

Hate Speech and Covert Discrimination on Social Media: Monitoring the Facebook Pages of Extreme-Right Political Parties in Spain

ANAT BEN-DAVID¹

The Open University of Israel, Israel

ARIADNA MATAMOROS-FERNÁNDEZ

Queensland University of Technology, Australia

This study considers the ways that overt hate speech and covert discriminatory practices circulate on Facebook despite its official policy that prohibits hate speech. We argue that hate speech and discriminatory practices are not only explained by users' motivations and actions, but are also formed by a network of ties between the platform's policy, its technological affordances, and the communicative acts of its users. Our argument is supported with longitudinal multimodal content and network analyses of data extracted from official Facebook pages of seven extreme-right political parties in Spain between 2009 and 2013. We found that the Spanish extreme-right political parties primarily implicate discrimination, which is then taken up by their followers who use overt hate speech in the comment space.

Keywords: social media, hate speech, covert discrimination, extremism, extreme right, political parties, Spain, Facebook, digital methods

Introduction

The rise in the prevalence of hate groups in recent years is a concerning reality. In 2014, a survey conducted by *The New York Times* with the Police Executive Research Forum reported that right-wing extremism is the primary source of "ideological violence" in America (Kurzman & Schanzer, 2015, para. 12), and Europe is dazed by the rise of far-right extremist groups (Gündüz, 2010). Online, hate practices are a growing trend and numerous human rights groups have expressed concern about the use of the Internet—especially social networking platforms—to spread all forms of discrimination (Anti-Defamation League, 2015; Simon Wiesenthal Center, 2012).

Anat Ben-David: anatbd@openu.ac.il

Ariadna Matamoros-Fernández: ariadna.matamorosfernandez@qut.edu.au

Date submitted: 2015-01-20

¹A version of this study was conducted while at the University of Amsterdam. Thanks are extended to Judith Argila, Stathis Charitos, Elsa Matamoros, Bernhard Rieder, and Oren Soffer.

Copyright © 2016 (Anat Ben-David & Ariadna Matamoros-Fernández). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <http://ijoc.org>.

Propagators of hate were among the early adopters of the Internet and have used the new medium as a powerful tool to reach new audiences, recruit new members, and build communities (Gerstenfeld, Grant, & Chiang, 2003; Schafer, 2002), as well as to spread racist propaganda and incite violence offline (Berlet, 2001; Chan, Ghose, & Seamans, 2014; Levin, 2002; Whine, 1999). The rise in the popularity of social media such as Facebook and Twitter has introduced new challenges to the circulation of hate online and to the targeting thereof. In the past decade, legislation and regulatory policy were designed to address explicit hate speech on public websites, and to distinguish between the criminalization of hate speech and the protection of freedom of expression (Banks, 2010; Foxman & Wolf, 2013); now, social media, operating as corporate platforms, define what hate speech is, set the accepted rules of conduct, and act on them.

For extremists, then, the use of social media platforms such as Facebook means that they must adapt their practices to the platforms' terms of use. Consenting to Facebook's authentic identity policy and community standards implies that extremists can no longer post anonymously or upload explicit content, as was previously done on public websites. Victims of hate, also consenting to the platform's terms of use, may report content they consider harmful, but the platform unilaterally decides whether the reported content is considered hate speech and, accordingly, whether or not to remove it.

This study brings together the rise in the popularity of social media with the rise in the popularity of political extremism to consider the ways that overt hate speech and covert discriminatory practices circulate on Facebook despite the platform's policy on hate speech. We join with critical studies of social media, which argue that the corporate logic of these platforms, alongside their technical intrinsic characteristics (algorithms, buttons, and features), condition the social interactions they host, as well as effect broader social and political phenomena (Gerlitz & Helmond, 2013; Gillespie, 2010; Langlois & Elmer, 2013). Following the logic of actor-network theory, which assigns equal agency to human and nonhuman agents in explaining sociotechnical phenomena (Latour, 2005), we argue that the study of the circulation of hate and discrimination on Facebook should not be limited to content analysis or to analyzing the motivations of extremists and their followers in using the platform, but also assign equal agency to the platform's policy and its technological affordances, such as the "like," "share," "comment," or "report" buttons. Whereas content analysis may expose instances of *overt hate speech* that may be regarded as violating Facebook's community standards, we argue that mapping the ties between the platform's policy, its technological affordances, and users' behavior and content better explain practices of *covert discrimination* that are circulated through the networks of association and interactive communication acts that the platform hosts. We support our argument with a case study that analyzed political extremism on Facebook in Spain, a country where racism on the Internet is "alarmingly increasing" (European Commission Against Racism and Intolerance, 2011, p. 22) and where anti-Semitic and extremist expressions on social media have become a pressing issue for legislators (Garea & Medina, 2014). Specifically, we present a longitudinal multimodal content and network analysis of the official Facebook pages of seven extreme-right political parties between 2009 and 2013, which are known to have radicalized their discourse against immigrants since the beginning of the economic crisis in 2008 (Benedí, 2013).

The rest of the article is structured as follows: We begin by outlining the literature that addresses overt hate speech and the circulation of covert discrimination on Facebook as two sets of tensions: (a) the regulatory and legislative challenges in targeting hate speech in public websites vis-à-vis the corporate logic of Facebook's hate speech policy, and (b) the role of Facebook's technological affordances in shaping social and political behavior vis-à-vis user practices. We then introduce the case study of the extreme-right political parties in Spain and our analysis of their Facebook pages, and conclude with a theoretical discussion of the platform's agency in contributing to the circulation of covert discriminatory practices, despite its official policy on hate speech.

Facebook's Corporate Logic and the Targeting of Hate Speech

The Internet presents new challenges for tackling the spread of hate speech, a concern usually related to freedom of expression—a right that is often used by those who advocate hatred to legitimize offenses, particularly against minorities (Gascón, 2012). The challenge in determining the boundary between legitimate free speech and criminal hate speech is apparent in the First Convention on Cybercrime, which was signed in 2001 by the Council of Europe in collaboration with other countries (Council of Europe, 2001), but left the criminalization of hate speech to further debates (Council of Europe, 2003).

Nevertheless, the general norm on the Internet is minimal intervention (Moretón, 2012), and governments often delegate the control of users' content to the Internet service providers. Service providers' terms of service are more flexible than those established by law, seeking limited liability for what users say (Brown, 2010; Gillespie, 2010).

With the emergence of social media, hate groups have added platforms such as Facebook and Twitter to their communicative networks (Jamieson, 2009; Simon Wiesenthal Center, 2012). However, unlike the regulation of hate speech on websites by Internet service providers, social networking platforms enjoy greater freedom to decide whether and how to address expressions of hate (Citron & Norton, 2011). Facebook both sets and enforces the criteria for the removal of hateful content. In its terms of service agreement, Facebook users agree to "not post content that: is hate speech, threatening, or pornographic; incites violence; or contains nudity or graphic or gratuitous violence" (Facebook, n.d.b, Section 3, para. 7). Facebook's definition of hate speech is further specified in the platform's community standards, as "content that directly attacks people based on their: Race, Ethnicity, National origin, Religious affiliation, Sexual orientation, Sex, gender, or gender identity, or Serious disabilities or diseases" (Facebook, n.d.a, Section 2, para. 3). Although the platform's community standards specifically state that "organizations and people dedicated to promoting hatred against these protected groups are not allowed a presence on Facebook" (Facebook, n.d.a, Section 2, para. 3), it also distinguishes between humorous and serious speech, and advocates for the freedom to challenge ideas, institutions, and practices.

Critics of the platform argue against the lack of transparency of its content removal policy (Gillespie, 2012). Facebook encourages its users to report content they consider harmful under various criteria,² and the platform determines, based on a set of internal rules, whether or not the reported content violated its community standards (Crawford & Gillespie, 2014). However, the decision to remove or keep reported content is not explained to the users (Heins, 2014). Moreover, Facebook uses a country-specific blocking system that acts in accordance with each country's legislation in terms of removing undesirable pages. For instance, Nazi content is forbidden in Germany but allowed in the United States (Arellano Toledo, 2012). Social networking platforms thus play a significant role as cultural intermediaries because their capacity to decide what content should be allowed is a "real and substantive" intervention into our understanding of public discourse and freedom of expression (Gillespie, 2010).

