

**BREAKING THE CHAINS OF HATE, FORGING NETWORKS OF SUPPORT:
YOUNG PEOPLE ONLINE AND ONLINE HATE SPEECH**

SUMMARY

**NO MORE
~~HATERS~~
Break the hate chain!**

Centro
Reina Sofía
sobre adolescencia
y juventud

fad

M MALDITA.ES

With the support of:

Google.org

BREAKING THE CHAINS OF HATE, FORGING NETWORKS OF SUPPORT: YOUNG PEOPLE ONLINE AND ONLINE HATE SPEECH

SUMMARY

Cómo citar el resumen:

Kuric, S. (2021). Young people online: breaking the chains of hate speech. Summary. Madrid: Centro Reina Sofía sobre Adolescencia y Juventud, Fad.
DOI: 10.5281/zenodo.4596326

Cómo citar el estudio completo:

Megías, I. (coord.); Amezaga, A.; García, M.C.; Kuric, S.; Morado, R. y Orgaz, C. (2020). Romper cadenas de odio, tejer redes de apoyo: los y las jóvenes ante los discursos de odio en la red. Madrid: Centro Reina Sofía sobre Adolescencia y Juventud, Fad.
DOI: 10.5281/zenodo.4288486

FAD, 2021

Edita:

Centro Reina Sofía sobre Adolescencia y Juventud
Fundación de Ayuda contra la Drogadicción (FAD)
Avda. de Burgos, 1 y 3
28036 Madrid
Teléfono: 91 383 83 48
fad@fad.es

Maquetación:

Jorge Albert Lenze

Traducción:

Paul Cassidy

DOI: 10.5281/zenodo.4596326

Licencia:

Creative Commons,
Attribution 4.0 International (CC BY 4.0)



With the support of:

Google.org

~~INDEX~~

1.- INTRODUCTION	04
2.- METHODOLOGY	08
3.- UNDERSTANDING HATE SPEECH	11
4.- CHARACTERISTICS AND MECHANISMS OF ONLINE COMMUNICATION THAT PROMOTE HATE SPEECH	15
5.- TYPES OF ONLINE HATE SPEECH	20
5.1.- Hatred of vulnerable collectives	21
5.2.- Hate in public-private spheres	22
5.3.- The normalisation of low intensity hate speech	22
5.4.- Hate speech from the perspective of victims	23
6.- RESPONSIBILITY FOR HATE SPEECH	25
7.- CONCLUSIONS AND PROPOSALS	28
8.- BIBLIOGRAPHY	32

**NO MORE
~~HATERS~~**

INTRODUCTION

~~INTRODUCTION~~

This text presents the principal results from the research project “Breaking the chains of hate, forging networks of support: young people and online hate speech”, developed by the *Centro Reina Sofía sobre Adolescencia y Juventud* (Reina Sofia Centre for Adolescence and Youth), part of FAD. The primary aim of the research was to analyse the perceptions, attitudes and experiences of young Spanish people in relation to their experiences of online hate speech.

Research on this topic and its impact on young people are essential. In 2015, the European Commission against Racism and Intolerance (ECRI) warned that hate speech was becoming increasingly habitual in online spaces¹. However, the commission also pointed out that despite this proliferation, very little is known about the real prevalence, form or characteristics of online hate speech. In addition, the lack of data makes any comparative analysis practically impossible. In setting out to research online hate speech and develop a typology of its forms and characteristics, one of the foremost issues is its complexity and lack of common international definitions. The ECRI defines hate speech as follows:

“Promotion or incitement, in any form, of the denigration, hatred or vilification of a person or group of persons, as well as any harassment, insult, negative stereotyping, stigmatization or threat in respect of such a person or group of persons and the justification of all the preceding types of expression, on the ground of “race”, colour, descent, national or ethnic origin, age, disability, language, religion or belief, sex, gender, gender identity, sexual orientation and other personal characteristics or status.”

