly to Boldrini, Irene Montero is depicted as a sort of "source of every evil," but if Salvini embodies in Boldrini the "immigrant loving elite," Abascal identifies in Montero the perfect representation of the "feminazi" (Bernardez-Rodal et al., 2022).

Exactly as in the previous case, she is mostly defined as "that lady/person" (again, someone who does not deserve to be called by her name or title), mocked as "*la marquesa de Galapagar*," referring to her being Pablo Iglesias' partner, and described as a danger for Spaniards (see Figure 6).

Table 4. Semantic triplets for Boldrini and Montero.

Variable	Salvini	Abascal
Subject	Laura Boldrini	Irene Montero
Definitions	Tiger eyes, lady, this	*La marquesa de Galapagar*, that lady, that person
Positive adjectives	—	—
Negative adjectives	*Buonista*, live on another planet, hopeless, obsessed (by Salvini)	Criminal, dangerous, ignorant, crazy, feminazi, threat (to our sons)
Positive actions	Like, defend, love, show tenderness	Empathize, defend
Object	Immigrants, illegals	Immigrants, rapists
Negative actions	Have no clue, do not care, hate, attack, has an obsession	Damage, harm, endanger, watch TV series, hate, fear, insult, accuse
Objects	Home, Italy, Italians, people, us, me	Spain, Spaniards, our homeland, us, Vox, the people

Figure 5. Matteo Salvini's Tweet. Note: "Didn't you miss tiger-eyed Boldrini?"

Figure 6. Santiago Abascal's Tweet. Note: "It's an international scandal that this person is still the minister. She is a threat to our children."

7. Conclusions

Our study has focused on two similar parties, Lega and Vox, operating in divergent political contexts. Our findings, on the one hand, show that, despite the differences, both parties identify the left as the main opponent. Because of the structure of the party system, these findings are not surprising in the Italian case. In fact, although in the analyzed time frame Lega formally supported the technocratic government led by Mario Draghi (Garzia & Karremans, 2021), Italian political competition is always based on the confrontation of the center-right versus the center-left block (Zulianello, 2020).

In the case of Vox, it is interesting to note that despite the unity of Spain being as central to the party as the right–left continuum (Barrio et al., 2021), for Abascal, the left epitomizes the political enemy.

In both cases, the discursive construction of "the Left" is based on two main strategies: demonization, the framing of opponents as "enemies of the people" who, along with dangerous "Others" such as immigrants, conspire against the "people" and are blamed for everything that is "wrong" in society; and character assassination of individual politicians through personal attacks, which aim to undermine their reputation and deflect attention from the real issues towards their personal traits and actions.

The demonization of political opponents, deprived of any positive qualities and emotionally blamed (Hameleers et al., 2017) for all the "evil," allows populist leaders to capitalize on the anger of those groups who perceive that their lives have been ignored, marginalized, or negatively affected by the actions of politicians (Horwitz, 2018). In this way, the subsequent victimization of the "people" offers populist leaders the chance to establish a powerful bond.

In other words, this discursive dynamic appears very similar to that identified by other studies (Carr & Haynes, 2015; Cervi et al., 2020) regarding the framing of immigrants as the "Others": Both dynamics consist of blaming the othered "Other" (Williams, 2010) through emotional appeal, avoiding any rational discussion and capitalizing on citizens' discontent.

In addition, in the specific case of Abascal, another political opponent is targeted, although its presence is less relevant from a quantitative perspective: independentist parties. These parties are not even dignified by being called by their official names; rather, they are referred to as *separatistas* or enemies of Spain and framed as a danger to the homeland. In the same vein, it is interesting to observe that these parties are more mentioned and called out as allies of the government rather than being the subject of tweets. On the other side, the Popular Party, which represents the main center-right option, is hardly ever mentioned, but when it is, it is treated as a legitimate political competitor.

As seen, therefore, while Salvini's populist dichotomous vision of the world is accompanied by a political system whose structure allows him to identify one political enemy to be blamed for everything, Abascal, who mainly focuses on the left (identified with the government), also has to deal with independentists, that are framed as enemies of the nation and with the Popular Party, positioned on the right side of the political spectrum and potentially representing a possible ally, is regarded as a legitimate competitor.

In addition, we have observed how these emotional attacks are mostly personalized. Personalization, besides being a central feature of much contemporary political communication (Garzia, 2011), is a key element of populism, which tends to construct charismatic leaders who claim to be the only authentic representatives of the people (Mudde & Kaltwasser, 2017).

In this sense, social media constitute a breeding ground for personalized politics: Communication can be filtered, re-framed, and re-contextualized, allowing populists to directly reach their audience (Groshek & Engelbert, 2012 and showcase their authenticity (Forchtner & Kølvraa, 2017) and closeness to the people.

Thus, character assassination becomes the ideal discursive strategy to fully exploit social media affordances to strengthen populist communication (Berti & Loner, 2021), focusing on opponents' personal lives and qualities, that is to say, on their personae, to undermine their reputation. In the same way, personalizing the "enemy" strengthen the "us vs. them" dichotomy that characterizes populist discourse.

Choosing a target that symbolically embodies all the "wrong" in society and attacking them through the use of mockery, insults, or impoliteness not only allows populists to distance themselves from establishment politics, characterized by moderation and issue-based arguments (Gerbaudo, 2018) but fosters polarization that has been proven to benefit them (Schulze et al., 2020).

In conclusion, our study has shed some light on PRR parties' discursive construction of the political "Other." Nonetheless, our results are inherently limited to the cases under analysis. Accordingly, future studies should extend the universe to prove if there is a common pattern outside Southern Europe.

Conflict of Interests

The authors declare no conflict of interests.

References

Aslanidis, P. (2018). Measuring populist discourse with semantic text analysis: An application on grassroots populist mobilization. *Quality and Quantity*, *53*, 1241–1263.

Barrio, A., de Oger, S. A. S., & Field, B. N. (2021). Vox Spain: The organisational challenges of a new radical right party. *Politics and Governance*, *9*(4), 240–251.

Bernardez-Rodal, A., Rey, P. R., & Franco, Y. G. (2022). Radical right parties and anti-feminist speech on

Instagram: Vox and the 2019 Spanish general election. *Party Politics*, *28*(2), 272–283.

Berti, C., & Loner, E. (2021). Character assassination as a right-wing populist communication tactic on social media: The case of Matteo Salvini in Italy. *New Media & Society*. Advance online publication. https://doi.org/10.1177/14614448211039222

Betz, H. G. (2017). Nativism across time and space. *Swiss Political Science Review*, *23*(4), 335–353. https://doi.org/10.1111/spsr.12260

Bobba, G. (2018). Social media populism: Features and "likeability" of Lega Nord communication on Facebook. *European Political Science*, *18*(1), 11–23. https://doi.org/10.1057/s41304-017-0141-8

Bracciale, R., & Martella, A. (2017). Define the populist political communication style: The case of Italian political leaders on Twitter. *Information, Communication & Society*, *20*(9), 1310–1329.

Capdevila, A., Moragas-Fernández, C. M., & Grau-Masot, J. M. (2022). Emergencia del populismo en España: Marcos metafóricos de Vox y de su comunidad online durante las elecciones generales de 2019 [Emergency of populism in Spain: Methaphorical frames of Vox and its online community during 2019 general elections]. *El Profesional de la Información*, *31*(3). https://doi.org/10.3145/epi.2022.may.17

Carr, J., & Haynes, A. (2015). A clash of racializations: The policing of "race" and of anti-Muslim racism in Ireland. *Critical Sociology*, *41*, 21–40.

Cervi, L. (2020a). Exclusionary populism and Islamophobia: A comparative analysis of Italy and Spain. *Religions*, *11*(10), Article 516. http://dx.doi.org/10.3390/rel11100516

Cervi, L. (2020b). *Veni*, *vidi*, Facebooked-live: Análisis del éxito de Matteo Salvini en Facebook [Veni, vidi, Facebooked-live: an analysis of Matteo Salvini's success on Facebook]. *Revista CIDOB d'Afers Internacionals*, *2021*(124), 99–122. https://doi.org/10.24241/rcai.2020.124.1.99

Cervi, L., & Carrillo-Andrade, A. (2019). Post-truth and disinformation: Using discourse analysis to understand the creation of emotional and rival narratives in Brexit. *ComHumanitas: Revista Científica de Comunicación*, *10*(2), 125–149. https://doi.org/10.31207/rch.v10i2.207

Cervi, L., García, F., & Marín-Lladó, C. (2021). Populism, Twitter, and Covid-19: Narrative, fantasies, and desires. *Social Sciences*, *10*(8), Article 294. http://dx.doi.org/10.3390/socsci10080294

Cervi, L., & Tejedor, S. (2020). Framing "the Gypsy problem": Populist electoral use of Romaphobia in Italy (2014–2019). *Social Sciences*, *9*(6), Article 105. http://dx.doi.org/10.3390/socsci9060105

Cervi, L., & Tejedor, S. (2021). "Africa does not fit in Europe": A comparative analysis of anti-immigration parties' discourse in Spain and Italy. *Migraciones*, *2021*(51), 207–239. https://doi.org/10.14422/mig.i51y2021.008

Cervi, L., Tejedor, S., & Alencar Dornelles, M. (2020). When populists govern the country: Strategies of legitimization of anti-immigration policies in Salvini's Italy. *Sustainability*, *12*(23), Article 10225. http://dx.doi.org/10.3390/su122310225

Civila, S., Romero-Rodríguez, L. M., & Civila, A. (2020). The demonization of Islam through social media: A case study of #StopIslam in Instagram. *Publications*, *8*(4), Article 52. https://doi.org/10.3390/publications8040052

Colantone, I., Di Lonardo, L., Margalit, Y., & Percoco, M. (2022). *The political consequences of green policies: Evidence from Italy*. SSRN. http://dx.doi.org/10.2139/ssrn.4048520

De Vreese, C. H., Esser, F., Aalberg, T., Reinemann, C., & Stanyer, J. (2018). Populism as an expression of political communication content and style: A new perspective. *The International Journal of Press/Politics*, *23*(4), 423–438.

Diamanti, G., & Pregliasco, L. (2019). *Fenomeno Salvini: Chi è, come comunica, perché lo votano* [The Salvini phenomenon: Who is he, how he communicates, why people vote for him]. Castelvecchi.

Ekström, M., Patrona, M., & Thornborrow, J. (2018). Right-wing populism and the dynamics of style: A discourse-analytic perspective on mediated political performances. *Palgrave Communications*, *4*(1), Article 4. https://doi.org/10.1057/s41599-018-0132-6

Engesser, S., Ernst, N., Esser, F., & Büchel, F. (2016). Populism and social media: How politicians spread a fragmented ideology. *Information, Communication & Society*, *20*(8), 1109–1126. https://doi.org/10.1080/1369118x.2016.1207697

Espejo, P. O. (2011). *The time of popular sovereignty: Process and the democratic state*. Penn State Press.

Fawzi, N. (2019). Untrustworthy news and the media as "enemy of the people"? How a populist worldview shapes recipients' attitudes toward the media. *The International Journal of Press/Politics*, *24*(2), 146–164.

Fernández-García, B., & Valencia Sáiz, Á. (2022). Nacionalismo y populismo en el contexto de las elecciones catalanas de 2021: El populismo al servicio de agendas nacionalistas enfrentadas [Nationalism and populism in the context of the Catalonian elections of 2021: Populism serving confronted nationalistic agendas]. *Revista Española de Ciencia Política*, *59*, 13–42. https://doi.org/10.21308/recp.59.01

Forchtner, B., & Kølvraa, C. (2017). Extreme right images of radical authenticity: Multimodal aesthetics of history, nature, and gender roles in social media. *European Journal of Cultural and Political Sociology*, *4*(3), 252–281.

Franzosi, R. (2010). *Quantitative narrative analysis*. Thousand Oaks.

Galston, W. (2017). *Anti-pluralism: The populist threat to liberal democracy*. Yale University Press.

Garzia, D. (2011). The personalization of politics in Western democracies: Causes and consequences on leader–follower relationships. *The Leadership Quarterly*, *22*(4), 697–709. https://doi.org/10.1016/J.LEAQUA.2011.05.010

Garzia, D., & Karremans, J. (2021). Super Mario 2: Comparing the technocrat-led Monti and Draghi governments in Italy. *Contemporary Italian Politics*, *13*(1), 105–115.

Gerbaudo, P. (2018). Social media and populism: an elective affinity? *Media, Culture & Society*, *40*(5), 745–753. https://doi.org/10.1177/0163443718772192

Gray, C. (2020). *Territorial politics and the party system in Spain: Continuity and change since the financial crisis*. Routledge.

Groshek, J., & Engelbert, J. (2012). Double differentiation in a cross-national comparison of populist political movements and online media uses in the United States and the Netherlands. *New Media & Society*, *15*(2), 183–202. https://doi.org/10.1177/1461444812450685

Hameleers, M. (2020). Augmenting polarization via social media? A comparative analysis of Trump's and Wilders' online populist communication and the electorate's interpretations surrounding the elections. *Acta Politica*, *55*(3), 331–350.

Hameleers, M., Bos, L., & De Vreese, C. (2017). They did it: The effects of emotionalized blame attribution in populist communication. *Communication Research*, *44*(6), 870–900.

Hawkins, K. A. (2010). *Venezuela's Chavismo and populism in comparative perspective*. Cambridge University Press.

Hawkins, K. A., Aguilar, R., Silva, B. C., Jenne, E. K., Kocijan, B., & Kaltwasser, C. R. (2019, June 20–22). *Measuring populist discourse: The global populism database* [Paper presentation]. EPSA Annual Conference, Belfast, UK.

Hawkins, K. A., & Kaltwasser, C. R. (2017). The ideational approach to populism. *Latin American Research Review*, *52*(4), 513–528.

Hawkins, K. A., Riding, S., & Mudde, C. (2012). *Measuring populist attitudes* (Working Paper No. 55). The Committee on Concepts and Methods. https://www.concepts-methods.org/Files/WorkingPaper/PC_55_Hawkins_Riding_Mudde.pdf

Homolar, A. (2022). A call to arms: Hero–villain narratives in US security discourse. *Security Dialogue*, *53*(4), 324–341. https://doi.org/10.1177/09670106211005897

Hooghe, L., & Marks, G. (2018). Cleavage theory meets Europe's crises: Lipset, Rokkan, and the transnational cleavage. *Journal of European Public Policy*, *25*(1), 109–135. https://doi.org/10.1080/13501763.2017.1310279

Horwitz, R. B. (2018). Politics as victimhood, victimhood as politics. *Journal of Policy History*, *30*(3), 552–574.

Jacobs, K., & Spierings, N. (2018). A populist paradise? Examining populists' Twitter adoption and use. *Information, Communication & Society*, *22*(12), 1681–1696. https://doi.org/10.1080/1369118x.2018.1449883

Kaltwasser, C. R., Taggart, P. A., Espejo, P. O., & Ostiguy, P. (Eds.). (2017). *The Oxford handbook of populism*. Oxford University Press.