Regulators' views are divided between those who predict that strict enforcement of the platforms' terms of use would be more effective than the law to fight hate practices online (Foxman & Wolf, 2013) and between those who argue the opposite (Banks, 2010; Citron, 2014; Rosen, 2011). Rosen (2011), for example, argues that "community standards will never protect speech as scrupulously as unelected judges enforcing strict rules about when speech can be viewed as a form of dangerous conduct" (p. 1537). As a result, the platform hosts controversial content it does not consider harmful. For example, although groups such as "Kill a Jew Day" were eventually removed from Facebook (Citron & Norton, 2011), it continues to host groups such as "Death to Islam," which, to date, has 868 members (see <https://www.facebook.com/groups/712317992197380/>).

We argue that although Facebook's terms of use function as a gatekeeper, the platform's corporate logic in deciding what content should be allowed (Grant, 2015; NewsOne, 2013) allows for the circulation of covert discrimination through its technological affordances and the communicative acts that it hosts, as described in the following section.

Platforms and Users: Equal Agency

Different scholars have studied the process by which digital technologies and their affordances affect users' self-presentation online (Ellison, Hancock, & Toma, 2012), their political tactics on social media (Jensen & Dyrby, 2013), and their racist practices. This literature accounts for the ways that technology shapes sociability from a user approach, in the sense that it describes how humans use the affordances of the medium to perform potential actions (Gibson, 1977). In the case of discriminatory practices, Nakamura (2014) examined how the sharing affordances of social media and image boards prompt users to mistakenly circulate racist visual content, and Rajagopal (2002) argued that the anonymity of the Web and the lack of censorship contribute to the dissemination of hate messages. Likewise, Warner and Hirschberg (2012) showed that extremists often modify their discourse online through deliberate misspellings or word choices, such as using *Zionists* instead of *Jews*, and Klein (2012) introduced an "information laundering theory" to explain the ways that hate groups legitimate their ideas

²The criteria are self-harm, bullying, and harassment; hate speech; violent graphic content; nudity and pornography; identity and privacy; intellectual-property-regulated goods; phishing; and spam. (Facebook, n.d.b., Section 3).

through a “borrowed network of associations” that spreads hate not only through words, images, and symbols, but also through “links, downloads, news threats, conspiracy theories, politics and even pop culture” (p. 428).

Similar practices are also found on Facebook, where users take advantage of the platform’s technological affordances to spread hate. For example, hate groups adapt a political discourse to legitimize anti-Semitic aims so as not to be removed from the platform (Oboler, 2008). The act of posting a link on Facebook that redirects the user to a webpage of extremist content could be interpreted as a deliberate strategy to circumvent the platform’s hate speech policy. Racist content materialized in a link can be published without being instantly detected by other users, filtering algorithms, or Facebook’s human editors. Hence, user practices also play a role in disguising hate speech.

At the same time, the technological shaping of sociability can also be studied from a critical perspective, by which objects and their technological affordances have equal agency as human actors (Latour, 2005). For example, Facebook’s features encourage people to “like” and “share” content, but its algorithms play an important role in the construction of sociability. Previous research has demonstrated that algorithms are far from being neutral and can discriminate (Datta, Tschantz, & Datta, 2015), support toxic cultures (Massanari, 2015), and enact a form of disciplinary power (Bucher, 2012).

Facebook’s algorithm EdgeRank, which tracks what users like and the links they click on, recommends similar information based on the user’s prior interests. Such algorithmic logic creates what Pariser (2011) describes as a “filter bubble” to refer to the increasing personalization of the Web. One consequence of such algorithmic logic is that a user’s racist behavior on Facebook triggers recommendations of similar content from the platform.

Moreover, previous research has shown that communication acts on social media simultaneously create new articulations at economic, technological, and cultural levels (Langlois & Elmer, 2013; Rieder, 2013). Gerlitz and Helmond (2013) explain that Facebook monetizes social interactions through its social plugins that are found across the Web, such as the like and share buttons. However, the platform’s decentralized presence on the Web not only creates economic profit to Facebook through the commodification of data, but also reorganizes communities and knowledge, and facilitates the creation of atmospheres in which users are prone to behave in a certain way than another. Similarly, Facebook’s community standards and policy on hate speech are also guided by the motivation to monetize interactions (Kuchera, 2014; Shepherd, Harvey, Jordan, Srauy, & Miltner, 2015).

We therefore argue that Facebook’s technological affordances and corporate logic not only change the dynamics of performances of hate online, but also contribute to the perception of hate rhetoric as legitimate information. Accordingly, the study of social media communication should focus on user practices and content, as well as on the specific features of the technology with which they interact. We support our argument with a longitudinal, multimodal content and network analysis of the Facebook pages of the extreme-right political parties in Spain, as discussed in the following sections.

The Spanish Extreme Right on the Web

The Spanish extreme right comprises several fringe parties that hold nostalgic fascist agendas, such as Alternativa Española (AES), Movimiento Social Republicano (MSR), Democracia Nacional (DN), and FE-La Falange. However, only two parties—España 2000 and Plataforma per Catalunya (PxC), whose agendas are more aligned with anti-immigration and new European right-wing populism—have gained electoral success in recent years (Casals, 2000).

Despite their relative marginal electoral success, these parties in recent years have used the Internet to reach new audiences, balance their underrepresentation in traditional media, and create a stronger sense of community (Caiani & Parenti, 2011; Jackson & Lilleker, 2009). One of the characteristics of such increased activity is the radicalization of the extreme-right parties' discourse against immigrants and minorities (Benedí, 2013).

Previous work on Spanish extreme-right groups and political parties online has focused on their media strategies in the digital media environment. Caiani and Parenti (2011) found that these groups commonly use the Web to launch anti-immigration campaigns and to spread messages in favor of the unity of Spain. The content on such websites generally expresses a nostalgic and closed political culture and contains hate symbolism. Sánchez Duarte and Rodríguez Esperanza (2013) found that España 2000 and DN pursued an ideological strategy on Facebook and that the public character of this social networking platform influenced their discourse, which is focused on anti-immigration and antiestablishment messages.

Our aim was to monitor the Spanish extreme-right activity on Facebook as a case study to conceptualize the circulation of covert discrimination on Facebook rather than account for its communicative strategy on this platform. Because at the time of study the majoritarian right-wing political party in Spain was Partido Popular (PP),³ a party with a Francoist legacy (Casals, 2000) that often sparks controversy (Ferrándiz, 2014; Sewell, 2013), our analysis of the Facebook pages of the extreme-right political parties includes a comparison with the page of the PP.

Method

In this study, we analyzed data extracted from the official Facebook pages of seven Spanish extreme-right political parties and from the official Facebook page of the governing party at the time of study, PP, between 2009 and 2013. Our methodological approach was guided by the work undertaken by Richard Rogers and the Digital Methods Initiative at the University of Amsterdam. Rogers (2013) sees the Web as a source of cultural and social information that requires methods suited for studying its specificity as a medium. The Digital Methods Initiative has developed tools that repurpose the Web's unique features (such as search engines, hyperlinks, and social media platforms) into methods for social research. Such tools aim to reveal both the social formations that are embedded within them and the influence of their technical specificities on the interactions they host (Rieder, 2013; Rogers, 2013).

³ PP was Spain's governing party between 2011 and 2015.

Our method combined network analysis and multimodal content analysis of text, images, and links (Kress & Van Leeuwen, 2001). We used different tools and analytical software for each type of analysis, and the data were interpreted using both distant and close reading methods, as explained next.

Data Collection and Data Set

We used the software tool Netvizz to retrieve data about content and interactions on specific Facebook pages in a manner similar to Facebook's own collection of data via its algorithms.⁴ The tool outputs data files in table and network formats, which include the interaction between content and users,⁵ and between the data and the technical properties of the platform, its systems, and the practices it hosts, such as likes and comments (Rieder, 2013).

The resulting data set included data retrieved from seven Facebook pages of the following Spanish political parties: España 2000 and PxC, two political parties with relevant seats in the municipal assemblies; MSR and AES, two parties with minor local representation; and the three parties that constitute the coalition La España en Marcha: FE-La Falange, Nudo Patriota Español (NPe), and DN.⁶ The Facebook page of the majoritarian party PP was also included in this investigation to compare whether there were similarities in instances of overt hate speech and covert discriminatory patterns between the pages of the extreme-right political parties and the governing party. For NPe, FE-La Falange, España 2000, and PP, the data set covered almost five years of activity. For the other parties that joined Facebook later, the data set included at least three years of interactions (see Table 1).