¹ European Commission against Racism and Intolerance (ECRI) (2015) General Policy Recommendation nº15 - On combating hate speech. Recuperado el 28/01/2021 de: <https://rm.coe.int/ecri-general-policy-recommendation-n-15-on-combating-hate-speech-adopt/16808b7904>

This very broad and complex definition requires that the concept be narrowed down and made more specific to the object of this study: online hate speech as it relates to young people.

To this effect, three essential aspects of hate speech can be identified (*Bazzaco et al., 2017, p. 8; Bustos Martínez et al., 2019, p. 26; Fundación Secretariado Gitano, 2017*). Firstly, it must contain inherently hostile expressions of hate based on the singling out or stigmatization of real or imaginary traits of specific social groups. Secondly, it can be articulated as direct or indirect attacks on members of these groups or the symbols that represent them, as well as inciting such attacks. Thirdly, the affected social groups should have a history of suffering from oppression or discrimination and be in a situation of recognised vulnerability. Nevertheless, it is also important to take into account that a person does not have to be part of a stigmatised group to be affected by hate speech.

The objectives of hate speech can be to silence, humiliate, intimidate, discriminate, exclude, degrade, belittle, dehumanise, persecute, threaten, harass, injure, discredit, vilify, incite violence, disseminate negative stereotypes, stigmatise, encourage insensitivity or brutality towards specific vulnerable social groups (*Richardson-Self, 2018*). For victims, the consequences of such acts and the harm that they cause may be observed at a personal and/or social level. On the one hand, they may experience psychological suffering, while at a social level negative effects may be evident in the way that victims freedom of movement is restricted through fear of exposing themselves to hate (*Jubany & Roiha, 2018, p. 36*).

From a global standpoint it is more difficult to quantify the magnitude of online hate speech, given that no standardised mechanisms exist with which to detect or measure it. Nevertheless, it is clear that the Internet and social media are now the primary platforms of dissemination. According to a recent report of the Spanish Ministry of the Interior (*Fernández-Oruña et al., 2019, p. 28*), 54.9% of hate speech acts detected by Spanish law enforcement agencies were

produced on the Internet, followed by 17.2% on social media. It is also notable that the profiles of users responsible for hate speech are highly diverse and heterogonous, adding to the visibility of hate speech over the last few decades (*Bazzaco et al., 2017*). Furthermore, the attempts to regulate online speech hate, such as the EU Code of Conduct on countering illegal hate speech online and The Network Enforcement Act (Germany) have achieved ambiguous results (*Bayer & Bárd, 2020*).

**NO MORE
~~HATERS~~**

METHODOLOGY

~~METHODOLOGY~~

Using a qualitative methodology, the research undertakes an analysis of the current situation in Spain. Apart from an extensive review of secondary data and literature, qualitative material was collected through four techniques, all employed in online contexts:

- Two in-depth interviews with experts with extensive experience in the analysis of hate speech from the perspective of theory and research (identified in the text as Expert 1 and 2).

- Two paired interviews with informants of special interest:
 - Specialists in community action and social work, involved in training, prevention and mediation interventions related to hate speech.

 - Representatives of specific social collectives that habitually experience hate speech (vulnerable collectives).

- Three asynchronous online communities established for the research, which consisted of groups of 10 to 14 people organised by gender and age group (18-20, 22-24, 26-29). Each group lasted one week.

- A process of ‘active listening’ through six one-to-one interviews with frequent Twitter users that focused on their experiences on the platform. The following criteria were used to construct the sample: level of vulnerability (greater in interviews 1, 3, 5, 6), number

of followers (greater in interviews 3, 4, 5), and the intensity of experiences of hate speech (Interviews 1, 3, 5, 6) or cyberbullying (interviews 2 and 4). Through these interviews, participants were presented with a series of examples of hate speech and subsequently a typology of hate speech was developed based on the characteristics of the sample messages, including the apparent motivation for the communication and other specificities. An intensive context analysis of the selected messages was also conducted through data extraction techniques (scraping) that permitted the identification of the context of enunciations (statistical analysis, network analysis) and utterances (discourse analysis).