Katz, J. (2018). The space between: Demonization of opponents and policy divergence. *Review of Policy Research*, *35*(2), 280–301. https://doi.org/10.1111/ropr.12279

Kefford, G., Moffitt, B., & Werner, A. (2022). Populist attitudes: Bringing together ideational and communicative approaches. *Political Studies*, *70*(4), 1006–1027. https://doi.org/10.1177/0032321721997741

Koopmans, R., & Muis, J. (2009). The rise of right-wing populist Pim Fortuyn in the Netherlands: A discursive opportunity approach. *European Journal of Political Research*, *48*(5), 642–664. https://doi.org/10.1111/j.1475-6765.2009.00846.x

Krzyżanowski, M., & Ekström, M. (2022). The normalization of far-right populism and nativist authoritarianism: Discursive practices in media, journalism and the wider public sphere/s. *Discourse & Society*, *33*(6), 719–729. https://doi.org/10.1177/09579265221095406

Laclau, E. (2005). *On populist reason*. Verso.

Lamont, M. (2018). Addressing recognition gaps: Destigmatization and the reduction of inequality. *American Sociological Review*, *83*(3), 419–444.

March, L. (2017). Left and right populism compared: The British case. *The British Journal of Politics and International Relations*, *19*(2), 282–303.

Marcos-Marne, H., Plaza-Colodro, C., & O'Flynn, C. (2021). Populism and new radical-right parties: The case of VOX. *Politics*. Advance online publication. https://doi.org/10.1177/02633957211019587

Maronitis, K. (2021). The present is a foreign country: Brexit and the performance of victimhood. *British Politics*, *16*(3), 239–253. https://doi.org/10.1057/s41293-020-00151-7

Mayer, F. W. (2014). *Narrative politics: Stories and collective action*. Oxford University Press.

Meny, Y., & Surel, Y. (2002). *Democracies and the populist challenge*. Palgrave Macmillan.

Michener, G., Amorim Neto, O., & Civitarese, J. (2021). The remoteness of democratic representation. *Party Politics*, *29*(1), 51–64. https://doi.org/10.1177/1354068821104954 5

Moffitt, B. (2016). *The global rise of populism: Performance, political style, and representation*. Stanford University Press.

Mudde, C. (2004). The populist zeitgeist. *Government and Opposition*, *39*(4), 541–563.

Mudde, C. (2007). *Populist radical right parties in Europe*. Cambridge University Press.

Mudde, C., & Kaltwasser, C. R. (Eds.). (2012). *Populism*

in Europe and the Americas: Threat or corrective for democracy? Cambridge University Press.
Mudde, C., & Kaltwasser, R. (2017). *Populism: A very short introduction* (2nd ed.). Oxford University Press.
Olivas, J. J. (2021). From chasing populists to deconstructing populism: A new multidimensional approach to understanding and comparing populism. *European Journal of Political Research*, *60*(4), 829–853.
Ostiguy, P. (2009). *The high-low political divide: Rethinking populism and anti-populism* (Working Paper No. 35). The Committee on Concepts and Methods.
Ostiguy, P. (2017). A socio-cultural approach. In C. R. Kaltwasser, P. A. Taggart, P.O. Espejo, & P. Ostiguy (Eds.), *The Oxford handbook of populism* (pp. 73–79). Oxford University Press.
Ostiguy, P., Panizza, F., & Moffitt, B. (Eds.). (2020). *Populism in global perspective: A performative and discursive approach*. Routledge.
Parmelee, J. H. (2013). The agenda-building function of political tweets. *New Media & Society*, *16*(3), 434–450. https://doi.org/10.1177/1461444813487955
Popping, R., & Roberts, C. W. (2014). Semantic text analysis and the measurement of ideological developments within fledgling democracies. *Social Science Information*, *54*(1), 23–37. https://doi.org/10.1177/0539018414553865
Romero-Rodríguez, L., Aguaded, I., & Gadea, W. (2015). De la demonización a la polarización: Un análisis desde el discurso digital del gobierno y la oposición venezolana [From demonization to polarization: An analysis of the digital discourse of government and opposition in Venezuela]. *Argos*, *32*, 97–117.
Sabatier, P. A., Hunter, S., & McLaughlin, S. (1987). The devil shift: Perceptions and misperceptions of opponents. *The Western Political Quarterly*, *40*(3), 449–476.
Samoilenko, S., Shiraev, E., & Keohane, J. (2016). Character assassination. In C. Carrol (Ed.), *The SAGE encyclopedia of corporate reputation* (pp. 115–118). SAGE.
Sartori, G. (2005). *Parties and party systems: A framework for analysis*. ECPR Press.
Schulze, H., Mauk, M., & Linde, J. (2020). How populism and polarization affect Europe's liberal democracies. *Politics and Governance*, *8*(3), 1–5. https://doi.org/10.17645/pag.v8i3.3460
Schwarzenegger, C., & Wagner, A. (2018). Can it be hate if it is fun? Discursive ensembles of hatred and laughter in extreme right satire on Facebook. *Studies in Communication | Media*, *7*(4), 473–498. https://doi.org/10.5771/2192-4007-2018-4-473
Smith, R. M. (2003). *Stories of peoplehood: The politics and morals of political membership*. Cambridge University Press.
Stavrakakis, Y., Katsambekis, G., Kioupkiolis, A., Nikisianis, N., & Siomos, T. (2018). Populism, anti-populism and crisis. *Contemporary Political Theory*, *17*, 4–27.
Stoppino, M. (2001). *Potere e teoria política* [Power and political theory]. Giuffrè.
Tarchi, M. (2015). Italy: The promised land of populism? *Contemporary Italian Politics*, *7*(3), 273–285.
Thurlow, C. (2010). Speaking of difference: Language, inequality and interculturality. In K. Thomas, T. Nakayama, & R. Halualani (Eds.), *The handbook of critical intercultural communication* (pp. 227–247). Wiley Blackwell.
Van Kessel, S., & Castelein, R. (2016). Shifting the blame: Populist politicians' use of Twitter as a tool of opposition. *Journal of Contemporary European Research*, *12*(2), 594–614.
Weyland, K. (2001). Clarifying a contested concept: Populism in the study of Latin American politics. *Comparative Politics*, *34*(1), 1–22.
Williams, M. H. (2010). Can leopards change their spots? Between xenophobia and trans-ethnic populism among West European far right parties. *Nationalism and Ethnic Politics*, *16*, 111–134.
Winberg, O. (2017). Insult politics: Donald Trump, right-wing populism, and incendiary language. *European Journal of American Studies*, *12*(2). https://doi.org/10.4000/ejas.12132
Wodak, R. (2015). *The politics of fear: What right-wing populist discourses mean*. SAGE.
Zulianello, M. (2020). Varieties of populist parties and party systems in Europe: From state-of-the-art to the application of a novel classification scheme to 66 parties in 33 Countries. *Government and Opposition*, *55*(2), 327–347. https://doi.org/10.1017/gov.2019.21

About the Authors

Laura Cervi is a Serra Húnter associate professor of the Department of Journalism and Communication Sciences at the Autonomous University of Barcelona, Spain. She holds a PhD in political science from the University of Pavia (Italy) and the Autonomous University of Barcelona. She is a member of the research group Gabinete de Comunicación y Educación, a recognized and consolidated research group at the Autonomous University of Barcelona. Her main research interest is political communication, with a focus on populism and its relations with new technologies.

Santiago Tejedor is a professor and director of the Department of Journalism and Communication Sciences at the Autonomous University of Barcelona. He holds a PhD in journalism and communication sciences from the Autonomous University of Barcelona and a PhD in project engineering from the Polytechnic University of Catalonia. He is currently the director and principal investigator of the research group Gabinete de Comunicación y Educación, a recognized and consolidated research group. His lines of research focus are, in particular, on media literacy, cyberjournalism, media convergence, and new transmedia narratives.

Mónica Gracia Villar has a master's degree in human resources and knowledge management from the European University of the Atlantic and a PhD in Project Engineering: Environment, Safety, Quality, and Communication from the Polytechnic University of Catalonia, Spain. She is currently a researcher and teacher of internal communication and the development of entrepreneurial social skills.

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 249–260
https://doi.org/10.17645/pag.v11i2.6465

Article

Accountability Issues, Online Covert Hate Speech, and the Efficacy of Counter-Speech

Fabienne Baider

Department of French and European Studies, University of Cyprus, Cyprus; fabienne@ucy.ac.cy

Submitted: 8 October 2022 | Accepted: 4 January 2023 | Published: 17 May 2023

Abstract
Concerning individual or institutional accountability for online hate speech, research has revealed that most such speech is covert (veiled or camouflaged expressions of hate) and cannot be addressed with existing measures (e.g., deletion of messages, prosecution of the perpetrator). Therefore, in this article, we examine another way to respond to and possibly deflect hate speech: counter-speech. Counter-narratives aim to influence those who write hate speech, to encourage them to rethink their message, and to offer to all who read hate speech a critical deconstruction of it. We created a unique set of parameters to analise the strategies used in counter-speech and their impact. Upon analysis of our database (manual annotations of 15,000 Twitter and YouTube comments), we identified the rhetoric most used in counter-speech, the general impact of the various counter-narrative strategies, and their specific impact concerning several topics. The impact was defined by noting the number of answers triggered by the comment and the tone of the answers (negative, positive, or neutral). Our data reveal an overwhelming use of argumentative strategies in counter-speech, most involving reasoning, history, statistics, and examples. However, most of these argumentative strategies are written in a hostile tone and most dialogues triggered are negative. We also found that affective strategies (based on displaying positive emotions, for instance) led to a positive outcome, although in most cases these narratives do not receive responses. We recommend that education or training—even machine learning such as empathetic bots—should focus on strategies that are positive in tone, acknowledging grievances especially.

Keywords
accountability; argumentative strategies; counter-speech; covert hate speech; emotional appeal

Issue
This article is part of the issue “Hate Speech, Demonization, Polarization, and Political Social Responsibility” edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

Concerning individual or institutional accountability for online hate speech, we argue that, because most online hate speech is covert, current measures regulating hate speech are insufficient. For example, the 2016 Code of Conduct from the EU Commission falls short in several regards (Konikoff, 2021), including concerns about the qualifications of those deleting hate messages and the fact that artificial intelligence models used to detect hate speech are 1.5 times more likely to flag tweets written by specific communities (Silva et al., 2016). Covert hate speech entails even more problems, as it uses implicit meaning and indirect discursive strategies to express hatred, including derogative metaphors (Musolff, 2015), inferences (Baider, 2022), and humor (Weaver, 2016). Such covert expressions fall outside the legal definitions of hate speech, and thus purveyors of such hateful speech remain unaccountable before the law. Examination of how hate speech is regulated by social media platforms such as Facebook and YouTube (Fortuna & Nunes, 2018; Hietanen & Eddebo, 2022) underscores the diverse interpretations of what constitutes hate speech. For this reason, we argue for greater emphasis on counter-speech rather than censorship (Strossen, 2018) as the best way to deflect or halt hate speech.

We suggest the use of counter-narratives, which we define as any form of expression that aims to influence those who sympathise with or take part in abusive speech. These narratives can encourage those who write hate speech to rethink their message, while at the same time, they offer a critical counter-argument to all who read the hate speech. They also offer another point of view and can potentially trigger positive feelings for victims of discriminatory narratives. The present study discusses the most frequent types of counter-narratives and their impact, based on analysis of our database of manual annotations of 15,000 Twitter and YouTube comments (collected within the IMsyPP EU program).

2. Addressing Online Hate Speech: Censorship vs. Dialogue

Since the 1990s, research targeting hate speech has noted the prevalence of hostile and aggressive content in online platforms, which might suggest that the medium itself is partially to blame—insofar as it offers anonymity, instantiation of communication, depersonalisation, deindividuation, etc. (cf. Baider, 2020). In fact, Wodak (2015, p. 207, emphasis added) concluded that "the more anonymous the genre, the more *explicit exclusionary rhetoric* tends to be."

2.1. Overt and Covert Hate Speech

Indeed, many of the advantages of digitisation, e.g., connectivity, access to new knowledge, and the creation of new relationships, have led to the rapid rise in cyber hate across the internet. As recently as 20 years ago, social media platforms, online fora, and group discussions were found to be prime locations for the collection and analysis of (violent) discriminatory discourse (Herring et al., 2002, p. 371). Such discourses manifest in various ways, including Twitter mobbing, trolling, cyberbullying, and sexting—all of which may fall under the umbrella term "hate speech" depending on the definition applied.

And herein lies part of the problem: the many, often contradictory, official definitions of hate speech. As early as 1965, the United Nations General Assembly Resolution 2106, in their International Convention on the Elimination of All Forms of Racial Discrimination, defined hate speech as "the promotion of racial hatred and discrimination" based on "ideas or theories of superiority of one race or group of persons of one colour or ethnic origin," and as speech that would incite "racial discrimination, or acts of violence…against any race or group of persons of another colour or ethnic origin" (United Nations, 1965, p. 3). The International Covenant on Civil and Political Rights, adopted by the United Nations General Assembly Resolution 2200A (XXI) and commonly used in court cases, defines hate speech as an "any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence" (United Nations, 1966, article 20). Most research studies are based on the broad definition suggested by the Office for Democratic Institutions and Human Rights (2009, pp. 37–46), which is the principal institution of the Organisation for Security and Cooperation in Europe dealing with the "human dimension" of security: "The expression of hatred towards an individual or group of individuals on the basis of protected characteristics, where the term 'protected characteristics' denotes a member of some specific social group that could, on its own, trigger discrimination.''

In EU countries, all judgments and social network regulations are based on the 2008 European Union Framework Decision, which delineates hate speech as statements "publicly inciting to violence or hatred directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin" and "publicly condoning, denying or grossly trivialising crimes of genocide, crimes against humanity and war crimes" (Council of the European Union, 2008, Article 1).

Further, there are some scholars (Gelber, 2019) who define hate speech as speech that causes harm to a degree sufficient to warrant government regulation, reasoning also followed by some legal judgments; for example, the case of a British National Party poster bearing the message "Islam out—Protect the British people," accompanied by a photo of the World Trade Center towers in flames and a sign with the Islamic crescent barred, which was legally judged as hate speech and banned (*Norwood v. the United Kingdom*, 2004). Another such example is *Leroy v. France* (2008), where Leroy was convicted of publicly condoning terrorism with his 2001 cartoon published in a Basque newspaper. The cartoon represented the attack on the World Trade Center with the caption: "We all dreamt of it….Hamas did it." Because the humorous use of a well-known catchphrase to express criticism against the United States was also glorifying extreme violence and death, the locutionary act was judged condemnable. However, the likelihood of causing harm, i.e., the perlocutionary effect, was also decisive in Leroy's condemnation since the Basque country is politically sensitive to terrorism. Thus, we can see that to determine whether the speech qualifies as hate speech, the courts must consider the social, historical, and cultural context, and "under what circumstances targets are vulnerable to harm" (*Leroy v. France*, 2008)—a definition that is very similar to that of the 2012 Rabat Plan of Action (United Nations, 2012).