⁴Netvizz uses Facebook's API to gather users' information, retrieve the content shared on the public pages, and capture Facebook's organizing principles, such as the "like" data.

⁵Netvizz anonymizes user information.

⁶The extreme-right political party Alianza Nacional also constitutes the La España en Marcha, but was not included in the analysis because it joined Facebook in 2013. It should be noted that NPe is not a political party but an association.

Table 1. Features of the Studied Political Parties' Facebook Pages.

Political party	Date joined Facebook	Posts analyzed ^a	Likes ^b
Alternativa Española (AES)	April 21, 2011	4,079	1,778
Nudo Patriota Español (NPe)	January 22, 2010	1,185	1,868
Plataforma per Catalunya (PxC)	April 27, 2011	3,362	2,434
Movimiento Social Republicano (MSR)	October 18, 2011	1,692	2,364
FE-La Falange	June 25, 2009	1,529	2,948
Democracia Nacional (DN)	February 16, 2011	1,794	4,241
España 2000	March 13, 2009	1,997	8,605
Partido Popular (PP)	March 6, 2009	32,343	61,446

^aPeriod of analysis: 2009–2013. ^bFigures as of April 2014.

Textual Analysis: Word Frequencies and Co-occurrences

To examine overt hate speech on the studied Facebook pages, we performed textual analysis to identify patterns and compare the most frequent words and co-occurrences used by each political party. *Co-occurrence* is defined as the “distance” between co-occurring terms (with a maximum of two words) because it may be informative of the actual relationship between words. For example, two terms that appear within a distance of -1 or +1 are considered more explanatory because they are often adjectives, nouns, or verbs that give further information about specific people or collectives.

To calculate word frequencies and co-occurrences, we first cleaned the data by removing stop words and changing special characters that contained diacritical marks. Subsequently, we used an unsupervised algorithm that counted the most frequent words and the top-10 terms that most co-occurred with them (Bird, Klein, & Loper, 2009).⁷

The co-occurrence results were stored in a NoSQL database and presented in four columns that showed the relative distance of each term. After a close reading of the analyzed Facebook posts, we identified nine topics that emerged from the text, and subsequently grouped the results into thematic clusters, which included words derived from the same stem, misspellings, and synonyms (see Table 2).⁸ To visualize the results, we normalized the frequencies of the terms and clusters by dividing them by the total number of posts and comments of each Facebook page per year. The clusters identified by the

⁷ Stop words were removed using the lists from the Natural Language Toolkit (<http://www.nltk.org/>).

⁸ The clusters contained words in both Spanish and Catalan.

automated method were further analyzed with a close reading of specific posts and comments on each page that contained controversial content.

Table 2. Clusters of Words Used for the Textual Analysis.

Cluster	Words included
Spain	[españa, españoles, español, española, españolas]
Immigration	[inmigración, immigració, immigracio, immigrants, immigrant, inmigracion, imigracion, extranjeros, extranjero, estranjeros, extranjero, extrangeros, extrangero, estrangeros, extranger, estrangero, extrangers, stranger, strangers, inmigrantes, inmigrante, inmigrante, inmigrantes]
Independence movement	[independentismo, independentistes, independentisme, independència, independentisme, independentistas, independencia, separatismo, separatistas]
Insults	[mierda, merda, putos, puto, asqueroso, asquerosos, hijoputa, hijoputas, apestosos, apestoso, parásitos, salvajes, parásito, parasitos, parasito, gentuza, rojos, cerdo, cerdos, chusma, marxista, marxistas]
Islam	[islam, islamistas, islámica, islamica, islamistes, islamització, islamización]
Moroccan	[magrebí, magrebies, magrebí, magrebins, magrebí, moro, moros, marroquí, marroquíes, marroquins, marroquina, marroquines]
Black people	[negros, negres, negro, subsaharianos, africano, african, negre, africanos, africà, subsahariano, subsaharià, negratas, negrata]
Romanians	[rumanos, rumano, romanesos, romanès, Rumanía]
South Americans	[panchitos, panchito, sudacas, sudaques, sudaca, sudamericano, sudamericanos]

Image and Link Analysis

For each political party and every year included in the study, we selected the 10 photos and links with the most engagement (number of likes, shares, and comments received). In total, 272 images and 306 links were manually analyzed and categorized by two independent coders, using a double classification scheme.⁹

Negative targeting. We identified the negative targeting of other groups, as derived from the content of an image or a link. Nine categories emerged from the data:

⁹ Cohen's kappa intercoder reliability is 0.661 for links and 0.957 for images.

1. *Anti-immigration*. Targeting immigrants as scapegoats of the Spanish population's problems.
2. *Antiestablishment*. Targeting the capitalist system, the ruling political class, bankers, globalization, and mainstream media.
3. *Anti-ETA*. Targeting the Basque terrorist group, Euskadi ta Askatasuna.
4. *Antileft*. Targeting people whom the extreme right considers to be "communists," such as unions and leftist parties.
5. *Antiseparatism*. Targeting the independence process that began in Catalonia in 2012.
6. *Party general information*. Information regarding the political party, such as electoral campaigns, political rallies, and corporate images.
7. *Religious symbolism/Catholic values*. Photos and links containing religious symbolism, as well as images defending Christian messages and values, such as the right to life.
8. *Spanish nationalism*. Images of the Spanish flag, demonstrations supportive of the unity of Spain, and acts of Spanish nationalism.
9. *Fascist symbolism/Francoist nostalgia*. Images and links containing fascist symbols, as well as symbols, personalities, and demonstrations that support Franco's dictatorship.

Source of the URLs. The URLs were classified according to the following categories: (1) mainstream media, (2) alternative media, (3) party website, (4) websites/blogs, and (5) social media.¹⁰ Following Klein (2012), we established this category to analyze the extent to which political extremists leverage their discourse on social media by redirecting the user to other sources of information beyond the platform.

In cases in which the same image fit in two or more categories, we chose the most relevant (i.e., immigration vs. electoral campaign). When in doubt, the decision was supported by reading the text that accompanied the photo. In addition, repeated images were ignored and broken links were grouped in a category by the same name.

Network Analysis

We performed network analysis to study the relationships between the political parties and the pages they liked. We generated a directed graph that displayed the relationships between the pages liked by the extreme-right political parties and compared them with those liked by the PP. We considered the act of liking as having cultural significance given that liking a page of a political party, extremist group, or organization is indicative of each party's politics of association and cultural preferences (Rieder, 2013). Such graphs reveal shared tastes or interests if two or more political parties like the same pages, as well as connections between the political parties if they like each other's pages.

¹⁰ These categories were not exclusive, in the sense that Category 3 (party website) could fall into Category 4 (websites/blogs). Nevertheless, we decided to establish an independent category for information coming from the parties' websites to identify their efforts to promote their political views. In Category 5 (social media), we included information coming from Facebook, Twitter, and YouTube.

Results

The rise in the popularity of social networks is evident in our data given that nearly all of the Facebook pages of the extreme-right political parties and the PP increased their number of posts, users,¹¹ and activity over the studied period (see Figure 1). In terms of content type distribution, we found that the extreme-right parties mostly shared links, followed by photos, rather than text. The PP, by contrast, mainly posted textual status updates (see Figure 2).