**NO MORE
~~HATERS~~**

**UNDERSTANDING
HATE SPEECH**

~~UNDERSTANDING HATE SPEECH~~

Amongst the participants there was general agreement about what the term ‘hate speech’ constitutes. The young people in the study viewed hate speech as an act of discrimination and abuse that attacks difference and represents a violation of human rights and dignity. Such acts are understood to be founded in prejudice, lack of knowledge of social reality, and to be self-reproducing.

For me hate speech is something that incites discrimination, abuse [...] of specific groups of people, because of their race, their religion, their ideas, something that makes them different. People that use these discourses see others, and ideas they don't agree with, as unnatural or think that they shouldn't exist.

Man, online community, 18-20 year olds

However, for the study participants there are also a series of factors that generate confusion around the conceptualisation of hate speech, in particular its polysemy, malleability and ambiguity about its scope and limits. This plurality and lack of an agreed definition means that attempts to regulate hate speech have become embroiled in various debates and controversies, most notably the dilemma around freedom of expression and its boundaries.

I think that there is a lot of confusion [...] I mean, [between] what is protected by freedom of expression and what should be punishable. I think that the boundaries are so confusing because really it isn't regulated nor... [it hasn't been] established what hate speech is, what freedom of expression is, what... yes, there's a lot of confusion and overlap from between one and the other.

Paired interview 2, vulnerable collectives

Another aspect that generates confusion amongst the study participants was the demarcation between online and offline behaviour. Far from being two distinct spheres, there are strong continuities between the two worlds. Both are spaces where culture is produced as well as being non-neutral cultural artefacts (*Hine, 2000*) that feed into each other while maintaining their own specificities. This complementarity and integration is present in general beliefs about the online and offline world as most people consider that they are 'themselves' in both spaces. However, of note amongst the discourses of the young people in the study is that they tend to attribute a higher degree of sincerity to online communication, as there are few established norms that govern interactions like in face-to-face communication. Furthermore, these same norms incorporate the idea that playing, to a certain degree, with deception or deceit is a characteristic of online communication.

Believing that what is said online (and any associated consequences) stays online is identified as a primary idea that encourages hate speech. Given the assimilation and complementarity of both worlds, this represents a paradox that forms the basis of the majority of youth discourses, which in some senses can be understood as a means of denying responsibility. Online attacks and insults are discontinued in the offline world and their significance is downplayed because only "face to face" communication is thought to be important (*Cordo & Megías, 2006*). However, online victims may suffer the psychological consequences in their day-to-day lives, as hate is transmitted and acts simultaneously in both spheres.

Finally, the analysis found that the huge amount and variety of online information also conditions perceptions and understandings of hate speech. From a very early age, adolescents and young people are faced with highly ideological and political discourses at a time when they are not necessarily able to assimilate them adequately and to distinguish between acceptable and intolerable ideas. Hence, as online users gain experience they go through a process of socialisation in which they learn the norms, codes, languages, genres and concepts that govern online spaces. As they discover what is permissible to say and what will be censored in different contexts they adapt to the medium and modify behaviour. In this sense, hate speech alters the way that young people engage with other users online, as well as influencing their choices about what to post or communicate.

**NO MORE
~~HATERS~~**

**CHARACTERISTICS AND
MECHANISMS OF ONLINE
COMMUNICATION THAT
PROMOTE HATE SPEECH**

~~CHARACTERISTICS AND MECHANISMS OF ONLINE COMMUNICATION THAT PROMOTE HATE SPEECH~~

The analysis shows that as soon as young people come to interpret that the online world has few limits or boundaries, they tend to view it as ungovernable, a “lawless state” where conflict is normalised and trying to steer clear of it is futile. In this respect, hate speech is seen as an unavoidable burden or the cost of being able to enjoy the freedoms offered by the Internet. Hence, almost all of the weight of regulation of people’s behaviour on the Internet is transferred to users, who learn by trial and error and through encounters with its darker side:

These days on the Internet there are absolutely no rules that people follow. People say and do whatever they want because social media let people remain anonymous. And of course, if no one knows who you are, what does it matter what you say, right? Personally, it seems to me that certain social media companies allow too much freedom, on Twitter for example. There is practically no censorship. And people are really toxic. [...]