Considering the various legal definitions, therefore, it is clear that covert hate speech is difficult to moderate and regulate, even though this disguised means of expressing hatred or calls for violence has sometimes been successfully addressed under present legislation (cf. *Norwood v. the United Kingdom* and *Leroy v. France*, among many). Indeed, it is not difficult to replace explicit stereotypes, which could be prosecuted, with implicit ones, and thus communicate hateful comments through other means. These covert means use disguised ways to

express racism, sexism, homophobia, or any bias against a specific community that could incite violence (Baider, 2020; Ben-David & Fernández, 2016; Matsuda, 1989). They are based on the same stereotypes and harmful prejudices as found in overt hate speech but use indirect strategies to express hateful sentiments, and/or a very negative stance towards specific communities. These covert strategies include metaphors (Musolff, 2015), sarcastic remarks or humour (Hill, 2008; Weaver, 2016), conspiracy theories (Baider, 2022), dog-whistling strategies (a strategy that refers to the use of words, phrases, and terminology that mean one thing to the public at large, but that carry an additional, implicit meaning only recognised by a specific subset of the audience; Bhat & Klein, 2020, p. 168), and memes (Askanius, 2021); even absence or silence can be used to invite hateful inferences (Hill, 2008, p. 41). They function as "Othering" mechanisms that breed anger, disgust, contempt, and fear towards a specific community—all emotions that are core to hatred, which in turn is core to extremism. It is extremely difficult to legally address these indirect expressions of hate speech; often speakers can escape accountability by pleading an excuse such as "I did not mean it" in cases of sarcasm for instance, or by using hedges such as "no offence, but" when uttering an insulting remark, etc. This type of hate speech is the most common form of racism, sexism, homophobia, etc., found on social media, supplanting overt hate speech by a huge margin (Bhat & Klein, 2020). In our data, less than 10% of comments are overt hate speech.

If the last ten years have seen an ever-growing dependency on automatic detection mechanisms to identify hate speech (Fortuna & Nunes, 2018), it is clear that online participants who want to express extremely negative attitudes have found solutions to circumvent censorship.

2.2. Hate Speech Management

Hate speech is typically managed by one or more of four responses (Benesch et al., 2016; Citron & Norton, 2011): deletion or suspension, inaction, education, and counter-speech. The response of inaction implies not responding to the abusive message and can lead to one of two consequences. On the one hand, ignoring a hate-filled message can lessen its impact, as it neither encourages nor feeds the debate and its (possible) ensuing thread. On the other hand, it can imply that such speech is acceptable. The education response involves training, media literacy, and national or international campaigns to inform the public, especially the youth, about hate speech, its consequences, and the best ways to address the messages. While this solution is important, it is more of a long-term investment.

A more immediate measure would be to increase accountability within computer meditated communication. Brown (2020, p. 32) has argued for different levels of governance regarding hate speech: the moderation level, which would primarily concern social media companies; and the regulatory level, which would concern agencies. The regulatory level is typically assumed by governments or their agencies, which guide internet governance and enact legislation. In fact, there are already a number of legal measures that enforce some degree of communication etiquette in computer meditated communication—e.g., international instruments such as the 2019 UN Action Plan on hate speech and the 2016 Code of Conduct enacted by the EU Commission, the latter being the most drastic measure, which is also closely monitored (Pingen, 2021). This EU Code of Conduct mandates that social media companies remove or disable access to—and within 24 hours—what has been deemed illegal hate speech on the basis of the 2007 Framework definition.

According to Brown (2020, p. 32), responsibility for the second type of governance, moderation, is to be assumed by internet platforms and ordinary citizens. Social platforms have therefore delineated their own limitations on freedom of speech. As an example, Facebook has a complex set of rules determining what constitutes hate speech, and even considers covert hate speech in some cases. They will take down straightforward animal metaphors such as "migrants are filthy cockroaches" and well-known wordplay such as "refugees and rape-fugees." However, they will differentiate between "migrants are so filthy" (non-violating—ignore) and "all English people are dirty" (violating—delete), or "fucking migrants" (non-violating—ignore) and "fucking Muslims" (violating—delete). To understand their evaluation of the statements, we have to bear in mind that:

1. A statement such as "migrants are filth" is deleted since the metaphor "migrants are DIRT" is an established metaphor in racist discourse. This example reveals that covert hate speech (here, a metaphor), even if not always identified as such by Facebook, is nevertheless covered in its anti-hate speech rules.
2. The statement "all English people are dirty" is deleted because condemning people based on their nationality violates hate speech laws, a rule consistent with the Council of Europe definition specifying nationality as a criterion that warrants the label of hate speech.

Therefore, while it is the task of artificial intelligence mechanisms to detect hate speech, these mechanisms follow the social media regulators' understanding and definition of hate speech. This raises both questions and concerns over the legitimacy of anyone other than trained lawyers deciding what is hate speech. Indeed, to evaluate what qualifies as such is hard enough for human beings, never mind artificial intelligence systems. Moreover, it appears that such rules are devised by socially homogeneous teams (Baider, 2020)

since artificial intelligence models have been found to flag tweets written by African Americans as offensive 1.5 times more often—in other words, a false positive—than tweets written by other communities (Sap et al., 2019). This is explained by the over-sensitivity of hate speech classifiers, e.g., "nigger" or "bitch," which do not signal hate speech when used in specific settings; rather, they can signify relational proximity in some communities of practice (Baider, 2020; Culpeper, 2021).

However, although deleting a message limits its spread on a specific network, it encourages the author to post the same message on another, less-censored network. It does not challenge the arguments or the resentment expressed in such messages.

Moreover, critics argue that blocking free speech is a dangerous precedent and that these measures may restrict freedom of expression, which is recognised by the European Court of Human Rights as a fundamental human right and a basic condition of democratic societies, as well as necessary for individual development. Indeed, the danger with such laws is that they could be used to curb dissent and pursue "persecution of minorities under the guise of anti-incitement laws" (United Nations, 2012); for example, blasphemy laws can threaten inter-religious dialogue and ban legitimate debate. Clearly, transnational regulation of hate speech is not an easy task (Burnap & Matthew, 2015).

Most relevant to our study is the first solution, deletion or suspension, which we argue cannot address covert hate speech. It does not respond to the need for an immediate answer to the millions of messages exchanged every day.

We suggest that counter-speech would be the best solution, since debate "is nearly always preferable to censorship and removal of content, including when dealing with extreme or radical content, whatever its origin" (Bartlett & Krasodomski-Jones, 2016, p. 5). In fact, over 20 years ago Richards and Calvert (2000) argued that the best way to combat hate speech would be to add more speech, i.e., to use counter-speech to tackle hate speech. Indeed, the advantages of counter-speech are far greater than those of deletion—one example is using specific argumentation to respond in particular social contexts, hopefully destabilising the presupposition on which hate speech is based (McGowan, 2009). While counter-speech should respect freedom of speech, we found in our data that it is often violent, ultimately defeating the purpose of halting spiraling violence. A message that responds to the stereotyping that is core to hate speech offers readers another point of view and a chance to "take back" cyberspace. In fact, it may be our responsibility as forum participants to address such issues rather than let them pass unattended, as the use of automatic bots to respond to hate speech is a very recent solution to counter the massive number of message exchanges (Ashida & Komachi, 2022).

Counter-speech is not an entirely new subject, and to date, there are a number of studies examining its effect on hate speech. The most important studies will be discussed in the next section.

2.3. Counter-Speech: Definition and Impact

Definitions of counter-speech vary; for example, Bartlett and Krasodomski-Jones (2016, p. 5) propose "a crowd-sourced response to hateful messages," which means a direct response to harmful speech using any form of expression, whether a text, a meme, a hyperlink, etc. The researchers focus on argumentation based on logic or affect, whereas counter-speech aims to deconstruct hate speech and weaken its impact. Other scholars posit that counter-speech can also take the form of an alternative narrative (Braddock & Horgan, 2016; Briggs & Feve, 2013). Alternative narratives make a deliberate choice to change the narrative, focusing on positive stories to promote tolerance and debunk the presupposition on which the hate speech is based. In any case, in this article, we use counter-speech as a hypernym that includes alternative narratives and counter-narratives.

Counter-narratives should attempt to affect the behavior and the thinking of those who sympathise with or take part in spreading prejudices. Most important, they should foster critical thinking, tackle the source of prejudice (McGowan, 2009), and provoke reactions (i.e., spark a dialogue even if it is fierce; Gemmerli, 2015; Silverman et al., 2016). At the same time, they should also point out the complexity of the issue, and facilitate exposure to alternative viewpoints (Bartlett & Krasodomski-Jones, 2016); they should encourage readers to condemn hateful comments, trigger positive feelings (such as empathy) for victims of discriminatory narratives, and/or trigger some doubt that could lead to a change in attitudes (Gemmerli, 2015; Silverman et al., 2016). While arguments exchanged between strangers may lead to a favourable change in discourse, this is very rare (Bartlett & Krasodomski-Jones, 2016; Benesch et al., 2016; Ernst et al., 2017; Konikoff, 2021; Schieb & Preuss, 2016; Wright et al., 2017). Most research is based on small experiments such as Munger's (2017), which attest to the power of in-group norms and the need to tackle this phenomenon if we want to reduce racism. The studies above argue that the most effective messages do not lecture the audience, rather, they must offer something to think about and reflect on (Braddock & Horgan, 2016; Gagliardone et al., 2015).

For that matter, Benesch and his colleagues, who define hate speech as "dangerous speech" (see Benesch, 2014; Benesch et al., 2016) were the first to suggest a series of strategies for writing counter-speech and reducing the impact of hateful comments: (a) present the facts in order to correct misstatements or misperceptions; (b) point out hypocrisy or contradictions to discredit the accuser; (c) warn of offline or online consequences of such action; (d) claim some affiliation to give weight to the counter-speech; (e) denounce the speech as hateful; (f) use humour and sarcasm to deescalate

conflict and encourage social cohesion; (g) adopt a positive tone to appeal to the other participants; (h) adopt hostile language to potentially persuade a participant to delete their message.

The few large-scale research projects that have focused on counter-narratives have used these parameters. For example, they were the basis for the Conan project in which Chung et al. (2019) created the first large-scale, multilingual dataset of hate speech and counter-narrative pairs, i.e., type of hate speech vs. type of counter-speech.

However, hate speech is foremost a type of argumentation, i.e., an attempt to persuade others that a specific community or individual is a danger to them; Stephan et al. (1999) examined hate-filled comments from a psychological perspective and concluded that the concept of "threat" was core to hate speech, especially racism.

Therefore, when writing a counter-narrative, it is important to first identify the strategies of argumentation and determine if they are dependent on the topic, and then identify their impact and focus on deconstructing the presuppositions of their comments. To determine the most effective type of counter-speech we will address two research questions that have not yet been answered with big data:

RQ1: What counter-speech argumentative strategies are used in situ on a large-scale basis?

RQ2: What is the impact of each of these different strategies in relation to the different identified topics?

3. Data, Methodology, and Results

To answer our research questions, we begin with a quantitative approach, drawing statistics from data that has been annotated. We focused on strategies of argumentation used in counter-speech to understand which strategies are widely used (RQ1) and the impact of these strategies (RQ2). Before presenting our results, we describe the data of our corpus and the annotation scheme we developed for the project.

3.1. Data and Teamwork

We worked with data available within the IMsyPP EU project (2020–2022): 15,000 annotated Facebook posts and YouTube comments focused on several topics known to trigger hate speech, i.e., migration, politics, and LGBTQ issues. The category "politics" is an umbrella term covering a variety of political topics (e.g., India helping Pakistan during the Covid-19 pandemic). The comments and posts referring to migration were collected from Facebook in 2015, when an unprecedented number of migrants flooded into Europe, while the comments related to LGBTQ and political issues were collected in 2020 from YouTube. Each comment was first annotated for triggering hate speech and offensive speech, whether it was covert or overt, resulting in 9,700 comments annotated by eight annotators working in pairs. All comments were then tagged twice for counter-speech and assessed for impact; ultimately, the idea was to offer recommendations. The datasets are all in English and are comparable, insofar as they have a similar number of comments. The datasets were all annotated by the same team for a period of one year (see Table 1 for a summary of the data).

3.2. Methodology

We had to first decide on a set of parameters to annotate the counter-narratives, so we turned to earlier research studies, notably Benesch (2014) and Benesch et al. (2016), whose parameters categorising counter-speech have been widely used (see, e.g., Braddock & Horgan, 2016; Chung et al., 2019; Tuck & Silverman, 2016). As noted earlier, these parameters are: presenting facts, pointing out hypocrisy, warning of consequences, claiming some affiliation, denouncing the speech as hateful, using humor and sarcasm, adopting a positive tone, and adopting hostile language. We also added using multimedia, as Benesch et al. (2016) advised.

As we noted above, counter-narratives must be tested and evaluated in terms of their strategies as well as their impact—for example, a measurable change in behavior. Therefore, we created a category titled impact, wherein we took note of the number of answers triggered by the comment and the tone of the answers (whether negative, positive, or neutral).

We next ran a two-week pilot study to test these criteria, which ultimately resulted in the creation of our own set of annotations. Our pilot study revealed several shortcomings in the criteria, as follows:

1. Some of Benesch's criteria were absent from our annotations, e.g., warning of consequences,

Table 1. Datasets used for annotations.

Data				
Number of comments	Number of annotated comments	Source of dataset	Topic	Language
5,873	3,700	Facebook	Migration	English
3,009	3,000	YouTube	Politics	English
5,979	3,000	YouTube	LGBTQ issues	English
Total: 14,861	Total: 9,700			

claiming some affiliation, and pointing out hypocrisy, or too difficult to be distinguished from other choices;
2. We found other elements being used, such as "acknowledging grievances," that were not present in Benesch's criteria;
3. The criteria "presenting facts" was found to be too broad, since merely offering data may not be sufficiently convincing. It does, however, show the audience that the accusations are not substantiated, so we subdivided the category by adding "using statistics," "using history," and "using examples or testimonies";
4. We found that emotional appeal should be annotated in its own right; it is a subdivision of classic rhetoric as is argumentation;
5. We included conspiracy theories, which came up as a variable in migration data and then in political debates.

Thus, we decided it was necessary to review criteria used in several other domains known for their work on counter-speech, i.e., the fields of data mining (for example, Fortuna & Nunes, 2018), psychology (Stephan, 1999; Stephan et al., 1999), discourse analysis and rhetoric (Baider, 2019, 2020; Wodak, 2015).

Eventually, we determined a new set of criteria, which we used to work on another set of annotations. The criteria now included:

1. The specific topic, since we had three data sets;
2. A rhetoric category, subdivided into argumentation (and further divided into logic or reasoning, statistics, examples, history, and other facts) and emotional appeal or affective rhetoric, to allow us to consider the affective dimension of rhetoric (subdivided into insult, personal attack, empathy with acknowledging grievances, displaying positive emotions, displaying negative emotions, and sarcasm);
3. A multimedia category, to identify the role of sharing links; we also included images and emoticons in this category;
4. An impact category (as mentioned above, this comprises the number of comments and the general tone of comments; however, we also annotated the tone of the counter-speech to have some correlation with the tone of answers to that specific counter-speech);
5. A notation of to which comment the counter-speech is addressed, so that at a later stage we can further correlate the variables of the comment and the counter-speech.