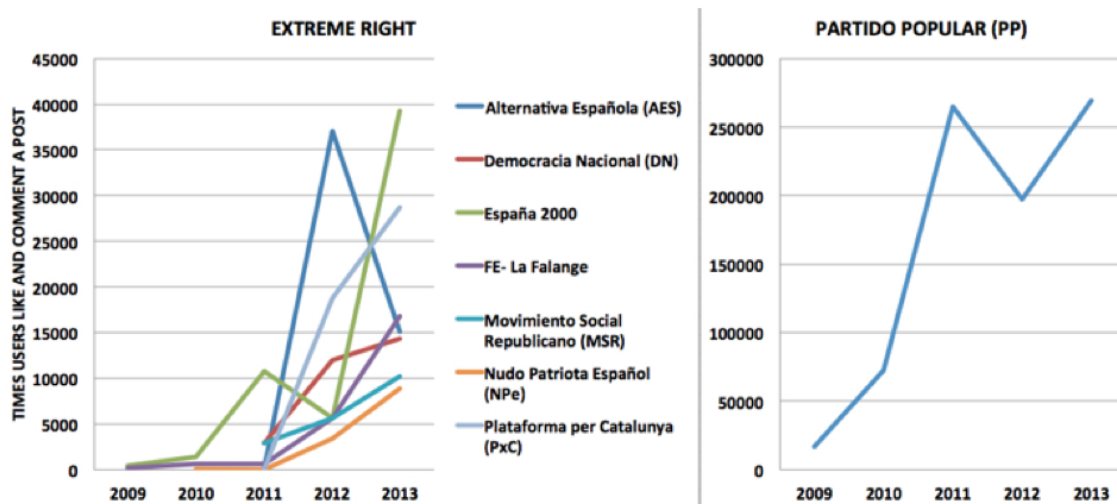


Figure 1. The number of times users "liked" and commented on posts on the Facebook pages of each of the studied political parties, per year. The scales are not normalized because setting the same scale for both graphs resulted in imperceptible values of the extreme right. The minimum value in the PP graph is the maximum value in the extreme-right graph.

¹¹The number of users on the AES page declined abruptly in 2013. This was due to a bug in the data. AES published a post in 2012 about a news event that brought many users to its page that were counted by Netvizz as AES's users.

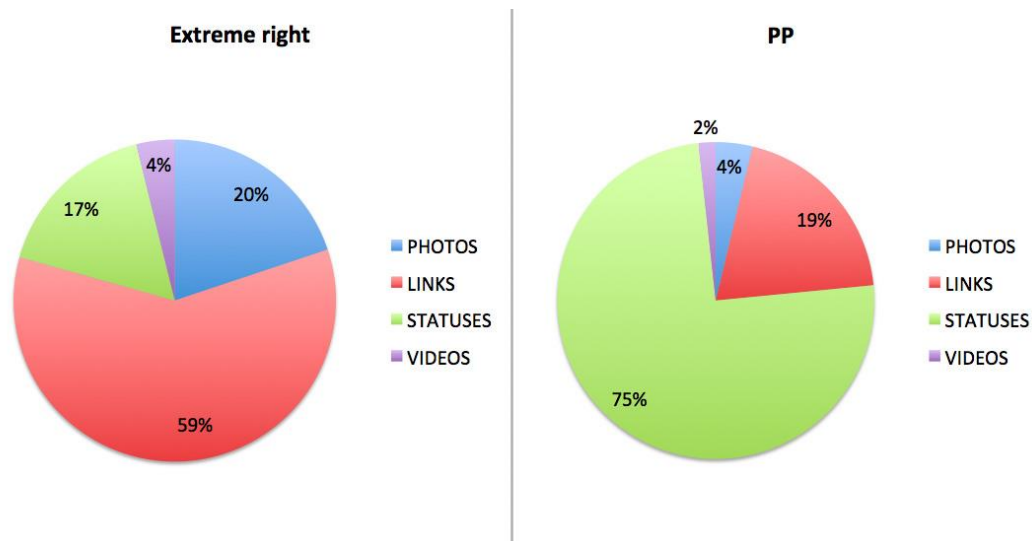


Figure 2. Content type distribution on the Facebook pages of the extreme right and the Partido Popular (PP).

Textual Speech: Hatred Against Immigrants and the Use of Slur

Figure 3 compares the frequency of the clusters of terms among all the extreme-right political parties with that of the PP. Unlike the PP, the extreme-right political parties increased their talk concerning the immigration issue over time, and their use of insults in the platform also rose.

An overview of the most frequent words revealed that the word *inmigrantes* (immigrants) was one of the top-four words on the PxC page throughout the studied years and the most frequent term on the España 2000 page in 2009. On both pages of these xenophobic parties, the words forming their anti-immigration slogans were the most frequent terms.¹²

¹² The slogans are “Primer els de casa” (“Natives come first”) for PxC and “Ni uno más. Los españoles primero” (“No one else. Spaniards come first”) for España 2000, both calling for the privileges of natives over migrants.

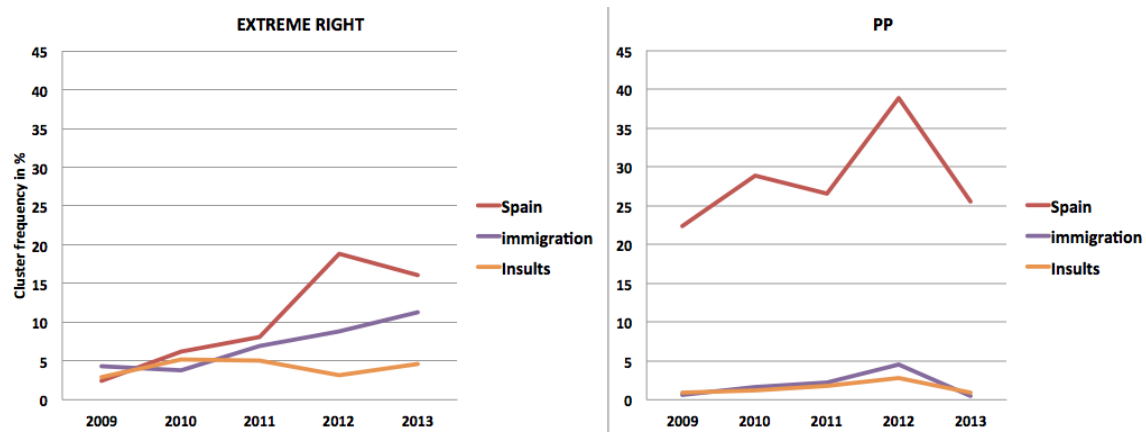


Figure 3. A comparison of the frequency of clusters of terms between the extreme-right political parties and the Partido Popular (PP).

Among the Facebook pages of the extreme right, PxC, España 2000, and DN substantially integrated the subject of immigration into their discourses.¹³ Whereas PxC constructed a clear discourse against immigration, España 2000's content was based on reactions to current events (i.e., the building of a mosque). The same was true of the DN page, although to a lesser extent. DN posted links that directed to its website, but the links triggered user engagement that generated anti-immigration discourse in the comments on the page. For instance, in 2012, an anti-immigration link posted by DN triggered the following user comment: "Spaniard unemployed, immigrant extradited. I think it's fair." By contrast, the median of mentions of the immigration cluster on the PP page was 1%; however, messages against immigrants were also found in users' comments¹⁴ (see Figure 4).

¹³The median of mentions for the cluster was 30% for the PxC party, 8% for España 2000, and 6% for DN.

¹⁴On February 28, 2012, an anonymous user posted a message on the PP's Facebook page that read, "One of the first things we should fix is the immigration issue, there are too many foreigners and they have more social rights than we [the Spaniards] do."

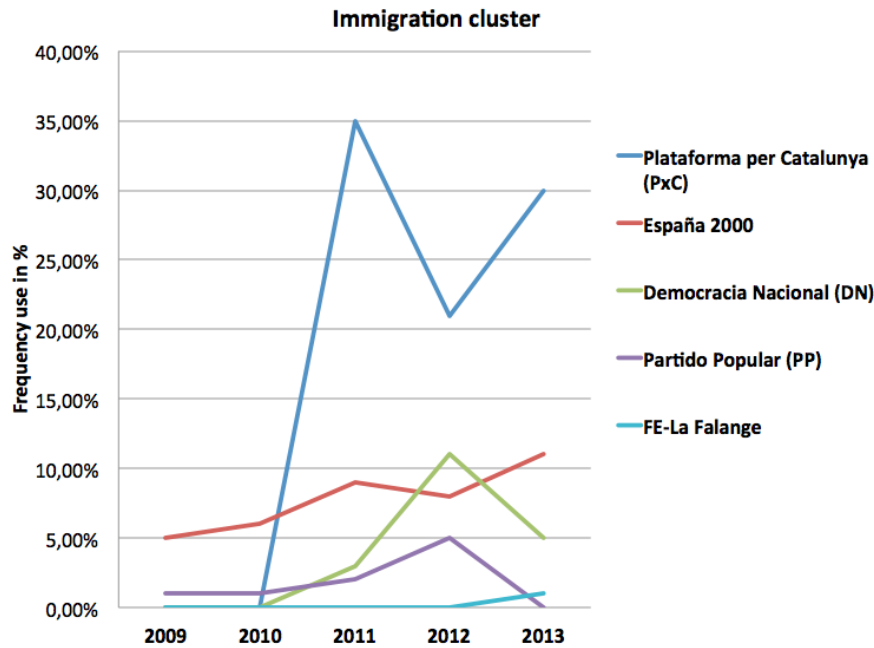


Figure 4. Use of the "immigration" cluster over time by PxC, España 2000, DN, PP, and FE-La Falange.