I think that we have gotten to the point where “anything is valid”. There is no respect for others. On the Internet everything is valid, or at least that’s what it seems like.

Participant, online community, 18-20 year olds

Beyond the technicalities of regulating hate speech on social media platforms, there are a number of characteristics of online communication that can help to explain the explosion in hate speech in recent years. Firstly, despite increasing awareness of digital footprints,

many users feel that their online anonymity is completely protected. As technology makes it possible for users to separate online interactions from any sense of physical presence, a direct link is established between an anonymised subjectivity and social reality. This limits external judgements or the influence of social norms from the offline world and creates a sense of unrealness and distance from online actions and behaviour. The idea that online behaviour is not transposed or consequential to the offline world contributes to the prevalence of depersonalising and dehumanising actions. Hence, as anonymity provides a mask that emboldens users and permits them to act with a level and form of aggression that is not consistent with typical face-to-face communication, lack of corporeality invisibilises the consequences of online behaviour on social reality (*Bustos Martínez et al., 2019, p. 36*).

Secondly, the Internet and social media platforms have a soapbox effect. Global reach, flexibility, immediacy, and transience offer users enormous capacity to construct online identities and personas far beyond the possibilities of the offline realm. Sharing information and opinions online can entail a heightened visibility of the online self and an almost immediate engagement with other users, but also means that publications remain indefinitely accessible as they can often only be removed by a judicial order (*Gómez Martín, 2020, p. 412*). The configuration and design of digital platform interfaces were also identified by the study participants as aspects of online communication that foment the dissemination and visibility of hate speech. As concordance between users tends to be communicated quantitatively through tools such as 'like' and 'retweet' buttons and discordance qualitatively (through comments), the consequence is that hate speech acquires greater visibility and appears to be more prevalent.

Additionally, virtual spaces are prone to acting as 'echo chambers'. This means that each user engages in a process of personalisation and filtration of the online content that they are exposed to, leading to the configuration of an online reality that is highly segregated and ideologically biased. At the same time, the automated learning or algorithms used by Internet search engines and social media

platforms to personalise content can create 'filter bubbles' that inadvertently contribute to increases in segregated ideology and the transmission of biased information (*DiFranzo & Gloria-Garcia, 2017*). These mechanisms can affect any Internet user, but they are especially relevant when analysing extremist discourses (Flaxman et al., 2016) as they can have a domino effect that makes hate speech seem more widespread than in reality. In essence, these technological processes create virtual spaces in which individuals with similar ideologies interact and mutually reinforce hate discourses, while also creating a sense of belonging and community (*Bustos Martínez et al., 2019, p. 36*). As hate speech is often emotionally highly charged it also makes it even more difficult to control or decouple from processes of group affiliation and acceptance. Hence, when users with similar profiles interact in essentially closed network groups it potentiates and legitimises hate speech as well as providing the defence mechanisms or sense of security that reduces users fears of losing anonymity.

Thirdly, because it is easier than ever to generate proprietary content and to access the content of other users (*Garmendia et al., 2016*), the capacity for generating and disseminating manufactured information, hoaxes or fake news has increased enormously. Previously, the dissemination of misinformation was restricted by the narrower range of communication media and the source vetting procedures of traditional publishers. However, part of the problem also relates to the fact that many producers of hate speech feel that much of their communications go unnoticed within a medium where millions of message vie for attention and misinformation is abundant.

In most hate speech cases, especially those carried out by political organisations, their arguments are based on fallacies, half-truths and even outright lies and falsehoods. To stop this hatred, the propagation of lies and falsehoods must be cut off at the roots on the social media platforms.