Weekly monitoring of the annotator results ensured consistency and coherence in the process. From a database of 14,861 comments, 1,500 (10%) were annotated as counter-narratives (Table 2).

4. Detailed Discussion of Results

4.1. Rhetoric Used in Counter-Speech Across Topics

In the next section are a number of graphs summarising our results. We keep the original (mis)spelling in the quotations. The first column of all graphs gives the statistics referring to the migration database (MIG); the second gives the results for the political issues database (POL); the third is for the LGBTQ database. The number before these abbreviations refers to the number of the comments in our database. In this section, we look at our results in terms of the questions we posed initially: What counter-strategy strategies are most effective? Are some strategies more effective for certain topics?

4.1.1. Use of Argumentation

The strategic use of reasoning is high across all topics discussed, with an average of 78% for all categories, and with the political issues database displaying the highest percentage (86%). This result is surprising, considering that most research into online speech has found high spontaneity and a lack of control (Herring et al., 2002; Yus, 2011). In response to this seeming inconsistency, we might suggest that those who engage in counter-speech will be less prone to outbursts in expressing their views, as perhaps they will have been educated or trained in the use of counter-narratives. The following examples show some of the reasoning strategies we found in the MIG and POL databases:

> You could be made a refugee at some point in your life, have some compassion, madam, or best stay silent. (a; 37 MIG)

> Sir, leave aside the jokes; state the truth, and use face masks and protect yourself and those near and dear to you from this harmful virus that is spreading. (b; 49 POL)

In (a) the argumentation uses the reversal of role tactic: "You could be in their shoes." In (b) several words of

Table 2. Numbers of annotated counter-narratives.

No. of comments analyzed	No. of comments annotated for triggering overt or covert hate speech	No. of annotated counter-narratives
14,861	9,700	1,500

advice are offered in an effort to halt a participant's sarcasm. The comment in (b) refers to a video posted by the government and its recommendations for Covid-19.

Statistics are predominant in the migration dataset (11%), which can be expected (Figure 1): The fact that the number of migrants or foreigners is hugely exaggerated is well-known (Wohlfeld, 2014). In the following quotation (c), the participant gives numbers to explain the plight of the asylum seekers from Syria and dismisses the stereotype of young men migrating to Europe from the Middle East with statistical facts:

> Most of them settled in Lebanon? You make it sound so nice. [In fact] 70% of the 1 m[illion] Syrian refugees in Lebanon live below the poverty line…36% of those entering Europe are children and women. But in your eyes even Muslim children are violent. (c; 77 MIG)

We find the use of examples, which are generally personal experiences, more prominent in the migration and LBGTQ datasets (Figure 2). The "history" argument is also quite common and is used to counter the idea that Muslims are prone to violence or that they condone ISIS violence, as in (d), or to counter LGBTQ stereotyping, as in (e):

> As a British Muslim, [I say to] people commenting on this, I thought I'd tell you all this has nothing to do with Islam! Islam goes against the killing of all innocent people! So do not think this is Islam! (d; 1466 MIG)

> I'm transgender. I served 14 years active duty. And I didn't do it for any purpose except to Serve My Country. (e; 510 LGBTQ)

In (f), by pointing out the history of the Vikings, the counter-argument dismisses the preceding allegations that Westerners bring civilisation to the countries they invade, and that today migrants only take advantage of the host nations. In (g), in response to hateful messages about gay marriage because of religious principles, the counter-speech argues that religion is a historical artifact and is based on tradition rather than truth or fact:

> You mean the Vikings who spent most of their time abroad raping and stealing all the goodies? And also took over a lot of other countries. :D :D (f; 4947 MIG)

> Religion is just a mix of history and the world in terms that humans can understand, especially for filling in the gaps. If an undiscovered tribe saw a helicopter fly over, they'd call it a flying beast because birds are the only thing they have to compare to. It wasn't any different 2000 years ago. It's just that now it's to do with tradition more than anything else. (g; 166 LGBTQ)

4.1.2. Use of Affect

Regarding the use of affect, Ernst et al. (2017) noted the degree of hostility often found in counter-speech. Our results point to a level of hostility and negativity in counter-speech equal to that of hate speech. If we

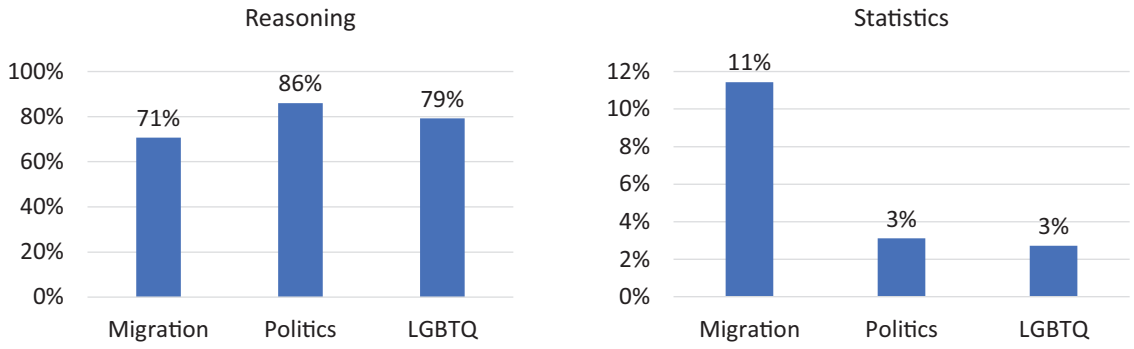


Figure 1. Reasoning vs. statistics used as arguments for the three topics.

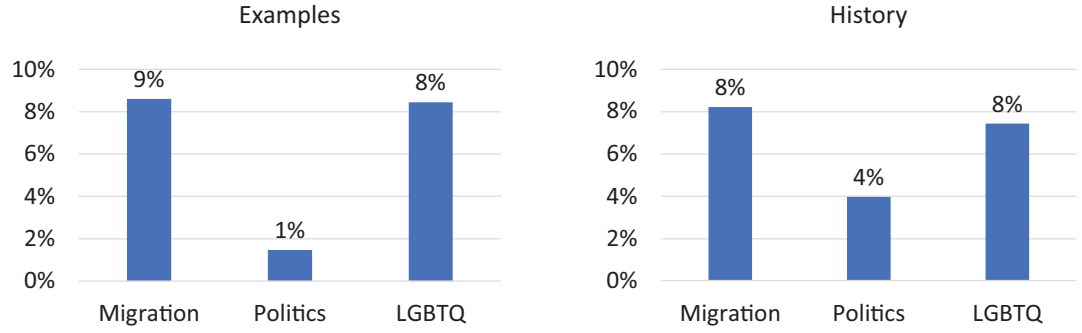


Figure 2. Examples and history used as arguments for the three topics.

group together the negative categories (insults, personal attacks, sarcasm, and display of negative emotions), we obtain an average of 73% for all topics of hostile counter-speech; respectively, 77% for migration, 66% for politics, and 76% for LGBTQ issues.

How hostility is evidenced, however, is specific to the topic debated (Figure 3): Insults are rare in political debates, while sarcasm prevails when countering LBGTQ "bashing" (33%), as seen in the following quotations:

> Have you been drinking? (20 LGBTQ)

> God can come have a chat with me about what people should have anger about [responding to a previous comment referring to the wrath of God]. (63 LGBTQ)

> Too absurd to bother commenting on. Number of gay people who have died as a result of the wrath of god 0. Find a sensible argument or give up. (73 LGBTQ)

Other signs of hostility, such as personal attacks and insults, were found in our migration data (Figure 4):

> You only see what you want to see. Pretty much like an ostrich. (h; 13 MIG)

> That's the most self-centered statement I've heard all day. No. We will continue to discuss them until they're safe from harm, have food and can work. (i; 71 MIG)

Although previous research targeting online speech generally, and online counter-speech specifically, has underlined the negativity of the messages or posts (Ernst et al., 2017), we nevertheless noted that almost a quarter of interventions were managed in a positive way. The rhetoric surrounding political issues is more often positive (34%), in comparison to migration (22%) or LGBTQ issues (24%); the two latter topics are more emotional and involve fundamental values such as religious values.

A more successful counter-speech strategy involves presenting positive emotions, as in the following examples. The speaker in (j) tries to appease verbal violence against a woman in a video wearing a veil in support of Muslim women, and in (k) the writer tries to derail racist rants against Pakistan by suggesting that India can help, if only on humanitarian grounds.

> Listen guys….The woman just wanted to show some love and solidarity with Muslim women. Don't make a big deal out of it. (j; 3309 MIG)

> On humanitarian grounds alone, we can help Pakistan also [the topic is about India giving medicine against Covid-19 to Pakistan]. (k; 739 POL)

To summarise the main counter-speech strategies, our data show a predominant use of argumentation, even though we know that the specific topics are better served by other types of counter-speech. We have also observed a notably limited use of statistics or historical

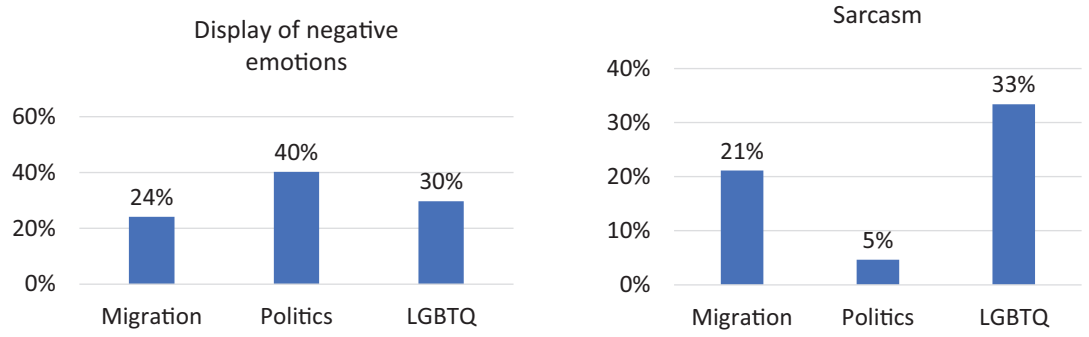


Figure 3. Hostile speech in counter-speech for the three topics.

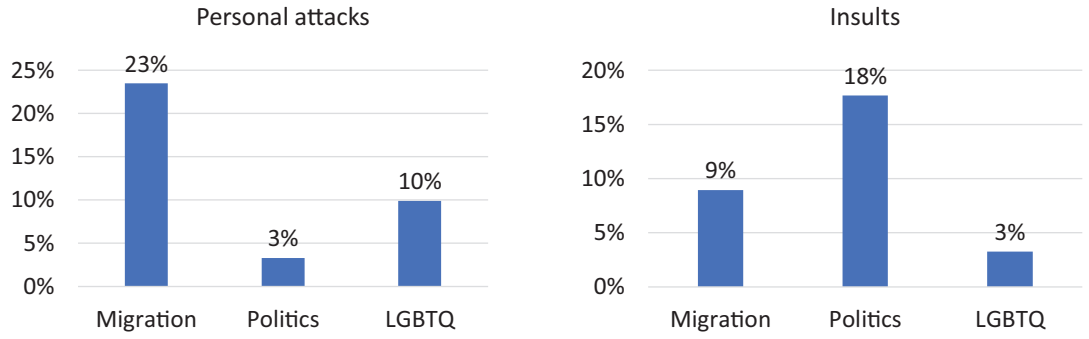


Figure 4. Personal attacks and insults used as arguments for the three topics.

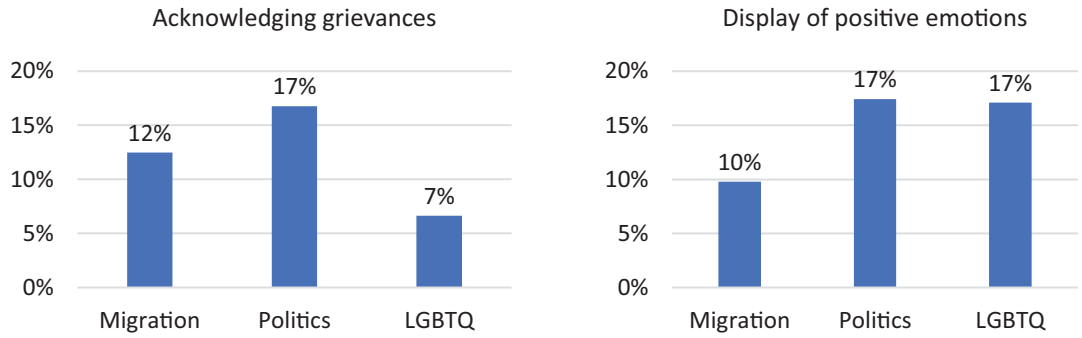


Figure 5. Positive comments used as arguments for the three topics.

examples when discussing political issues, but a high use of logical arguments. In contrast, the migration topic seems to favour the use of statistics, which should inform educational training programs in counter-speech: Knowing facts and statistics are important when responding to racist comments. We noted heavy use of sarcasm and personal attacks in responses to homophobic (and racist) comments.

The most-recorded strategy, whether in argumentation or use of affect, was a hostile stance. This is among the tactics recommended by Benesch et al. (2016), as it can make the commentator feel embarrassed because of their statement. The power of hostile comments to generate dialogue was confirmed in our study.

4.2. Impact of Counter-Speech Strategies

As explained in our methodology section, we evaluated the success of counter-speech strategies in relation to two variables and in line with the literature (de Latour et al., 2017; Silverman et al., 2016): (a) whether the comment initiated a dialogue, i.e., triggered a response (whether positive or negative is not important); and (b) what the tone of the responses triggered by the counter-speech was.

4.2.1. Number of Answers on Average

In broad terms, almost half the counter-speech strategies generated comments for all categories considered; moreover, we found no statistically significant difference among the three datasets concerning the impact of counter-speech. Therefore, we give the average numbers across all topics (see Table3).

Table 3. Average number of answers to counter-speech.

Number of answers to counter-speech	Percentage
0	46%
1–5	51%
5–10	0%
More than 10	0%

4.2.2. Correlations Between Strategies and Number of Answers

We found no evidence of a statistically significant relationship between the different argumentative strategies and generating a dialogue (i.e., triggering several answers between 1–5), although some emotional appeal (affective) strategies were found to be correlated with generating a dialogue.

We found a positive correlation between generating a dialogue and using personal attacks, while in contrast, counter-speech that is less aggressive, e.g., displaying positive emotions and acknowledging grievances, is less likely to generate a dialogue.

Additionally, the correlation coefficients suggest that there is no real relationship between generating dialogue and the strategies of using insults, sarcasm, or displaying negative emotions.