Moreover, on the pages of PxC, España 2000, and DN, immigration was regularly associated with crime, as well as with the opinion that there are too many foreigners in Spain. By contrast, the PP did not directly criminalize this group. The stigmatization of Muslims was particularly relevant on the PxC page, where Islam was directly targeted as a threat to Western culture and immigration was frequently related to words such as *ilegales* (illegals), *islamistas* (Islamists), *amenazan* (to threaten), and *trabajan* (to work).

A closer reading of the results showed that stigmatization of immigrants was also evident in specific posts. An official post from June 1, 2012, on the PxC page read, "To arrive in Spain without papers should be a crime. We ask for the inclusion of the illegal immigration crime in the penal code." Similarly, in a post from May 23, 2013, on the España 2000 page, the party stated, "The number of immigrant prisoners per habitant quadruple the number of Spanish in jail . . . more immigrants, more delinquency."

In most cases, anger was mainly expressed against Moroccans, Islamists, and Black people because there was an increase in the mention of words included in these clusters over time.¹⁵ These clusters were associated with concepts such as trouble, danger, and crime. For example, on the PxC page,

¹⁵ For example, the clusters "Islam" and "Black people" were mentioned a median of a 4% on the PxC page in 2012. On the España 2000 page, "Moroccans" represented nearly a 3% of the talk in 2011.

peligrosos (dangerous) co-occurred 48 times with the cluster “Islam” in 2012 and 67 times in 2013. This observation was confirmed again by a close reading of specific posts. In 2011, España 2000 organized various demonstrations against the building of a mosque in Onda, which provoked multiple messages from users, some of them calling for violent action against Muslims, such as this comment posted on March 27, 2013: “You should be throwing Molotov cocktails to those mosques, or are you planning to do it when they surpass the Pyrenees? . . . We not only have to wake up to this reality, but we must also act.” Such direct threats and calls for violent action are also contradictory to Facebook’s definition of its community standards.

As shown in Figure 5, the use of slurs increased over time, especially for España 2000, NPe, and PxC. España 2000 and PxC mostly directed their strong language against immigrants, whereas NPe and PP mainly targeted Catalans and separatists. The average use of the “insults” cluster on the España 2000, DN, and PxC pages was 7.0%, 6.7%, and 4.6%, respectively. Among the most co-occurring words were the terms *moros*—an offensive word referring to Moroccans—and offensive terms against other ethnicities such as *sudacas* (South Americans).

On the DN Facebook page, the word *catalanes* (Catalans) was frequently mentioned next to an insult. On the PP’s page, the most frequent terms next to an insult were *catalanes* in 2011 and *nacionalistas* in 2012, and the term *moritos*, an offensive word used to designate Moroccans, also occurred. These slurs appeared in the comments space. Thus, although the studied political parties did not post textual status updates that contained overt hate speech, they repeatedly stigmatized immigrants, Muslims, and Catalans by association, which covertly contributed to promote hatred against these groups.

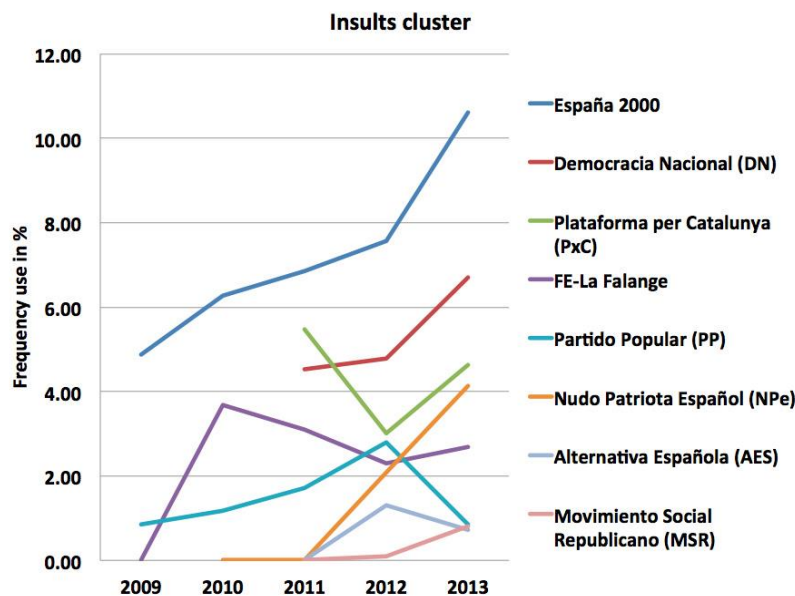


Figure 5. Use of the cluster “insults” over time by all political parties.

Discriminatory Visual Content

Among the different emerging categories, anti-immigration stood out on the pages of the extreme right, with 18.71% of the images that were most liked, shared, and commented on. By contrast, the PP's followers reacted mainly favorably (80% of the images) to messages containing general information about the party (i.e., meetings and press conferences). Apart from the anti-immigration trend, the extreme-right community engaged with messages related to Spanish nationalism, party general information, antiestablishmentarianism, and fascist symbolism.

Table 3 shows the distribution of the most popular images categorized by content for each party. The most popular images posted by PxC and España 2000 stand out as having the highest percentage of images criticizing immigrants (75% and 31%, respectively), a popularity that increased over time. Two examples of such discriminatory visual content are presented in Figure 6, in which images posted by PxC and España 2000 stigmatize Muslims by using abject humor to portray Muslims as pigs or as primitives that are in favor of stoning women.

Table 3. Percentages of the Most Popular Images Categorized by Content on the Facebook Pages of the Extreme Right and the Partido Popular.

Category	Political party							
	PxC	DN	FE-La Falange	AES	MSR	España 2000	NPe	PP
Anti-immigration	75	11	4	3	7	31	0	0
Antiestablishment	4	13	10	7	37	17	12	0
Antileft	0	7	4	7	0	0	16	16
Anti-ETA	0	3	4	10	0	3	4	2
Spanish nationalism	0	30	37	13	10	29	24	0
Party general information	21	20	10	17	23	17	8	80
Religious symbolology/Catholic values	0	0	0	33	0	0	0	0
Fascist symbolology/Francoist nostalgia	0	3	31	3	13	0	20	0

Note. PxC = Plataforma per Catalunya; DN = Democracia Nacional; AES = Alternativa Española; MSR = Movimiento Social Republicano; NPe = Nudo Patriota Español; PP = Partido Popular; ETA = Euskadi ta Askatasuna.



Figure 6. Discriminating images of Muslims posted by Plataforma per Catalunya (left) and España 2000 (right).

Links to Discriminatory Content

Among the links shared by the extreme-right political parties, 25% of the links that were most liked, shared, and commented on fell under the anti-immigration category. By contrast, the PP did not upload any link containing anti-immigration information on its Facebook page.

Regarding the sources of the most popular URLs, whereas the PP used the official party website as the primary source of information on its page (91% of the links), the extreme right relied on a variety of sources such as blogs, social media, and other news media outlets alongside the links to the official party websites (41% of the links; see Figure 7). Some of links shared on the extreme right pages redirected the users to websites of dubious legitimacy, such as links¹⁶ posted by DN in 2013 that redirected the users to websites containing anti-immigration information or links¹⁷ posted in 2009 by FE-La Falange from Falangists organizations.