Woman, online community, 18-20 year olds

Recent research with young people in Spain illustrates just how much conventional media have been displaced by social media platforms as the primary distributors of information to young people: three online platforms occupy second to fourth positions as the principal sources of information and news: social media platforms such as Twitter and Facebook (41.4%); family members, friends and acquaintances on instant messaging platforms (37.1%), and specific people on social media (32.7%). The webpages of traditional media occupied fifth place at 28.8% of respondents (Sanmartín et al., 2020). This new communication ecosystem has generated a crisis of confidence without precedence in journalism (Terol-Bolinches & Alonso-López, 2020) and highlights how much people need new competences when accessing, selecting and evaluating the trustworthiness of online information and news.

**NO MORE
~~HATERS~~**

TYPES OF ONLINE HATE SPEECH

~~TYPES OF ONLINE HATE SPEECH~~

5.1 Hatred of vulnerable collectives

As a hierarchical and unequal system that directly affects the configuration of hate speech, the virtual environment reproduces power relations and situations of vulnerability instituted in the offline world. Even if hate is not directed exclusively at vulnerable collectives, it does assume specific forms in these cases and can have the capacity to echo the institutional violence that these collectives suffer. On the contrary, when vulnerability is not as great, hate speech tends to be constructed in more explicit terms and present clearer expressions of threat and harassment. This is to say that hate speech directed at vulnerable groups does not have to contain overt expressions of hate, but can be constructed in neutral language (or images) that communicate hate through the context of their production, namely the power relations that they invoke.

Finally, it is important to take into account that framing a victim of hate within a vulnerable collective can encourage those that observe hate speech to personalise the violence they have witnessed. Paradoxically, the visibility of a collective can also impact the vulnerability of individuals at the same time as the collective loses capacity to generate empathy by personalising specific incidents of abuse or harassment. However, by situating the group at the centre, there is also a risk of attributing hate speech exclusively to group and public dynamics, which may invisibilise individual victims. Regardless, it is evident that it is necessary to visibilise vulnerable collectives as well as the structural and power dynamics that give rise to discrimination, marginalisation and abuse. But at the same time, it is also necessary to take into account how specific users, operating from their computers or mobile phones, bring these power relations into effect at local level.

5.2 Hate in public-private spheres

One of the key dimensions of hate speech is its impact in both the public and private spheres. Through viralisation, social media platforms make it possible to circulate messages quickly and far beyond personal circles. However, when a post containing hate speech goes viral it often produces a moment of perplexity for the poster. Initially, they may not even realise why they are receiving so many responses, notifications or comments. Subsequently, they may try to recover some control, to minimize impact, and to stop the exponential spread of the post by eliminating or silencing the publication (Expert 2). The impossibility of controlling a post once it has reached a certain magnitude means that it can generate a feeling of impotence as well as demonstrating how social media platforms can strip individuals of power.

Although hate speech is mostly communicated for the purposes of public effect, the people interviewed also spoke of how such messages reach them privately. In the case of women, this can be particularly intimidating, the content of messages is often more aggressive and personal; a palpable threat designed to generate fear. Apart from the form or explicit content of a message, such as insults, threats, etc., a number of other factors can construe messages as hate speech or harassment, including: the message format or the medium and manner of emission, such as sending of the same message repeatedly; the identity or social position of the sender or receiver may in itself be sufficient to constitute a message as hateful; and the intention of the communication, a component that is especially relevant to the impact of hate.

5.3 The normalisation of low intensity hate speech

Given the perception of the Internet as a lawless, difficult to regulate space where self-interest and ideological bias drives disinformation and fake news, a very significant risk exists that the presence of hate

speech becomes normalised in online spaces. Normalised discourses can be imitated and their significance underestimated, downplayed or stripped of racial, sexist (or other) charge. They can also be reconfigured within specific ideologies or reinforced through the process of sharing and viralisation. Hate generates hate, amplifying behaviour that would otherwise remain marginal. In this respect, hate speech becomes socially integrated and normative within radical subcultures. By consequence, people who grow up in such contexts may find it difficult to identify hate speech as it becomes interiorised at a subconscious level. This hidden disposition toward hate speech is therefore normalised by 'low intensity' cultural patterns. This form of hate employs stereotypes and prejudices that are particularly dangerous because of its high level of assimilation within cultural practices. This includes, for example, discrimination that is rooted in paternalism or the benevolent sexism that reinforces gender-based attitudes and ideas of women as dependent. Another example would be the use of racist premises or stereotypes. The normalisation of hate speech is also supported by its legitimisation through various social institutions and high profile social actors (media personalities, politicians and influencers, amongst others) that far from setting an example, give carte blanche to discriminatory and vexatious attitudes.