4.2.3. Impact of the Tone Used in Counter-Speech

We found that the tone of the counter-speech may influence the tone of the response: There is a positive correlation between a negative tone in the counter-narrative and a negative tone in the response; however, there is no correlation between a negative tone in the counter-narrative and a positive tone in its response. These correlations suggest that since most counter-narratives are classified as having a negative tone, they will generate a dialogue with negative answers.

In contrast, a counter-narrative with a positive tone correlates with a positive tone in the response. This finding indicates that, although counter-speech that is positive in tone is less likely to generate a dialogue, when it does, the resulting exchange is likely to be positive.

Counter-narratives that are classified as positive in tone are more likely to use statistical facts as part of their argumentative rhetoric, and to acknowledge grievances as part of their affective rhetoric (Table 4). We can observe in Table 4 a positive correlation between acknowledging grievances and generating answers as well as between acknowledging grievances and using a positive tone.

Table 4. Correlations for the strategy of acknowledging grievances.

Correlations for acknowledging grievances	
Number of answers	27%
Positive tone	71%

We conclude, therefore, that if a counter-speech aims to generate a positive dialogue, it would be most effective if it used facts, statistics, or acknowledged grievances while avoiding a hostile tone.

4.3. Summarising Results

Our first research question examined the counter-speech strategies used in situ in a large sample. Our analysis revealed that the majority of counter-narratives expose or ridicule the authors of offensive comments. We found this to be true for the majority of narratives, which used argumentative strategies (70% of the chosen strategies, on average for the three topics under investigation), as well as for the 30% of strategies that used affective rhetoric, where most narratives were highly negative, featuring insults, personal attacks, and negative emotions. We found that although humor is seldom used, sarcasm is prevalent in both covert hate speech and counter-speech. Thus, whether argumentative or affective, these strategies exacerbate verbal violence and fuel the negativity, further polarising the debate. Importantly, analysis of our findings led us to conclude that the particular strategy selected in counter-arguments is highly influenced by the social context as well as by the topic under discussion.

Our second research question measured, from a quantitative perspective, the impact of the various strategies on the different topics. The general results reveal that dialogue is rarely sparked: most often the counter-speech is ignored. Moreover, in the case that dialogue is generated by counter-speech, it is usually because of its hostile tone, especially if it contains a personal attack. It would appear that positive dialogues, the ultimate aim of counter-speech, are only generated by acknowledging grievances or displaying positive emotions, two strategies that are not often encountered in online heated debates.

5. Concluding Remarks

In summary, our results indicate that the tone of the counter-narrative is highly important and should be the first consideration when responding to hate speech. In contrast, we found, in our data, that most counter-speech took a hostile tone, and although this is a strategy recommended by Benesch et al. (2016), our results show that this is ineffective: It only puts the "opponent" on the defensive and often leads to continued verbal violence. Our results, therefore, confirm a number of earlier studies that found hateful posts were most often responded to with disagreement, conflict, and derision (Bartlett & Krasodomski-Jones, 2016; Maity et al., 2018). Nevertheless, we did identify some argumentative strategies that led to a positive outcome: the use of historical facts and/or personal examples correlated with generating dialogue, even when the tone was negative. Positive-toned responses—which we consider a marker of the effectiveness of a counter-narrative—resulted when the comment acknowledged the writer's grievances or used a positive emotional tone. Yet we found that very few counter-narratives (on average 10%) used these strategies. We, therefore, recommend that educational training—even machine learning and empathetic bots—should focus on such strategies.

Acknowledgments

This article is based on data collected within the Innovative Monitoring Systems and Prevention Policies of Online Hate Speech (IMsyPP) project, co-funded by the European Union's Rights, Equality and Citizenship Programme (2014–2020). The project IMsyPP, grant no. 875263, is available at http://imsypp.ijs.si. The European Commission's support for the production of this publication does not constitute an endorsement of its contents, which reflect the views of the authors alone. The Commission cannot be held responsible for any use which may be made of the information contained in this article.

Conflict of Interests

The author declares no conflict of interest.

References

Ashida, M., & Komachi, M. (2022). Towards automatic generation of messages countering online hate speech and microaggressions. In A. Narang, A. Davani, L. Mathias, B. Vidgen, & Z. Talat (Eds.), *Proceedings of the sixth workshop on online abuse and harms* (pp. 11–23). Association for Computational Linguistics.

Askanius, T. (2021). On frogs, monkeys, and execution memes: Exploring the humor-hate nexus at the intersection of neo-nazi and alt-right movements in Sweden. *Television & New Media*, *22*(2), 147–165.

Baider, F. (2019). Le discours de haine dissimulée; le mépris pour humilier [Covert hate speech; using contempt to humiliate]. *Déviance et Société*, *43*(1), 71–100.

Baider, F. (2020). Pragmatics lost? Overview, synthesis and proposition in defining online hate speech. *Pragmatics and Society*, *11*(2), 196–218.

Baider, F. (2022). Covert hate speech, conspiracy theory and anti-semitism: Linguistic analysis versus legal judgement. *International Journal for the Semiotics of*

Law, *35*, 1–25. https://doi.org/10.1007/s11196-022-09882-w

Bartlett, J., & Krasodomski-Jones, A. (2016). *Counter-speech on Facebook*. Demos. https://demosuk.wpengine.com/wp-content/uploads/2016/09/Counter-speech-on-facebook-report.pdf

Ben-David, A., & Fernández, A. M. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, *10*, 1167–1193.

Benesch, S. (2014). *Countering dangerous speech: New ideas for genocide prevention*. United States Holocaust Memorial Museum.

Benesch, S., Ruths, D., Dillon, K. P., Saleem, H. M., & Wright, L. (2016). *Counterspeech on twitter: A field study*. The Dangerous Speech Project. https://dangerousspeech.org/counterspeech-on-twitter-a-field-study

Bhat, P., & Klein, O. (2020). Covert hate speech: White nationalists and dog whistle communication on Twitter. In G. Bouvier & J. Rosenbaum (Eds.), *Twitter, the public sphere, and the chaos of online deliberation*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-41421-4_7

Braddock, K., & Horgan, J. (2016). Towards a guide for constructing and disseminating counternarratives to reduce support for terrorism. *Studies in Conflict & Terrorism*, *39*(5), 381–404.

Briggs, R., & Feve, S. (2013). *Review of programs to counter narratives of violent extremism. What works and what are the implications for government?* Institute for Strategic Dialogue. https://www.publicsafety.gc.ca/lbrr/archives/cn28580-eng.pdf

Brown, A. (2020). *Models of governance of hate speech*. Council of Europe.

Burnap, P., & Matthew, L. W. (2015). Cyber hate speech on Twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy Internet*, *7*(2), 223–242.

Citron, D. K., & Norton, H. (2011). Intermediaries and hate speech: Fostering digital citizenship for our information age. *Boston University Law Review*, *91*, 1436–1460.

Chung, Y. L., Kuzmenko, E., Tekiroglu,S., & Guerini, M. (2019). CONAN—COunter NArratives through Nichesourcing: A multilingual dataset of responses to fight online hate speech. In P. Nakov & A. Palmer (Eds.), *Proceedings of the 57th annual meeting of the Association for Computational Linguistics* (pp. 2819–2829). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1271

Council of the European Union. (2008). Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law. *Official Journal of the European Union*, *51*(L 328/55). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32008F0913

Culpeper, J. (2021). Impoliteness and hate speech: Compare and contrast. *Journal of Pragmatics*, *179*, 1–11 https://doi.org/10.1016/j.pragma.2021.04.019

Ernst, J., Schmitt, J. B., Rieger, D., Beier, A. K., Vorderer, P., Bente, G., & Roth, H. J. (2017). Hate beneath the counter speech? A qualitative content analysis of user comments on YouTube related to counter speech videos. *Journal for Deradicalization*, *10*, 1–49.

Fortuna, P., & Nunes, S. (2018). A survey on automatic detection of hate speech in text. *ACM Computing Surveys (CSUR)*, *51*(4), 1–30.

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000233231

Gelber, K. (2019). Differentiating hate speech: A systemic discrimination approach. *Critical Review of International Social and Political Philosophy*, *24*(4), 393–414. https://doi.org/10.1080/13698230.2019.1576006

Gemmerli, T. (2015). *The challenges of propaganda war: A guide to counter-narratives in the prevention of radicalisation* (Policy Brief). Danish Institute for International Studies. https://www.diis.dk/node/6900

Herring, S., Kirk, J. S., Scheckler, R., & Barab, S. (2002). Searching for safety online: Managing trolling in a feminist forum. *Information Society*, *18*, 371–384.

Hietanen, M., & Eddebo, J. (2022). Towards a definition of hate speech with a focus on online contexts. *Journal of Communication Inquiry*. Advance online publication. https://doi.org/10.1177/01968599221124309

Hill, J. (2008). *The everyday language of white racism*. Wiley Blackwell.

Konikoff, D. (2021). Gatekeepers of toxicity: Reconceptualizing Twitter's abuse and hate speech policies. *Policy Internet*, *13*, 502–521.

de Latour, A., Perger, N., Salaj, R., Tocchi, C., & Viejo Otero, P. (2017). *WE CAN! Taking action against hate speech through counter and alternative narratives*. Council of Europe.

Leroy v. France, 36109/03 (2008).

Maity, S. K., Chakraborty, A., Goyal, P., & Mukherjee, A. (2018). Opinion conflicts: An effective route to detect incivility in twitter. In K. Karahalios, A. Monroy-Hernández, A. Lampinen, & G. Fitzpatrick (Eds.), *Proceedings of ACM on Human-Computer Interaction* (Vol. 2, pp. 1–27). Association for Computing Machinery.

Matsuda, M. (1989). Public response to racist speech: Considering the victim's story. *Michigan Law Review*, *87*(8), 2320–2381.

McGowan, M. K. (2009). Oppressive speech. *Australasian Journal of Philosophy*, *87*(3), 389–407. https://doi.org/10.1080/00048400802370334

Munger, K. (2017). Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Political Behavior*, *39*(3), 629–649.

Musolff, A. (2015). Dehumanizing metaphors in UK immigrant debates in press and online media. *Journal of Language Aggression and Conflict*, *3*(1), 41–56.

Norwood v. the United Kingdom, 23131/03 (2004).

Office for Democratic Institutions and Human Rights (2009). *Hate crime laws.* https://www.osce.org/odihr/36426

Pingen, A. (2021). *EU Commission 6th evaluation of code of conduct on countering illegal hate speech online*. Eucrim. https://eucrim.eu/news/commission-6th-evaluation-of-code-of-conduct-on-countering-illegal-hate-speech-online

Richards, R. D., & Calvert, C. (2000). Counterspeech 2000: A new look at the old remedy for "bad" speech. *BYU Law Review*, *2*(2), 553–586.

Sap, M., Dallas, C., Saadia, G., Choi, Y., & Smith, N. A. (2019). The risk of racial bias in hate speech detection. In P. Nakov & A. Palmer (Eds.), *Proceedings of the 57th annual meeting of the Association for Computational Linguistics* (pp. 1668–1678). Association for Computational Linguistics.

Schieb, C., & Preuss, M. (2016, June 9-13). *Governing hate speech by means of counterspeech on Facebook* [Conference session]. 66th ICA annual conference, Fukuoka, Japan.

Silva, L., Mondal, M., Correa, D., Benevenuto, F., & Weber, I. (2016). Analyzing the targets of hate in online social media. *Proceedings of the International AAAI Conference on Web and Social Media*, *10*(1), 687–690.

Silverman, T., Stewart, C. J., Amanullah, Z., & Birdwell, J. (2016). *The impact of counter-narratives. Insights from a year-long cross-platform pilot study of counter-narrative curation, targeting, evaluation, and impact*. Institute for Strategic Dialogue. https://www.isdglobal.org/wp-content/uploads/2016/08/Impact-of-Counter-Narratives_ONLINE_1.pdf

Stephan, W., Stephan, C., & Gudykunst, W. (1999). Anxiety in intergroup relations: A comparison of anxiety/uncertainty management theory and integrated threat theory. *International Journal of Intercultural Relations*, *23*(4), 613–628.

Strossen, N. (2018). *Hate: Why we should resist it with free speech, not censorship*. Oxford University Press.

Tuck, H., & Silverman, T. (2016). *The counter-narrative handbook*. Institute for Strategic Dialogue. https://www.isdglobal.org/wpcontent/up

United Nations. (1965). *International convention on the elimination of all forms of racial discrimination.* https://www.ohchr.org/sites/default/files/cerd.pdf

United Nations. (1966). *International covenant on civil and political rights*. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights

United Nations. (2012). *Rabat plan of action* (HRC/22/17/Add.4). https://www.ohchr.org/sites/default/files/Rabat_draft_outcome.pdf

Weaver, S. (2016). Humor and race. In J. Stone, R. M. Dennis, P. S. Rizova, A. D. Smith, & X. Hou (Eds.), *The Wiley Blackwell encyclopedia of race, ethnicity and nationalism* (pp. 1–3). John Wiley & Sons.

Wodak, R. (2015). *The politics of fear: What right-wing populist discourses mean*. SAGE.

Wohlfeld, M. (2014). Is migration a security issue? In O. Grech & M. Wohlfeld (Eds.), *Migration in the Mediterranean: Human rights, security and development perspectives* (pp. 61–77). Mediterranean Academy of Diplomatic Studies.

Wright, L., Ruths, D., Dillon, K. P., Saleem, H. M., & Benesch, S. (2017). Vectors for counterspeech on Twitter. In Z. Waseem, W. Hui Kyong Chung, D. Hovy, & J. Tetreault (Eds.), *Proceedings of the first workshop on abusive language online* (pp. 57–62). Association for Computational Linguistics.

Yus, F. (2011). *Cyberpragmatics. Internet-mediated communication in context*. John Benjamins.

About the Author

Fabienne Baider is a professor of linguistics and gender studies at the Department of French and European Studies, University of Cyprus. Her current research focusses on online hate speech from a linguistic perspective. Her recent work on this topic has been published in the *Handbook of Intercultural Pragmatics*, *Pragmatics and Society* as well as the *International Journal for the Semiotics of Law*.

cogitatio

Politics and Governance (ISSN: 2183–2463)
2023, Volume 11, Issue 2, Pages 261–271
https://doi.org/10.17645/pag.v11i2.6550

Article

A Corpus-Based Discourse Analysis of Liberal Studies Textbooks in Hong Kong: Legitimatizing Populism

Yulong Li [1,2], Yuxi Wu [3,*], and Mingfeng Xiao [1]

[1] School of Education, City University of Macau, Macau
[2] School of Education, University of Glasgow, UK
[3] Faculty of Humanities and Social Sciences, City University of Macau, Macau

* Corresponding author (jessiewu@cityu.mo)

Submitted: 30 November 2022 | Accepted: 24 January 2023 | Published: 17 May 2023

Abstract
Researchers have discussed Hong Kong's localist identities, nativist sentiments, and populism, but have not widely examined the extent to which populism could be perceived in education in Hong Kong. As the chief participants of the Occupying Central and the radical Anti-Extradition Bill movements in Hong Kong were students, this suggests the need to explore the relationship between populism and education, particularly the then-controversial liberal studies textbooks. According to contemporary news reports, liberal studies textbooks contained much content stigmatising the Chinese mainland. Previous studies of liberal studies textbooks applied qualitative discourse analysis methods. In this study, mixed-method analysis was applied to a specialised corpus comprising seven commercial liberal studies textbooks containing 248,339 Chinese characters in total to explore the extent to which liberal studies textbooks contain information concerning the key features of populism—the heightened division between the inner and outer groups. A division was found between positive images of Hong Kong and negative images of China in the narratives of commercial liberal studies textbooks. Accordingly, the textbooks can be understood to contain populism. The present study advocates that relevant educational watchdogs in Hong Kong provide more guidance on the writing and publishing of liberal studies textbooks in the future, keeping the enquiry-based spirit of the liberal studies course fulfilled, and urges stakeholders of Hong Kong education to consider teaching peace education and developing a more inclusive environment.