¹⁶ For example, <http://elcasar.blogspot.com.es/2013/01/asi-nos-tratan-los-queridos-inmigrantes.html>
<http://valladolidisidente.blogspot.com.es/2013/08/la-semilla-de-la-ira-diaria-3-magrebies.html>,

¹⁷For example, <http://auxilio-azul.blogspot.com.au/2009/11/es-ahora-cuando-hacemos-falta.html>,
<http://www.sindicatotns.es/falange/?p=491>

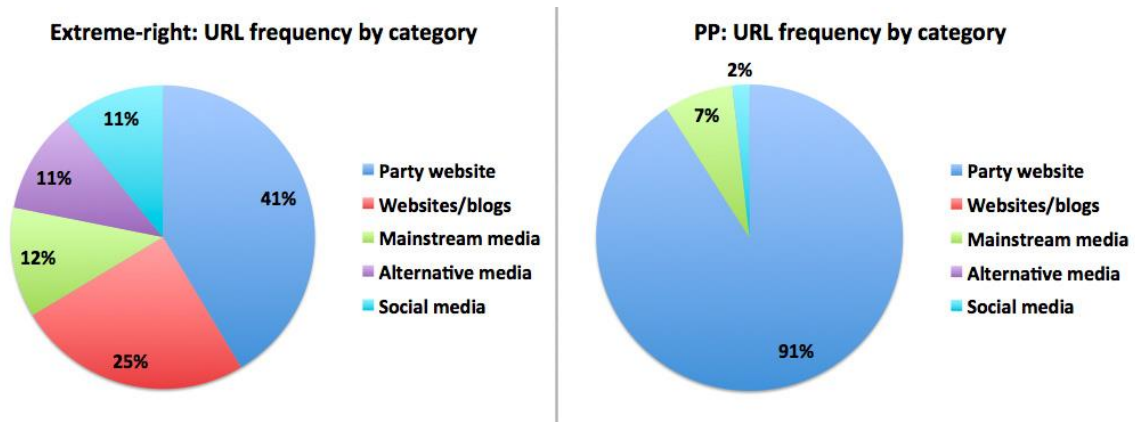


Figure 7. Sources of the most popular URLs on the Facebook pages of the Partido Popular and the extreme right.

Network Analysis: Liking as a Cultural Practice

The immediate network of likes of each political party showed that the extreme-right political parties were not interrelated, apart from a cluster formed by DN, FE-La Falange, and NPe, which replicated their part in the extreme-right coalition La España en Marcha. AES, DN, and NPe liked some Spanish nationalistic pages, MSR liked an anti-immigration page, and the PP and AES liked antiseparatism pages. In addition, MSR and España 2000 liked extremist international organizations, such as the Swedish antimulticulturalism organization, Nationaldemokraterna; the French right-wing populist political party, Front National; and an extreme-right party from Portugal, National Renovator Party.

Despite the insignificant internal liking among the extreme-right parties, a network analysis of the second layer of pages liked by the studied political provided a broader perspective on their networks of associations. The visualization in Figure 8 shows that the PP and the extreme-right parties form different clusters. Unlike the PP, which is linked to a cluster of Spanish mainstream media and to a cluster of the European People’s Party, the extreme right is not linked to the Spanish mainstream media, but to a broader network of far-right nationalistic political parties, most of them supporting anti-immigration agendas.¹⁸ Thus, the networks of association of the extreme right are more aligned with international far-right actors than with domestic counterparts.

¹⁸ These foreign parties did not “like” them back, however.

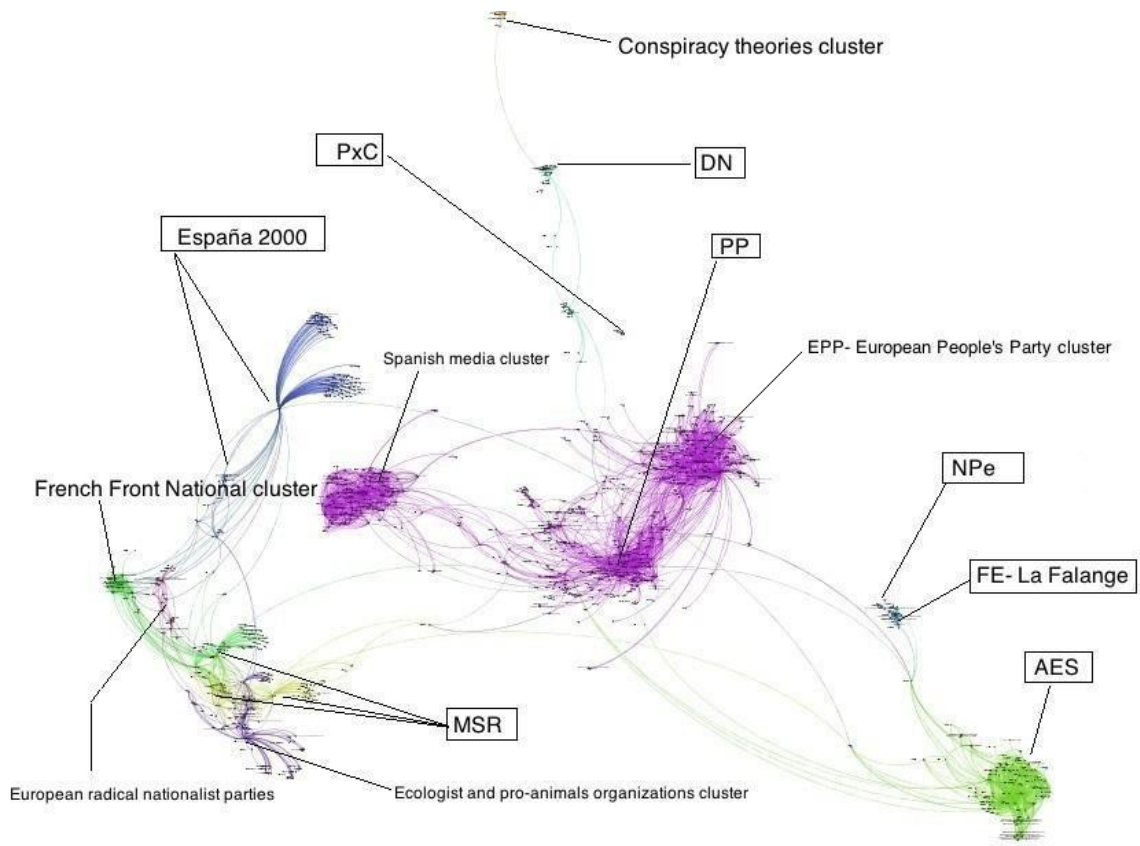


Figure 8. A network of "likes" of the extreme right and the Partido Popular. PxC = Plataforma per Catalunya; DN = Democracia Nacional; AES = Alternativa Española; MSR = Movimiento Social Republicano; NPe = Nudo Patriota Español; PP = Partido Popular.

Discussion

Although previous research has discussed the distributed spread of online hate through activities such as hyperlink networks and downloads (Klein, 2012), this research moves beyond the focus on deliberate rhetoric, motivations, and actions of the propagators of hate and their followers to study the ways that overt hate speech and covert discrimination circulate on Facebook through the platform's policy and design, the content types that it hosts, and the people who engage with them. In choosing to analyze the Facebook activities of the extreme-right political parties in Spain, and to compare them with those of the governing party, PP, this study also shifts the focus from analyzing racist practices of hate groups (Klein, 2012; Nakamura, 2014; Oboler, 2008; Rajagopal, 2002; Warner & Hirschberg, 2012) to examine the introduction and circulation of hate and discrimination on Facebook within the legitimate boundaries of the Spanish political discourse. Our analysis demonstrates that both forms of overt hate speech and covert discrimination are found on the platform. Overt hate speech was found in direct threats against immigrants and in derogative associations and insults to foreign people, which were identified after a close reading of the political parties' posts and their user comments. Covert discrimination was found in repeated co-occurring textual associations of immigrants with danger and crime, as well as in popular images that denigrate Muslims and in popular links to anti-immigration websites. Moreover, the network of likes between the studied parties revealed their association with international xenophobic and extremist parties. It should be noted that instances of overt hate speech and covert discrimination that were identified in the official Facebook pages of the extreme-right political parties were not found in the official page of the PP, with the exception of textual hate speech against Catalans that was posted by followers of the PP in the page's comment space.

Whereas PxC, España 2000, and DN posted explicit multimodal content that discriminated immigrants, the rest of the extreme-right political parties were primarily implicating discrimination (as detected by the associated words, links, imagery, and repeated messages). Although we did not investigate whether the mitigation of overt hate speech was an outcome of a deliberate strategy of the studied political parties to adhere to Facebook's community standards, and although we cannot penetrate the black box of the platform's internal corporate logic on determining whether or not specific content violates its community standards, our findings clearly show that explicit derogative hate speech toward immigrants and insults against Catalans are particularly prevalent in the user comment space of all studied political parties. Given that the parties were liable to Facebook for the content they post, but not for the users' comments on their pages, we therefore witnessed a chain of liability transference with regards to hate speech: First, Facebook leverages the liability of hosting hate speech from the platform to its users through its community standards (Gillespie, 2010); subsequently, owners of Facebook pages (in this case, the extreme-right political parties in Spain) transfer the liability of overt hate speech to their followers through the platform's technological affordances.