5.4 Hate speech from the perspective of victims

As mentioned earlier, the pervasiveness of hate speech forces Internet users to adapt to its presence. The fear and anxiety generated by hate modifies the way that users behave by forcing them to change the way that they interact online and, in particular, to carefully choose what they share with others. The outcome is a form of self-censorship as these users stop (Interview 6) or restrict the publication of content that has previously attracted hate speech, or they only post content when they feel emotionally strong enough to handle any backlash (Interview 1). Some users also attempt to restrict their exposure to hate speech by removing or blocking the access of particular perpetrators to their accounts (Interview 4). Others eliminate their

accounts and create new ones (interview 2) or stop using social media altogether. Changing accounts also means the creation of a new identity, user name and forms of posting. In other words, the construction of another avatar that the user hopes will not attract hate speech. However, users who create new accounts often include the old handle name in the subtitle of the new account in order to maintain a connection between the two 'selves' (Interview 2). Some of the most common actions mentioned by the study participants to avoid hate speech were to not upload personal photographs with friends or family (interview 3), not publishing images of their physical appearance or voice (interview 6), and reducing or stopping online activity, for example, by not publishing opinions, ignoring comments, and refusing to engage with users that they don't know (Interview 5).

Often, users construct such modification of behaviour as simply 'ignoring' hate speech or maintaining that they are 'not affected' by it or viewing it as a response to something that is 'not worth the trouble' of direct engagement. On the one hand, these strategies of avoidance have to be understood as forms of self-care that minimise exposure and soften the impact of hate speech. On the other hand, such modifications are also perceived as a form of imposed censorship. However, such adaptation sets out from an idealised premise of hate speech typologies and the assumption that it is easily identified. In reality, hate speech comes from multiple sources and directions and doesn't conform to any specific pattern that users can easily detect. In this sense, rather than being a process of adaptation, what users experience is a dilemma between avoiding self-censorship as a product of hatred and avoiding hatred by self-censoring. As such, beyond the individual experience of fear, this gives rise to a social and political construction that organises systems of oppression and that takes advantage of the idea that hate can extend beyond the screen to the offline world; one of the greatest fears of the study participants.

**NO MORE
~~HATERS~~**

**RESPONSIBILITY FOR
HATE SPEECH**

~~RESPONSIBILITY FOR HATE SPEECH~~

The question of who should bear the weight of responsibility for online hate speech is highly controversial. While there is a prevalent notion that the Internet is self-regulating and appeals to individual responsibility, there is also a common feeling of a lack of protection in online spaces. On the one hand, users are encouraged to report hate speech on the basis that it will make abuse and injustice visible and help to combat it. This is based on the assumption that the greater the volume of reported abuse the more effect it will have. On the other hand, this strategy may only be effective when it happens within close social circles. From a micro perspective, it is believed that if everyone was capable of controlling hate speech within their own close circles, we would be in a better position to control hate in general.

At a local level, a dichotomy can be observed in opinions on how hate speech should be dealt with. Some advocate that reporting abuse is not sufficient and that public intervention and rebuttal of hate speech is necessary. Others maintain that participation in such discussion only contributes to a feedback loop that encourages hate speech and makes it more visible, as well as running the risk of exposing people to violence. Another approach, which excludes perpetrators, is to place the focus on victims. The idea being that by standing by them publically, visibility is raised at the same time as providing support and reducing vulnerability. While there is general agreement that this approach is appropriate and responsible, it can also be problematic when the victim's close social network are not aware of the circumstances of the abuse and the perpetrators are unknown.