Keywords
corpus linguistics; Hong Kong; liberal studies; populism; textbook

Issue
This article is part of the issue "Hate Speech, Demonization, Polarization, and Political Social Responsibility" edited by Luis M. Romero-Rodríguez (Rey Juan Carlos University), Pedro Cuesta-Valiño (University of Alcalá), and Bárbara Castillo-Abdul (Rey Juan Carlos University).

1. Introduction

Since 2014, when the Occupying Central movement broke out in Hong Kong, researchers have discussed Hong Kong localist identities and activities, nativist sentiments, and populism. Synonyms for these phrases increased until the radical Anti-Extradition Bill movement in 2019 (e.g., Chow et al., 2020; Chun, 2020; Li & Xiao, 2020; Lowe & Tsang, 2017; Ma, 2018; Ng & Kennedy, 2019; Sautman & Yan, 2015; Veg, 2017; Vukovich, 2020; Zamecki, 2018). However, scholars have not widely examined the extent to which the influence of populism could be perceived in education in Hong Kong. It is noteworthy that the chief participants in the Occupying Central and the radical Anti-Extradition Bill movements were university and pre-tertiary school students (see Li & Wu, 2022) and student unions. For example, the student union at the University of Hong Kong was among the radical populist groups (Ng & Kennedy, 2019), indicating the need to explore the possible seepage of populist culture in

some educational fields of Hong Kong. The present study adopts a critical perspective on education, employing Bourdieu's (Bourdieu & Passeron, 2000) theory of education as a field of cultural arbitrariness, with textbooks perceived as a medium of cultural arbitrariness, rather than couriers of objective truth.

Accordingly, the present study examines textbooks for the mandatory liberal studies (LS) course in Hong Kong secondary schools for students wishing to enter university. Launched in 2009 by the Education Bureau (EDB), the LS course was a symbol of Hong Kong's educational system reform, transitioning from the British mode of 3–2–2–3 (three-year junior secondary school, two-year senior secondary school, two-year pre-university, and three-year university study) to the Chinese mode of 3–3–4 (three-year junior secondary school, three-year senior secondary school, and four-year university study; Li et al., in press). The LS course had six themes: "personal development and interpersonal relationships, Hong Kong today, modern China, globalization, public health, and energy technology and the environment" (Li & Wu, 2022, p. 130). The original purpose of the LS course was to nurture lifelong learners with critical thinking abilities and inclusive spirit by having students apply enquiry-based learning to social topics (Chiu et al., 2018). Since 2012, the LS course has been suspected of politicising students who participate in radical protests, despite the fact that Chiu et al. (2018) surveyed 2,896 students from 25 secondary schools in Hong Kong found that most students did not develop radical thinking and seldom participated in political activities.

However, the LS course became the focus of significant controversy in 2019. According to contemporary news reports ("Xiuli fengbo yizhounian," 2020), the textbooks contained much content stigmatising the Chinese mainland, potentially misleading students; some critics like Tu (2020) also reported that LS teachers in St Stephen's College had students comment in an exam on the advantages of participating in radical protests rather than evaluating the protesters' behaviour holistically. Noting that the LS course, designed to nurture students' critical thinking, in fact inculcated biases, Wu (2021) suggested that the EDB had delegated LS textbook composition to several commercial publishers and maintained no censorship over the textbooks before 2019. Li and Wu (2022) added that the EDB did not specify official LS textbooks (LST), allowing teachers to choose any textbook from commercial publishers. Such measures and attitudes of the EDB thus allowed bias to enter (Li & Wu, 2022). Although the EDB requested that commercial publishers voluntarily submit their LST for review, only eight publishers submitted their LST to the EDB in 2020 (Li & Wu, 2022).

To pinpoint possible biases in commercial LST, Li et al. (in press), Li and Wu (2022), and Wu and Li (2022) analysed several LST, identifying contrasts between "politically correct" Hong Kongers and "morally stained" Chinese mainlanders, the "developed and Westernised" Hong Kong and the "chaotic and back-watered" Chinese mainland, and "commendable" radical protests and "ineffective" peaceful demonstrations. These authors found that the textbooks highlighted the dangers of Hong Kong's assimilation to the Chinese mainland, inculcating in readers a rejection of almost everything about China. These findings echo those of Silberberg and Agbaria (2021) regarding the heightened division between the insider group (Hong Kongers) and the outsider group (those of the Chinese mainland) from a populist viewpoint, as well as Moffit and Tormey's (2014) descriptors of the key elements of populist style: the notion of crisis (depicting China's assimilation of Hong Kong in the LST) and disseminating "bad manners" in political discourse (commending radical protests in the LST). These LST may thus be suspected of contributing to populist views.

However, earlier research on LST (e.g., Li et al., in press; Li & Wu, 2022; Wu, 2021; Wu & Li, 2022), although informative, applied qualitative discourse analysis, unavoidably "cherry-picking" texts that conveyed particular meanings (Mautner, 2007), rather than providing a verbatim overview of the texts. In this study, mixed-method analysis was applied to a specialised corpus, comprising seven commercial LST containing 248,339 Chinese characters to explore:

> RQ1: How were the semantic preferences and prosody of China and its synonyms manifested in the LST corpus?
>
> RQ2: How were the semantic preference and prosody of Hong Kong manifested in the LST corpus?

The answers to these research questions should facilitate a more detailed comparison of the depictions of the inner and outer groups.

2. Literature Review

2.1. Populism and Education

"Populism" derives from the Latin *populus*, "the people," although today populism is often referred to as a political doctrine (Heywood, 2013) centred around tensions between groups claiming to represent the will of the "pure people" against the "corrupt elite" or the regime (Laclau, 2005; Mudde, 2004). Bergmann (2020, p. 36) stated that populism appeals to people across social strata to move against the dominant elite, who are "siding with international actors against the nation and the people." However, Mudde (2004) highlighted the "thin-centred ideology" of populism, contending that populism parasitises other host ideologies, e.g., nationalism, socialism, liberalism, religion, racism, and even neoliberalism. Therefore, populism today is not only about the masses resisting the elite, but also describes the building of boundaries in the name of the masses to exclude the people (i.e., majority-insiders) from the

threatening others (i.e., minority-outsiders). Relevant examples include Israel, where religious populism led the majority insider-Jewish citizens to exclude Arab citizens (Silberberg & Agbaria, 2021); Zimbabwe, where neoliberal populism excluded anyone, including Zimbabwe nationals, with Western backgrounds in favour of insider-majority Black Africans with no Western education (Hwami, 2013); and the US, where Trump-led racist populism excluded Asian international students and immigrants from membership with insider-white majority US citizens (Kirby, 2021).

The relationship between populism and education is somewhat debated. Norris and Inglehart (2019) and Spruyt (2014) opined that people with higher education are less likely to hold populist opinions, and Saurette and Gunster's (2011, p. 199) explanation of the source of academic knowledge may indicate why highly educated people are unlikely to be populists, as academics belong to the elite. However, recent studies, including those by Hwami (2013) and Silberberg and Agbaria (2021), have demonstrated that populism has infiltrated higher and pre-tertiary education systems. Meanwhile, Hong Kong university students' participation in radical populist groups (Ng & Kennedy, 2019) and the divisive contrasts between Hong Kongers and Chinese mainlanders in LST (Li & Wu, 2022; Wu & Li, 2022) suggest that populism might have infiltrated some aspects of Hong Kong's educational system.

2.2. A Critical Perspective on Education: System and Textbook

Employing a critical perspective, Bourdieu and Passeron (2000) demonstrated how education, or what they termed "pedagogic action," could be understood as a system of symbols imposed by the dominant class or through cultural arbitrariness, whereby the culture of the dominant class not only becomes the legitimate form of knowledge accepted in society but devalues and marginalises the symbols and culture of the lower classes (Bourdieu & Passeron, 2000). Therefore, students whose family backgrounds and cultural symbols are closer to those of the dominant class experience smoother schooling than their peers from a lower social class (Bourdieu & Passeron, 2000). It is easier for students from higher than from lower social classes to achieve higher education; those from lower social classes must either abandon their innate cultural symbols for academic achievement or are eliminated from school through academic failure that locates them in their home culture (Bourdieu & Passeron, 2000). In this way, education systems participate in maintaining or reproducing existing social and power structures:

> [Pedagogic actions] tend to reproduce the system of cultural arbitrariness characteristic of that social formation, thereby contributing to the reproduction of the power relations which put that cultural arbitrary into the dominant position. (Bourdieu & Passeron, 2000, p. 10)

Bourdieu and Passeron (2000) characterise such education as a form of symbolic violence against people from the lower classes, indicating that a precondition for symbolically violent education to prevail is that people remain unaware of its nature through camouflage by pedagogic authorities. They identify two layers of pedagogic authority delegates: institutional (e.g., education ministries, bureaus, schools, and degrees) and vicarious—the teachers and the regulatory apparatus (syllabuses and textbooks). An example of an encounter with pedagogical authorities is that of a student who works hard on a subject from his/her admiration for a teacher, without considering the hidden cultural arbitrariness that the teacher may be delivering. Bourdieu and Passeron (2000) defined pedagogic authority as a default position that legitimises everything students learn in school.

Therefore, from a Bourdieusian perspective, education delivers cultural arbitrariness, as do textbooks. As Apple and Christian-Smith (1991) asserted, textbooks are a site of the reconciliation of social, cultural, and political power, and reflect their writers' preferences and interests. However, for ordinary students, textbooks are frequently perceived as absolute authoritarian and legitimate sources of knowledge (Gulliver, 2010).

3. Research Methodology

Corpus-based critical discourse analysis (CDA) is used here to examine the portrayal of China and its synonyms (e.g., the Chinese mainland, continent) and Hong Kong across the seven LST targeted in this study. Fairclough (2015) believed that the relationship between language and society is both internal and dialectical. The goal of CDA is to go beyond the text, considering sociocultural contexts (Bednarek & Caple, 2012). Fairclough's (1995) three-dimensional conception of discourse (i.e., text, discourse practice, and sociocultural practice) is adopted herein. However, as CDA has been criticised for its arbitrary selection of small-scale research texts (Fowler, 1996; Stubbs, 1997; Widdowson, 2000), cherry-picking (Baker & Levon, 2015; Mautner, 2007), or decontextualization of meaning, the corpus linguistics method is included in view of its empirical and quantitative potential (Cheng, 2013) to handle large-scale textual data. The present study should thus be understood, following Tognini-Bonelli (2001, p. 65), as corpus-based rather than corpus-driven, where the former involves "mak[ing] use of the corpus itself to expound, test, and exemplify theories and descriptions that were generated before a large corpus become available to inform language study." In contrast, the corpus-driven approach functions inductively, using the data and patterns observed in the corpus itself to identify regularities and exceptions (Baker, 2006). Baker and McEnery (2005, p. 223) point out that the

corpus-based approach has the advantage of affording discourse analysts a more complete and detailed understanding of the meanings and functions of certain word choices in texts. Corpus-based CDA is particularly effective for examining ideological issues in textbooks; a good example is the work of Hong and He (2015). Here, a corpus-based study seems appropriate for gaining a relatively detailed understanding of the meanings of the terms "China" and its synonyms and "Hong Kong" in the chosen commercial LST. Indeed, commercial LST were once reported to contain their writers' subjective preferences (Wu, 2021), thus not entirely embodying the intended enquiry-based learning spirit of the LS course.

In practice, a specialised corpus, named LST, was compiled from seven commercial LST, totalling 248,339 words. Convenience sampling was used in selecting the textbooks, although this inevitably risks introducing subjectivity. However, the authors were based elsewhere than Hong Kong, and as by the time the research was conducted, border travel was restricted due to Covid-19, they only had access to the LST in their local libraries. When selecting which LST to include in the corpus, the first criterion was that it considers the theme of "Hong Kong today" or "modern China," because these two themes of LST have been found to be controversial (see Li & Wu, 2022; Wu, 2021). The second criterion was that different versions of a textbook from the same publisher (like the third and fourth *New Horizon Liberal Studies: Today's Hong Kong*, published in 2016 and 2020, respectively, by the Hong Kong Educational Publishing Company) should contain obvious revisions. Supplementary File 1 provides details of the LST corpus and information on the textbooks included.

The Corpus tool AntConc Version 3.59 (Anthony, 2020) was used to process the data and generate concordance lines containing the keywords. Each concordance line was analysed manually to identify the semantic preference and prosody of the keywords. Semantic preference refers to "the restriction of regular co-occurrences to items which share a semantic feature" (Sinclair, 2004, p. 142) and semantic prosody is the determiner of the meaning as a whole, expressing attitudinal and pragmatic meaning (Sinclair, 2004). Both semantic preference and prosody belong to Sinclair's (2004) descriptive model of the co-selection of lexical items. The present study examined lexical items because individual words do not create meaning; instead, "meaning resides in extended units" (Cheng, 2012, p. 154). The coding scheme of descriptions of Hong Kong and China's synonyms in the corpus is based on Chen (2014) and Ross and Caldwell's (2020) articulation of Martin and White's (2005) language appraisal system for interpersonal meaning at the level of discursive semantics. Of Martin and White's (2005) key elements of attitude, engagement, and graduation of the appraisal system, Chen (2014) particularly emphasised attitude in labelling texts' negative and positive meanings, and subdivided attitude into three semantic regions: affect, judgement, and appreciation. Affect is the polarity of emotional response of an expression, which Ross and Caldwell (2020, p. 16) exemplified thus: "Trump is calm/angry." Judgement presents a moral evaluation of behaviours, which Ross and Caldwell (2020, p.16) exemplified with "Trump is cool/weak." Appreciation is defined as the aesthetic quality of semiotic texts and natural phenomena, exemplified thus: "His speech was brilliant/horrible." Table 1 presents examples of the coded texts for affect, judgement, and appreciation in the present study, and lists the coding author's general judgement of the texts as positive or negative.

Table 1. Examples of text coding (highlighted words are indicators of positive or negative expressions).