Put differently, we argue that content analysis and the study of users' motivations are suited for studying overt hate speech, but do not suffice to explain and unravel practices of hate and circulation of covert discrimination on Facebook. Had we performed only textual analysis, we would have found that the extreme-right political parties do not use hate speech on Facebook, and the covert discriminatory practices identified in this study could not have been detected. Instead, by assigning equal agency to

human and nonhuman actors, we have shown that the simultaneous meanings of communicative acts on Facebook, which were previously discussed in cultural and economic contexts (Gerlitz & Helmond, 2013; Gillespie, 2010; Langlois & Elmer, 2013) also extend to hate and discrimination. Our findings reveal the platform's dual agency in shaping a Web space that officially prohibits explicit manifestations of hate, on one hand, and on the other, provides a networked infrastructure for the circulation and accumulation of subtle associations of hate and discrimination. For example, Facebook's policy on hate speech relates to content, but not to practices of its circulation through the like, comment, and share buttons, the technological affordances through which we found that most covert discriminatory practices circulate. Moreover, the platform's default design allows users to upload any content, and although the company employs moderators who remove pornographic images (Chen, 2014), most of the content that violates Facebook's community standards is not moderated unless it is actively flagged or reported by other users (Crawford & Gillespie, 2014). In addition, we have shown that the content uploaded by the Spanish extreme-right political parties circulated in isolated clusters of broader far-right networks, which confirms both a "filter bubble" effect (Pariser, 2011) and the marginality of the studied political parties (Caiani & Parenti, 2011; Jackson & Lilleker, 2009), but at the same time decreases the likelihood that such content will be reported by users who find it offensive. As a result, and as evidenced by our longitudinal analysis, discriminatory associations accumulate over time, and may have a long-term effect on the introduction of extremist ideas as a legitimate form of online culture (Klein, 2012).

Nevertheless, there are several limitations to the analysis. One of the challenges relates to the scope of the extracted data, which documented the Facebook activity of eight political parties over a time range of three to five years. The comparative and longitudinal approach excluded in-depth analyses of the nuanced practices of each party's page, which can be seen as subjects for future research. In addition, our findings are limited to Facebook, and cannot be readily applied to other social networking platforms, which employ different policies, attract different audiences, and design interactions through different features. Future cross-platform research could expand the scope of analysis to other social networking platforms and to other case studies. In addition, future research on Facebook could further elaborate on the ties between the platform's policy, its technological affordances, user behavior, and the circulation of overt hate speech and covert discrimination by empirically studying the platform's reactions to flagged content and the contribution of the EdgeRank algorithm to the circulation of discriminatory content on users' news feeds, and by investigating the motivations of both extremists and their followers in posting overt hate speech or covertly disguising it with regards to the platform's community standards.

Conclusion

This study considered the ways that Facebook's corporate logic, its technological affordances, and user practices come together to enable the circulation of hate despite the platform's official policy. Throughout the study, we analytically distinguished between instances of overt hate speech, which could be regarded as violating the platform's community standards (such as comments calling to violent action against immigrants), and between covert practices, which were not addressed by the platform's community standards but nonetheless discriminated through the interaction between users and the platform's technological affordances (such as liking a link that redirects to an external anti-immigration website). Although our distinction does not purport to be commensurable with Facebook's internal rules

for considering what does or does not violate its community standards, or with regulatory and legislative criteria that determine whether specific content or activities are legitimate instances of free speech, we aimed to show that the current disparities between the regulatory and legal frameworks that attempt to address performances of hate online, and between the platform's corporate policy and technological affordances, create a void in which discriminatory practices are prevalent, untargeted, and undetected. In the case of the Spanish extreme-right political parties, we have shown that Facebook hosts an increasing volume of covert discriminatory practices that not only circulate data and content, but also trigger overt hate speech by the parties' followers. However, the circulation of covert discrimination is not limited to the Spanish case and may have broader implications on the role Facebook plays as a cultural intermediary (Gillespie, 2010). Because Facebook is primarily designed to favor and share content that increases the company's revenue from data activity—and until new regulatory means and emerging user practices fundamentally change the way hate speech and discrimination are defined and targeted on Facebook—hate and discrimination will continue to circulate through the interaction between users and the platform's technological affordances.

References

- Anti-Defamation League. (2015). *Report of the Anti-Defamation League on confronting cyberhate*. New York, NY: Author. Retrieved from <http://www.antidef.org.au/news/report-of-the-anti-defamation-league-on-confronting-cyberhate/>
- Arellano Toledo, W. (2012). *La sociedad de la información en Iberoamérica. Estudio multidisciplinar* [The information society in Latin America. A multidisciplinary study]. Mexico City, Mexico: Fondo de Información y Documentación para la Industria, INFOTEC.
- Banks, J. (2010). Regulating hate speech online. *International Review of Law, Computers & Technology*, 24(3), 233–239.
- Benedí, S. (2013). *ENAR shadow report*. Brussels, Belgium: European Network Against Racism. Retrieved from <http://cms.horus.be/files/99935/MediaArchive/publications/shadow%20report%202011-12/Spain.pdf>
- Berlet, C. (2001, April). *When hate went online*. Paper presented at the Northeast Sociological Association Spring Conference. Fairfield, CT: Sacred Heart University.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. Sebastopol, CA: O'Reilly Media.
- Brown, I. (2010). *Internet self-regulation and fundamental rights* (SSRN Scholarly Paper No. 1539942). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1539942

- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society, 14*(7), 1164–1180.
- Caiani, M., & Parenti, L. (2011). The Spanish extreme right and the Internet. *Análise Social, 46*(201), 719–740.
- Casals, X. (2000). La ultraderecha española: Una presencia ausente (1975–1999) [The extreme-right in Spain: An absent presence]. *Historia y Política: Ideas, Procesos y Movimientos Sociales, (3)*, 147–174.
- Chan, J., Ghose, A., & Seamans, R. (2014). *The Internet and hate crime: Offline spillovers from online access* (SSRN Scholarly Paper No. 2335637). Rochester, NY: Social Science Research Network. Retrieved from <http://papers.ssrn.com/abstract=2335637>
- Chen, A. (2014, October 23). The laborers who keep dick pics and beheadings out of your Facebook feed. *Wired*. Retrieved from <http://www.wired.com/2014/10/content-moderation/>
- Citron, D. K. (2014). *Hate crimes in cyberspace*. Cambridge, MA: Harvard University Press.
- Citron, D. K., & Norton, H. L. (2011). *Intermediaries and hate speech: Fostering digital citizenship for our information age* (SSRN Scholarly Paper No. 1764004). Rochester, NY: Social Science Research Network. Retrieved from <http://papers.ssrn.com/abstract=1764004>
- Council of Europe. (2001). Convention on cybercrime. Retrieved from <http://conventions.coe.int/Treaty/en/Treaties/Html/185.htm>
- Council of Europe. (2003). Additional protocol to the Convention on Cybercrime, concerning the criminalization of acts of a racist and xenophobic nature committed through computer systems. Retrieved from <http://conventions.coe.int/Treaty/en/Treaties/Html/189.htm>
- Crawford, K., & Gillespie, T. (2014). What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society*. Advanced online publication. doi:1461444814543163
- Datta, A., Tschantz, M. C., & Datta, A. (2015). Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination. *Proceedings on Privacy Enhancing Technologies, 2015*(1), 92–112. Retrieved from <http://doi.org/10.1515/popets-2015-0007>
- Ellison, N. B., Hancock, J. T., & Toma, C. L. (2012). Profile as promise: A framework for conceptualizing veracity in online dating self-presentations. *New Media & Society, 14*(1), 45–62.
- European Commission Against Racism and Intolerance. (2011). ECRI report on Spain. Retrieved from <http://www.coe.int/t/dghl/monitoring/ecri/Country-by-country/Spain/ESP-CBC-IV-2011-004-ENG.pdf>

- Facebook. (n.d.a). Facebook community standards. Retrieved from <https://www.facebook.com/communitystandards>
- Facebook. (n.d.b). Statement of rights and responsibilities. Retrieved from <https://www.facebook.com/legal/terms>
- Ferrándiz, J. P. (2014, February 5). ¿Hay espacio más allá? [Is there any space great beyond?]. *El País*. Retrieved from http://sociedad.elpais.com/sociedad/2014/02/05/actualidad/1391634791_203513.html
- Foxman, A. H., & Wolf, C. (2013). *Viral hate: Containing its spread on the Internet*. New York, NY: St. Martin's Press.
- Garea, F., & Medina, M. A. (2014, May 21). Spanish government asks state attorney to crack down on Twitter hate speech. *El País*. Retrieved from http://elpais.com/elpais/2014/05/21/inenglish/1400663993_676484.html
- Gascón, A. (2012). Evolución jurisprudencial de la protección ante el discurso del odio en España en la última década [Case law evolution of the protection of hate speech in Spain in the last decade]. *Cuadernos Electrónicos de Filosofía del Derecho*, 26, 310–340. Retrieved from <https://ojs.uv.es/index.php/CEFD/article/view/2002/3021>
- Gerlitz, C., & Helmond, A. (2013). The like economy: Social buttons and the data-intensive Web. *New Media & Society*, 15(8), 1348–1365.
- Gerstenfeld, P. B., Grant, D. R., & Chiang, C.P. (2003). Hate online: A content analysis of extremist Internet sites. *Analyses of Social Issues and Public Policy*, 3(1), 29–44.
- Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 67–82). Hillsdale, NJ: Erlbaum.
- Gillespie, T. (2010). The politics of "platforms." *New Media & Society*, 12(3), 347–364.
- Gillespie, T. (2012). The dirty job of keeping Facebook clean. *Culture Digitally*. Retrieved from <http://culturedigitally.org/2012/02/the-dirty-job-of-keeping-facebook-clean/>
- Grant, B. (2015, July 24). Facebook refuses to shut down "racist" "Humans of Sunnybank" group which cruelly mimics Asian accents and labels the Brisbane community "dog meat eaters." *Daily Mail Australia*. Retrieved from <http://www.dailymail.co.uk/news/article-3172851/Facebook-refuses-shut-racist-Humans-Sunnybank-group-crudely-mimics-Asian-accents-labels-Brisbane-community-dog-meat-eaters.html>

- Gündüz, Z. Y. (2010). The European Union at 50—Xenophobia, Islamophobia and the rise of the radical right. *Journal of Muslim Minority Affairs, 30*(1), 35–47.
- Heins, M. (2014). The brave new world of social media censorship. *Harvard Law Review, 127*(325), 325–330. Retrieved from <http://harvardlawreview.org/2014/06/the-brave-new-world-of-social-media-censorship/>
- Jackson, N. A., & Lilleker, D. G. (2009). Building an architecture of participation? Political parties and Web 2.0 in Britain. *Journal of Information Technology & Politics, 6*(3–4), 232–250. doi:10.1080/19331680903028438
- Jamieson, A. (2009, April 23). Facebook “must do more to tackle racism” say campaign groups. *The Telegraph*. Retrieved from <http://www.telegraph.co.uk/technology/facebook/5205426/Facebook-must-do-more-to-tackle-racism-say-campaign-groups.html>
- Jensen, T. B., & Dyrby, S. (2013). *Exploring affordances of Facebook as a social media platform in political campaigning* (ECIS 2013 Completed Research Paper 40). Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1263&context=ecis2013_cr
- Klein, A. (2012). Slipping racism into the mainstream: A theory of information laundering. *Communication Theory, 22*(4), 427–448.
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London, UK: Edward Arnold.
- Kuchera, B. (2014, July 30). Twitter can fix its harassment problem, but why mess with success? *Polygon*. Retrieved from <http://www.polygon.com/2014/7/30/5952135/twitter-harassment-problems>
- Kurzman, C., & Schanzer, D. (2015, June 16). The growing right-wing terror threat. *The New York Times*. Retrieved from <http://www.nytimes.com/2015/06/16/opinion/the-other-terror-threat.html>
- Langlois, G., & Elmer, G. (2013). The research politics of social media platforms. *Culture Machine, 14*. Retrieved from <http://www.culturemachine.net/index.php/cm/article/view/505>
- Latour, B. (2005). *Reassembling the social: An introduction to actor–network theory*. Oxford, UK: Oxford University Press.
- Levin, B. (2002). Cyberhate: A legal and historical analysis of extremists’ use of computer networks in America. *American Behavioral Scientist, 45*(6), 958–988. doi:10.1177/0002764202045006004
- Massanari, A. (2015). #Gamergate and The Fappening: How Reddit’s algorithm, governance, and culture support toxic technocultures. *New Media & Society*. Advance online publication. doi:1461444815608807

- Moretón, M. (2012). El "ciberodio," la nueva cara del mensaje de odio: Entre la cibercriminalidad y la libertad de expresión [Cyberhate, the new face of hate speech: Between cybercrime and freedom of expression]. *Revista Jurídica de Castilla y León*, 27, 1–18.
- Nakamura, L. (2014). "I WILL DO EVERYthing that am asked": Scambaiting, digital show-space, and the racial violence of social media. *Journal of Visual Culture*, 13(3), 257–274.
- NewsOne. (2013). Facebook refuses to remove racist "Black Baby as Monkey" post. *News One for Black America*. Retrieved from <http://newsone.com/2187515/facebook-black-baby-as-monkey/>
- Oboler, A. (2008). The rise and fall of a Facebook hate group. *First Monday*, 13(11). Retrieved from <http://firstmonday.org/ojs/index.php/fm/article/view/2254>
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. New York, NY: Penguin.
- Rajagopal, I. (2002). Digital representation: Racism on the World Wide Web. *First Monday*, 7(10). Retrieved from <http://uncommonculture.org/ojs/index.php/fm/article/view/995>
- Rieder, B. (2013). Studying Facebook via data extraction: The Netvizz application. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 346–355). New York, NY: Association for Computing Machinery. Retrieved from <http://dl.acm.org/citation.cfm?id=2464475>
- Rogers, R. (2013). *Digital methods*. Cambridge, MA: MIT Press.
- Rosen, J. (2011). Deciders: The future of privacy and free speech in the age of Facebook and Google. *Fordham Law Review*, 80(1525). Retrieved from <http://ir.lawnet.fordham.edu/flr/vol80/iss4/1/>
- Sánchez Duarte, J., & Rodríguez Esperanza, S. (2013). La extrema-derecha en Facebook. España 2000 y Democracia Nacional durante la campaña electoral de 2011 [The extreme right in Facebook. España 2000 and Democracia Nacional during the election campaigns of 2011]. *Revista Mediterránea de Comunicación*, 4(1). Retrieved from <http://dx.doi.org/10.14198/MEDCOM2013.4.1.10>
- Schafer, J. A. (2002). Spinning the web of hate: Web-based hate propagation by extremist organizations. *Journal of Criminal Justice and Popular Culture*, 9(2), 69–88.
- Sewell, A. (2013, August 21). Spanish PP politicians in trouble over Facebook fascist salutes. *Digital Journal*. Retrieved from <http://www.digitaljournal.com/article/356845>
- Shepherd, T., Harvey, A., Jordan, T., Srauy, S., & Miltner, K. (2015). Histories of hating. *Social Media + Society*, 1(2). doi:2056305115603997

- Simon Wiesenthal Center. (2012, March 27). Social media must do more to thwart subculture of hate fueling lone wolf terrorism – Simon Wiesenthal Center debuts 2012 Digital Hate Report. *Simon Wiesenthal Center*. Retrieved from <http://www.wiesenthal.com/site/apps/nlnet/content2.aspx?c=lsKWLbPJLnF&b=4441467&ct=11675937#.VrA6-aL89GG>
- Warner, W., & Hirschberg, J. (2012). Detecting hate speech on the World Wide Web. In *Proceedings of the Second Workshop on Language in Social Media* (pp. 19–26). Montreal, Quebec: Association for Computational Linguistics. Retrieved from <http://dl.acm.org/citation.cfm?id=2390377>
- Whine, M. (1999). Cyberspace—A new medium for communication, command, and control by extremists. *Studies in Conflict & Terrorism*, 22(3), 231–245. doi:10.1080/105761099265748