Semantic regions of attitude	Sentence examples from the LST	Coding author's judgement
Affect	After the "*black-hearted* vaccines" outbreak in China in 2016, parents could not pursue the case and were eventually forced to protest on the streets, while some were jailed (Ming Pao Educational Publications Editorial and Advisory Board, 2020, p. 181)	Negative description
Judgement	In 2016, China's Gini coefficient reached 0.465, suggesting *a severe gap* between the rich and poor (Zang et al., 2019, p. 49)	Negative description
	To a certain extent, *the spirit of the rule of law* protects the civil rights of Hong Kong residents and motivates them to fulfil their civic obligations (Chan et al., 2013, p. 132)	Positive description
Appreciation	China has a *serious* acid rain problem (Wu et al., 2013, p. 89)	Negative description
	Hong Kong's culture has been influenced by Chinese and Western cultures such that it has developed into a *unique* local culture (Hung et al., 2020, p. 47)	Positive description
	The court system in Hong Kong is *clear*, its commercial rules are *thorough*, and there are laws governing how to invest (Ng et al., 2016, p. 112)	Positive description

4. Results

4.1. China and Its Synonyms

4.1.1. About China

Four primary semantic preferences and prosodies were identified across 289 instances in the concordance lines relating to China. The first was "China's environmental issues" (189 out of 289 instances, 65.39%), as China was co-selected with phrases related to several types of environmental pollution, such as "environmental pollution," "sea water pollution," "groundwater pollution," "air pollution," "water pollution," and "the largest polluter," showing that China's environmental issues are frequently discussed in the LST corpus. Regarding "China's environmental issues," "China" is found to co-occur with phrases and clauses with negative connotations (96 out of 189 instances, 50.79%), such as "China's looming environmental crisis," "the country with the most serious soil erosion," "disaster," "still deteriorating," "seriously damaged," "deterioration of ecological environment," "worsening pollution problem," "severe pollution of underground water," "deteriorating air pollution," "acid rain," and "a serious threat to the health of civilians." All the co-occurring phrases and clauses used with China in the LST corpus highlight China's severe environmental pollution; thus, the semantic prosody is unfavourable, confirming that China's environmental situation is depicted negatively in the LST. Some examples are shown in Table 2. In contrast, only three concordances were found to mention China's environmental issues positively, representing 1.58% (three out of 189).

The second semantic preference regarding China in the LST corpus concerns "social issues in China," and represents 9.69% (28 of 289 instances), as evidenced by co-occurrence with phrases concerning social events in China. Corresponding semantic prosody concerning "social issues in China" was "negative" (17 out 28, 61.71%), as evidenced by its co-occurrence with phrases with negative meanings, such as "black heart/contaminated food," "black-heart vaccines," "unscrupulous merchants," "unsafe drinking water," "imbalance in gender ratio," and "negative social events," whose semantic prosody highlights negative aspects of China. Examples are given in Table 3.

The third semantic preference noted in reference to China relates to "China's economic development," amounting for 14.19% (41 out of 289) of instances with the surrounding phrases "China's economic development," with the semantic prosody of "problematic and full of challenges," accounting for 63.41% (26 out of 41 instances), e.g., "the side effect of population stagnation," "low birth rate," "facing challenges," "the emergence of a series of social problems," "disparity between the urban and rural areas," "regional disparity," "gap between the rich and the poor," "social polarization," "a series of environmental problems," "meeting a bottleneck," "having traditional culture shock," "corruption," "pose a serious problem," and "a massive waste of resources." The prosody of this language use shows China's "reform and opening up" led to many problems and challenges for China. Examples are presented in Table 4. By comparison, only five out of 41 concordance lines (12.20%) portray China favourably.

The final semantic preference relating to China is the "Chinese legal system," which represents 14.19% (31 out of 289) of the cases, as evidenced by co-selection with the words "legal system" and "judicial system." The corresponding prosody contends that the system is "flawed and needs to be improved," as evidenced by co-occurrences with phrases such as "full of challenges," "denial of justice," "many problems and shortcomings in the Chinese judicial system," "needs to be improved," "the rule of law has not yet become widespread," "China's judicial system is not complete," and "flawed

Table 2. Examples of negative descriptions of China's environment in the LST corpus.

1	*China's groundwater pollution* is also severe (Ming Pao Educational Publications Editorial and Advisory Board, 2020, p. 146)
2	In recent years, *air pollution in China* has been a significant issue, with sandstorms ravaging the nation, with the first resulting from preventing and controlling sandstorms. However, a new issue has emerged (Ming Pao Educational Publications Editorial and Advisory Board, 2020, p. 147)
3	China's economic recovery, which relies heavily on the steel, cement, and energy industries, has caused *air pollution* (Wu et al., 2013, p. 89)
4	*Water pollution* has also led to the emergence of "cancer villages" in China, with many residents around water sources drinking untreated sewage discharged into the rivers by enterprises upstream, resulting in large-scale cancer in villages (Zang et al., 2019, p. 141)
5	*Water pollution* has grave consequences for human life. Approximately 30% of China's freshwater is no longer drinkable and cannot even be utilised to irrigate farms; there is also a risk of agricultural product contamination (Zang et al., 2019, p. 141)
6	China has a serious *acid rain* problem (Wu et al., 2013, p. 89)

Note: Authors' English translations.

Table 3. Examples of negative descriptions of China's social issues in the LST corpus.

1	After the "*black-hearted vaccines*" outbreak in China in 2016, parents could not pursue the case and were eventually forced to protest on the streets, while some were jailed (Ming Pao Educational Publications Editorial and Advisory Board, 2020, p. 181)
2	In China, counterfeiters and low-quality goods are common, and *unscrupulous merchants* show no regard for people's health as it is not to their benefit (Zang et al., 2019, p. 67)
3	Three hundred and twenty million *Chinese citizens drink unsafe drinking water* (Wu et al., 2013, p. 88)
4	China's long-term economic development and overall national strength may be adversely affected by the decline in population. In contrast, other demographic issues such as an insufficient birth rate, a rising elderly population, and *an imbalance in gender ratio* will seriously undermine China's sustainable development (Ming Pao Educational Publications Editorial and Advisory Board, 2020, p. 91)

Note: Authors' English translations.

legal system." A further semantic prosody of the "legal system" is "needing improvement," only accounting for 54.83% (17 out 31 instances). Examples are presented in Table 5. Notably, only one out of 31 concordance lines (3.23%) mentioned the Chinese legal system positively.

4.1.2. About the Chinese Continent

In the LST corpus, we found two other synonyms for China: the Chinese continent and the Chinese mainland. "Chinese continent" appears 14 times in the LST corpus, co-occurring with negative phrases in 50% of cases (seven out of 14 instances): "the deteriorating human rights in China," "contaminated food is rampant on the continent of China," "against indiscriminate admission of continental students to universities in Hong Kong," "Chinese continent tourists' clamour," and "Hong Kongers fight the continentals in the metro." This reveals that the term "continent" is portrayed negatively. Examples are given in Supplementary File 2 and Table 6. However, no concordance lines represented "Chinese continent" positively.

4.1.3. About Mainland China

"Mainland China" is another term used to refer to China occurring 121 times in the LST corpus, tending to co-occur with words and phrases having predominantly negative connotations, representing 12.40% (15 out of 121 instances), such as "fake products," "unscrupulous

Table 4. Examples of negative descriptions of China's economic development in the LST corpus.

1	China's long-term economic growth and overall national strength may be influenced by a declining population and consequently *low birth rate* (Ming Pao Educational Publications Editorial and Advisory Board, 2020, p. 91)
2	An increasing *disparity between urban and rural areas* and unequal social development accompanies China's rapid economic expansion (Hung et al., 2020, p. 178)
3	*Corruption occurs endemically among the Chinese authorities*, and corruption cases continue unabated (Zang et al., 2019, p. 64)
4	In 2016, China's Gini coefficient reached 0.465, suggesting a severe *gap between the rich and poor* (Zang et al., 2019, p. 49)

Note: Authors' English translations.

Table 5. Examples of negative descriptions of China's legal system in the LST corpus.

1	Increased corruption has resulted from *China's flawed legal system* and the inadequate supervision of officials, who are not required to declare their assets (Zang et al., 2019, p. 80)
2	Coupled with *China's flawed legal system* and lack of routes of recourse, disputes can quickly lead to widespread rallies and demonstrations, which does not benefit social stability (Zang et al., 2019, p. 80)
3	The awareness of *the rule of law has not yet become widespread in China*, and many people still look to petitions and other channels to defend their rights (Zang et al., 2019, p. 132)
4	*Corruption among officials in China* is still a severe problem, and government officials are holding back the pace of judicial reform (Zang et al., 2019, p. 131)

Note: Authors' English translations.

Table 6. Examples of negative descriptions of Chinese continent in the LST corpus.

1	A social media page titled "Against *Indiscriminate Admission of Continental Students* to Universities in Hong Kong'' asks universities to restrict the number of mainland China students accepted (Ng et al., 2016, p. 208)
2	Some *continental Chinese guests were angry* because they could not obtain chips from the China Pavilion. In contrast, others *rushed into* the China Pavilion and *assaulted* the public security agents distributing chips. The situation was somewhat chaotic (Wu et al., 2013, p. 164)
3	The human rights situation in *continental China continued to deteriorate* in 2017 (Zang et al., 2019, p. 70)

Note: Authors' English translations.

businessmen," "fake drugs," "flaws exist in the mainland Chinese judicial system," "numerous reports of poisoned food in mainland China," "poisoned milk powder," "corruption in the mainland," "counterfeit tourist attractions," "poor petrol quality in the mainland," and "the mainland tourists behave in a uncivilized manner." The Chinese mainland is thus represented unfavourably in the LST corpus (Table 7). As with "Chinese continent," no corresponding concordance lines favourably portray the Chinese mainland.

4.1.4. About Chinese

In the 119 concordance lines relating to Chinese, we found an apparent semantic preference for "identity," accounting for 29.41% of cases (35 out 119 instances), as in phrases such as identity, multiple identities, and identity recognition. Analysis of the concordance lines shows that Chinese national identity and Hong Kong identity are juxtaposed for comparative purposes, indicating that identity issues are prominent in the LST corpus. The predominant semantic prosody is "contradictory," accounting for 37.14% (13 times out 35 instances), as in phrases such as "a strong contrast between the identity of being Hong Kongers and of being Chinese mainlanders," "the conflicts between Hong Kong people and Chinese," "the younger generations' sense of belonging toward Chinese national identity is weakened," "the younger generations' sense of belonging toward Chinese national identity is reduced," "therefore, they began to reject their Chinese identity," "the younger generations do not necessarily recognize their Chinese identity," "they hold a reserved attitude towards Chinese identity," "their recognition of their 'Chinese' identity is relatively weak," "Hong Kong inhabitants began to distance themselves from being Chinese," "make them distance themselves from being Chinese," "it inevitably reduced Hong Kong people's self-identification as Chinese," and "Hong Kong people tend to think of themselves as Hong Kongers rather than Chinese."

The semantic prosody of the Chinese people regarding "identity" in the LST corpus indicates a division between two identities (Chinese national identity versus Hong Kong identity) presented as contradictory. Chinese national identity is projected as distanced from Hong Kong's identity. A concordance analysis demonstrates that the LST corpus considers Hong Kong people's perceptions of their Chinese national identity to be negative, so much so that they have rejected it. This division is further underscored by emphasising Hong Kongers' unique identity as distinct from Chinese identity. This is illustrated in the following examples (see Table 8).

An LST cited an anonymous survey as a method to discount Chinese identity:

> A mainland website conducted a survey on "if there is an afterlife, would you like to be Chinese again?" The results showed that 65% of the respondents no longer wanted to be Chinese. What a surprise! We cannot imagine Chinese mainlanders rejecting their Chinese national identity. (Chan et al., 2013, p. 248)

By contrast, only one concordance line mentioned Chinese favourably, by quoting a positive view of Chinese identity.

4.2. About Hong Kong

Two primary semantic preferences and prosodies are found in 117 concordance lines mentioning *Hong Kong*.

Table 7. Examples of negative descriptions of Chinese mainland in the LST corpus.

1	There have been numerous reports of *poisoned food in mainland China* (Hung et al., 2020, p. 84)
2	*Corruption in mainland China* is a significant issue (Ming Pao Educational Publications Editorial and Advisory Board, 2020, p. 199)
3	Combined with the *low quality and high sulphur content of mainland China*'s gasoline, this has compounded the environmental damage caused by car emissions (Zang et al., 2019, p. 142)
4	*Income disparity among mainland Chinese* households is a severe problem (Wu et al., 2013, p. 61)

Note: Authors' English translations.

Table 8. Examples of negative descriptions of Chinese identity in the LST corpus.

1	Some young people believe it is impossible to create a trusting relationship with those in authority and propose severing links with China; as a result, their sense of *Chinese identity is eroded* (Hung et al., 2020, p. 181)
2	In recent years, the increasing number of Individual Visit Scheme visitors from mainland China has impacted the daily lives of Hong Kong residents and eroded *their sense of belonging* to a group with Chinese identity (Hung et al., 2020, p. 184)
3	In terms of economic development, Hong Kong is ahead of mainland China. The residents' superior financial standing has led them to view themselves as separate from the "Chinese" and to discriminate against and demean mainland Chinese, *preventing them from identifying as* "*Chinese*" (Ng et al., 2016, p. 220)
4	Before the handover, the British Hong Kong government systematically "removed the sense of national identity'' by infrequently discussing China's modern development in primary and secondary school curricula, which developed *feelings of unfamiliarity with Chinese identity* (Ng et al., 2016, p. 224)

Note: Authors' English translations.

First, *Hong Kong* often appears with phrases such as "law-ruled region," "the rule of law," and "judicial system," such that there is a semantic preference for "Hong Kong's judicial system" in 91 out 117 instances (82.05%), in which the dominant semantic prosody regards "Hong Kong's judicial system" as "solid and good" 57 times (62.64% of the 91 instances). It co-occurs with phrases with positive connotations, such as "autonomous judicial system," "good legal foundation," "fairness," "that Hong Kong ranks the top in terms of the rule of law," "the rule of law is the core value of Hong Kong," "Hong Kong people's evaluation of the law has maintained a high level for many years," "the government of Hong Kong attaches great importance to and actively upholds the rule of law," "high transparency in the judicial system of Hong Kong," "Hong Kong has always been a society ruled by law," "Hong Kong has a solid foundation in the rule of law and a sound judicial system," "the court system in Hong Kong is clear," and "making Hong Kong a fair and clean metropolis." Both Hong Kong and China share a semantic preference for "judicial system," but with different semantic prosody; while China's is "flawed and problematic," Hong Kong's is "good and solid." Examples are presented in Table 9. Only one out of 91 concordance lines (1.09%) portrays Hong Kong's legal system negatively: "I don't think Hong Kong can be called a society ruled by law as Hong Kong's laws cover the rich" (Chan et al., 2013, p. 114).

In addition, Hong Kong appeared with phrases associated with "culture," "different cultures," "unique local culture," "Western cultures," "diverse culture," "inclusive and diverse society," and "diverse cultural environment." This shows that Hong Kong's second semantic preference (26 out 117, 22.22%) is "Hong Kong's local culture." Of these, 17 instances (65.38%) carry a meaning of "diversity and inclusion," e.g., "Hong Kong, a confluence of Eastern and Western cultures, has formed a diverse, free, and pluralistic culture," "Hong Kong is a pluralistic and inclusive society," "Hong Kong has a pluralistic culture," and "Hong Kong's cultural diversity." Table 10 provides examples.

Table 9. Examples of descriptions of Hong Kong in the LST corpus.

1	Hong Kong has *an autonomous judicial system*, so its citizens have the law to follow, and the law constrains the power of those in authority (Hung et al., 2020, p. 103)
2	Hong Kong's *legal and judicial institutions maintain the rule of law*, and law enforcement agencies (such as the Police, Customs and Excise Department, Independent Commission Against Corruption, etc.) have collaborated to make Hong Kong a fair and clean metropolis (Chan et al., 2013, p. 114)
3	The court system in Hong Kong is *clear*, its commercial rules are thorough, and there are laws governing how to invest (Ng et al., 2016, p. 112)
4	According to the World Justice Project's Rule of Law Index, a non-governmental organisation, the international community *recognised* the legal situation in Hong Kong (Hung et al., 2020, p. 38)
5	To a certain extent, *the spirit of the rule of law* protects the civil rights of Hong Kong residents and motivates them to fulfil their civic obligations (Chan et al., 2013, p. 132)
6	Hong Kong has *a strong foundation in the rule of law and a robust judicial system*, which enabled it to maintain a high standard of the rule of law both before and after the handover (Ng et al., 2016, p. 32)

Note: Authors' English translations.

Table 10. Other example descriptions of Hong Kong in the LST corpus.

1	Hong Kong is an *inclusive and diverse* society (Chan et al., 2013, p. 266)
2	Regarding food, architecture, language, and religion, Hong Kong's *cultural diversity* offers a unique blend of the old and the new, the East and the West (Hung et al., 2020, p. 45)
3	Hong Kong's culture has been influenced by Chinese and Western cultures such that it has developed into *a unique* local culture (Hung et al., 2020, p. 47)
4	Hong Kong-style cafés are accessible to all classes in Hong Kong, and the food is a blend of cuisines, reflecting Hong Kong's *diverse* culinary culture (Hung et al., 2020, p. 47)

Note: Authors' English translations.

5. Conclusions

The LST corpus shows a division between positive images of Hong Kong and negative images of China's synonyms in the LST, corroborating Li et al. (in press) and Li and Wu (2022). The LST corpus analysis also revealed language discouraging the self-acknowledgement of Chinese identity, also echoing Li and Wu (2022), who asserted that LST directly called for Hong Kong people to distance their identity from that of mainland Chinese. These findings demonstrate a key feature of populism, the heightened division between the inner and outer groups (Silberberg & Agbaria, 2021). Therefore, the textbooks can be understood as containing populist elements that exclude the outer group, China, from the local inner group, Hong Kong. This echoes Ng and Kennedy's (2019, p. 112) framing of Hong Kong's emerging populism as localism-based "populist radical regionalism.''

By August 2020, eight LST publishers had submitted their LST to the EDB for revision, and these newly revised LST are now in use (Li & Wu, 2022), containing less language distancing Hong Kong from the Chinese mainland and including references to Chinese national role models, Nobel Prize winners, and Olympic champions (Li & Wu, 2022) to inculcate Chinese nationalism among Hong Kong students. The National Security Law of China has also been enacted in Hong Kong, and many radical populists, such as Joshua Chi-Fung Wong and Jimmy Chee-Ying Lai, have been prosecuted. Some Hong Kong tertiary institutes recently included the ceremony of raising the national flag and singing the national anthem in their weekly assemblies for the first time. Chinese nationalism seems to have won over populism in Hong Kong, but what is noteworthy is that nationalism is often a symbiont of populism (Mizushima, 2018), also creating a distinct inner group and outer group (Bergmann, 2020). Perhaps nationalism could be seen more clearly as quasi-populism. Thus, using one kind of populism to counter another, as appears to be happening now, will temporarily alleviate but likely not resolve chronic social cleavages in Hong Kong. Rather than developing anti-populist fantasies, the stakeholders of Hong Kong education are advised to develop peace education and a more inclusive environment. This is certainly an area that demands further study in the future. As the LS course was originally designed as an enquiry-based subject, encouraging students' inclusiveness and critical thinking (Chiu et al., 2018), biased textbooks do not represent the true spirit and original educational intention of the course; however, Bourdieu and Passeron (2000) warned that the textbook as a pedagogical authority serves to legitimise the content it delivers, even if the contents involve cultural arbitrariness. Therefore, it is equally important for educational watchdogs in Hong Kong to provide more guidance on the writing and publishing of LST in the future.

Acknowledgments

The authors would like to thank the journal editors and anonymous reviewers for their valuable suggestions and help.

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

Anthony, L. (2020). AntConc (Version 3.59) [Computer software]. Waseda University. https://www.laurenceanthony.net/software.html

Apple, M. W., & Christian-Smith, L. K. (1991). The politics of the textbook. In M. W. Apple & L. K. Christian-Smith (Ed.), *The politics of the textbook* (pp. 1–21). Routledge.

Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.

Baker, P., & Levon, E. (2015). Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. *Discourse & Communication*, *9*(2), 221–236.

Baker, P., & McEnery, T. (2005). A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *Journal of Language and Politics*, *4*(2), 197–226.

Bednarek, M., & Caple, H. (2012). *News discourse*. Continuum.

Bergmann, E. (2020). Understanding nativist populism. In E. Bergmann (Ed.), *Understanding nativist populism: Neo-nationalism* (pp. 29–52). Palgrave Macmillan.

Bourdieu, P., & Passeron, J. C. (2000). *Reproduction in education, society, and culture*. SAGE.

Chan, J. W., Hung, S. F., Wong, T. H., & Ziu, G. F. (2013). *New century liberal studies—Today's Hong Kong* (2nd ed.). Ling Kee Publishing.

Chen, Y. (2014). Exploring the attitudinal variations in the Chinese English-language press on the 2013 air pollution incident. *Discourse & Communication*, *8*(4), 331–349. https://doi.org/10.1177/1750481314537574

Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Routledge.

Cheng, W. (2013). Corpus-based linguistic approaches to critical discourse analysis. In C. A. Chapelle (Ed.), *The encyclopaedia of applied linguistics* (pp. 1–8). Wiley-Blackwell.

Chiu, W., Ruan, Y., & Liang, Y. (2018). Tongshike ruhe yingxiang xianggang zhongxuesheng? [How did the liberal arts module influence secondary school students of Hong Kong?]. *Hong Kong and Macao Journal*, *1*(18), 63–73.

Chow, S. L., Fu, K. W., & Ng, Y. L. (2020). Development of the Hong Kong identity scale: Differentiation between Hong Kong "locals" and mainland Chinese in cultural and civic domains. *Journal of Contemporary China*, *29*(124), 568–584.

Chun, A. (2020). "Nativist" social movement to "post-colonial" democratization. *Arena Journal*, *2020*(1), 43–49.

Fairclough, N. (1995). *Critical discourse analysis: A critical study of language*. Longman.

Fairclough, N. (2015). *Language and power* (3rd ed.). Routledge.

Fowler, R. (1996). On critical linguistics. In C. R. C. Coulthard & M. Coulthard (Eds.), *Texts and practices: Readings in critical discourse analysis* (pp. 3–14). Routledge.

Gulliver, T. (2010). Immigrant success stories in ESL textbooks. *TESOL Quarterly*, *44*(4), 725–745.

Heywood, A. (2013). *Politics* (4th ed.). Palgrave Macmillan.

Hong, H., & He, X. (2015). Ideologies of monoculturalism in Confucius Institute textbooks: A corpus-based critical analysis. In X. L. C. Christiansen & C. Weninger (Eds.), *Language, ideology and education: The politics of textbook in language education* (pp. 90–108). Routledge.

Hung, C. L., Zau, G. K., & Heoi, D. M. H. (2020). *New horizon liberal studies—Today's Hong Kong* (4th ed.). Educational Publishing Company.

Hwami, M. (2013). The nativist turn and the crisis in university education in Zimbabwe. *International Journal of Education*, *5*(4), 122–144.

Kirby, D. (2021). Authoritarian populism and the rhetoric of neo-racist nativism in the United States: Implications for higher education internationalization. *Literacy Information and Computer Education Journal*, *12*(1), 3491–3495.

Kong, F. (2020, June 9). Xiuli fengbo yizhounian jingju liucheng weichengnianren shaonian jiudubeibu xiaoliusheng zaoju [The first anniversary of the Extradition Bill Amendment Disturbance: Police arrested one thousand and six hundred teenagers, young man was arrested nine times, a grade six school student got detained]. *Hong Kong 01*. https://www.hk01.com/社會新聞/483467/修例風波一周年-警拘千六未成年人-少年9度被捕-小六生遭拘4

Laclau, E. (2005). *On populist reason*. Verso.

Li, Y., & Wu, Y. (2022). Why are the liberal studies textbooks that stigmatized China spread in Hong Kong? A textual analysis from Foucauldian order of discourse. *International Critical Thought*, *12*(1), 116–132.

Li, Y., & Xiao, D. (2020). Mandarin discrimination in Hong Kong: Four mainland Chinese sojourn teachers' experience of symbolic violence. *Journal of Education Culture and Society*, *11*(2), 499–520.

Li, Y., Xiao, D., & Liu, X. (in press). On the problems and causes of the past versions of Hong Kong liberal studies textbooks for secondary schools—Taking the "Hong Kong Today" series as an example. *Studies on Hong Kong and Macau*.

Lowe, J., & Tsang, E. Y. H. (2017). Disunited in ethnicity: The racialization of Chinese mainlanders in Hong Kong. *Patterns of Prejudice*, *51*(2), 137–158.

Ma, N. (2018). Changing identity politics: The democracy movement in Hong Kong. In W. M. Lam & L. Cooper (Eds.), *Citizenship, identity and social movements in the new Hong Kong* (1st ed., pp. 34–50). Routledge.

Martin, J., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.

Mautner, G. (2007). Mining large corpora for social information: The case of *elderly*. *Language in Society*, *36*(1), 51–72.

Ming Pao Educational Publications Editorial and Advisory Board. (2020). *Ming Pao senior secondary liberal studies: Modern China*. Ming Pao Education Publications.

Mizushima. (2018). *Mincui shidai* [Populist era]. Eurasian Press.

Moffit, B., & Tormey, S. (2014). Rethinking populism: Politics, mediatisation and political style. *Political Studies*, *62*, 381–397.

Mudde, C. (2004). The populist zeitgeist. *Government and Opposition*, *39*(4), 542–563.

Ng, H. Y., & Kennedy, K. J. (2019). Localist groups and populist radical regionalism in Hong Kong. *China: An International Journal*, *17*(4), 111–134.

Ng, L. F., Wong, M. L., & Wong, T. C. (2016). *Aristo senior secondary liberal studies—Today's Hong Kong* (3rd ed.). Aristo Educational Press.

Norris, P., & Inglehart, R. (2019). *Cultural backlash: Trump, Brexit, and authoritarian populism*. Cambridge University Press.

Ross, A. S., & Caldwell, D. (2020). "Going negative": An APPRAISAL analysis of the rhetoric of Donald Trump on Twitter. *Language & Communication*, *70*, 13–27.

Saurette, P., & Gunster, S. (2011). Ears wide shut: Epistemological populism, argutainment and Canadian conservative talk radio. *Canadian Journal of Political Science/Revue Canadienne de Science Politique*, *44*(1), 195–218.

Sautman, B., & Yan, H. (2015). Localists and "locusts" in Hong Kong: Creating a yellow-red peril discourse. *Maryland Series in Contemporary Asian Studies*, *2*(1), 2–86.

Silberberg, R., & Agbaria, A. (2021). Legitimising populist education in Israel: The role of religion. *British Educational Research Journal*, *47*(2), 316–331.

Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. Routledge.

Spruyt, B. (2014). An asymmetric group relation? An investigation into public perceptions of education in-based groups and the support for populism. *Acta Politica*, *49*(12), 123–143.

Stubbs, M. (1997). Whorf's children: Critical comments on critical discourse analysis (CDA). *British Studies in Applied Linguistics*, *12*, 100–116.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins Publishing Company.

Tu, H. (2020). *Jiujiu huanbing de Xianggang* [Save the ill Hong Kong]. Chung Hwa.

Veg, S. (2017). The rise of "localism" and civic identity in post-handover Hong Kong: Questioning the Chinese nation-state. *The China Quarterly*, *230*, 323–347. https://doi.org/10.1017/s0305741017000571

Vukovich, D. F. (2020). A city and a SAR on fire: As if everything and nothing changes. *Critical Asian Studies*, *52*(1), 1–17.

Widdowson, H. G. (2000). On the limitations of linguistics applied. *Applied Linguistics*, *21*(1), 3–25.

Wu, W. (2021). Politics, textbooks, and the boundary of "official knowledge": The case of liberal studies in Hong Kong. *Pedagogy, Culture & Society*, *29*(4), 537–554.

Wu, Y., & Li, Y. (2022). How China and Chinese people are visually represented: The case of a series of liberal studies textbooks in Hong Kong. *Journal of Education Culture and Society*, *13*(1), 403–417.

Wu, D. O., Chan, G. S., Leong, J. M., Zeng, G. J., & Chan, B. M. (2013). *Longman new senior liberal studies: Modern China*. Pearson.

Zamecki, L. (2018). Hong Kong youth radicalization from the perspective of relative deprivation. *Contemporary Chinese Political Economy and Strategic Relations: An International Journal*, *4*(2), 451–485.

Zang, L. G., Lei, G. M., Chan, H. J., Lo, P. K., Ma, G. W., Zung, M. W., Zyu, C. W., & Ziu, G. F. (2019). *New thinking liberal studies: Modern China* (3rd ed.). Ling Kee Publishing.

About the Authors

Yulong Li is assistant professor at the School of Education, City University of Macau. He is also an EdD candidate at the School of Education, University of Glasgow. His research interest includes educational philosophy and sociology. Before moving to Macau, Yulong Li worked as a lecturer at the Education University of Hong Kong, where he obtained a PhD in applied linguistics in 2017.

Yuxi Wu is an assistant professor at Faculty of Humanities and Social Sciences, City University of Macau. She obtained her doctoral degree in applied language sciences from the Hong Kong Polytechnic University in 2019. Her research interests include corpus linguistics, critical discourse analysis, and translation studies. She has publications in the areas of applied linguistics and translation studies.

Mingfeng Xiao is a postgraduate student at the School of Education, City University of Macau. Her research interests include educational theory and discourse studies.

POLITICS AND GOVERNANCE
ISSN: 2183-2463

Politics and Governance is an internationally peer-reviewed open access journal that publishes significant and cutting-edge research drawn from all areas of political science.

Its central aim is thereby to enhance the broad scholarly understanding of the range of contemporary political and governing processes, and impact upon of states, political entities, international organisations, communities, societies and individuals, at international, regional, national and local levels.

cogitatio

www.cogitatiopress.com/politicsandgovernance