In terms of the responsibility of institutions and social media platforms, young people in the study believe that it is very difficult to be knowledgeable on the regulations and legal implications of hate speech, in particular as this relates to their own rights and responsibilities. Furthermore, they point out that a lot of hate speech goes unnoticed by the platform monitors responsible for reviewing content. Amongst the participants, there is a demand for filters that can sift out fake news or facilitate users to do so, but without the necessity of entering into disputes or having to expend too much energy attempting to identify the truth. While social media platforms argue that freedom of expression limits their capacity to control hate speech, the study participants view this as a refusal to accept responsibility, which ultimately contributes to its proliferation.

**NO MORE
~~HATERS~~**

**CONCLUSIONS
AND PROPOSALS**

~~CONCLUSIONS AND PROPOSALS~~

Hate speech is nothing more than the exteriorisation and reproduction, in a specific space, of problems and discrimination underlying the social world. There are some situations that generate greater vulnerability than others and some collectives that suffer much greater discrimination and intolerance compared to other groups. There is little doubt that controlling and regulating online hate speech is very difficult and that much work needs to be done. For one, it is particularly difficult to bring together social institutions from various levels (European, State, Local) with the aim of creating universal regulations while taking into account local and cultural particularities and diverse attitudes toward the control of speech and rights to freedom of expression. While some countries view legislative interventions as fair and necessary, others see them as overly zealous and an intolerable form of censorship. Additionally, it is very difficult to monitor and control online communication that is highly diverse but also characterised by patterns of hate speech that are not just textual but often only recognisable through the analysis of specific messages. Tied up in this is the problem of making moral, value-based and aesthetic judgements. To a large degree, these difficulties feed the collective imaginary with a discourse that constructs the task as indefinable and impossible to address. Furthermore, a recurrent idea persists amongst young people that social media platforms have forgone their responsibility to control online hate speech or at the very least are not doing enough. Another important factor that conditions actions is the generalised view that the Internet provides more advantages than disadvantages. From this perspective, the bad intentions of some people are perceived as a necessary cost of freedom of expression, which means, in effect, that freedom to express hate is defended.

Hence, more passive attitudes are evident amongst those that maintain that hate is a characteristic of the Internet that we must learn to live with. This is commensurate with the belief that acting or responding directly achieves little, particularly as the mechanisms for reporting and controlling hate speech are inefficient. This leaves mechanisms of emotional distancing that impedes identification with persons who experience hate as the only viable response.

Taking this into account, a number of strategies for dealing with on-line hate speech have been proposed (Bayer & Bárd, 2020). From a structural perspective these include activism, pedagogical proposals and institutional strategies designed to protect the vulnerable. There are also efforts to generate support networks that can mitigate the impact of hate speech and to give greater prominence to vulnerable communities in the development of appropriate responses.

Amongst the pedagogical strategies there are a number of important proposals and aims:

- Neutralising the elements that sustain hate through training in critical thinking, conflict management, learning that freedom of expression comes with responsibilities, managing difference and feelings of frustration, the possibility of being wrong and how to rectify or make amends, as well as the management of emotions and self-esteem.
- Setting out the bases of a technological and emotional education in how to use the Internet and social media that transcends the operative by taking a transversal approach in conjunction with educational values and the management of emotions in online contexts. This also implies that education and training in technology moves beyond questions of privacy, cyber security and addiction.

- Promoting alternatives to reactive discourses that are based on horizontal messages and mediated action within peer groups. In this case, if adolescents and young adults feel that they do not have a legitimate leading role and that actions or strategies are not their own, they will remain suspicious of any process of prevention and education that identifies them as affected.
- Use first person narratives that visibilise the personal consequences of hate speech.
- Placing an emphasis on the capacity and responsibility of individuals to break the chains of hidden and normalised hate within close social networks where a silent majority creates a stable environment for the persistence of hate. In other words, in the face of low intensity hate, generate awareness and high-density support.

Ultimately, there is a demand for institutionalised hatred to be countered by prevention and protection strategies from public bodies and the education system. This requires that all situations of vulnerability are treated equally and that human rights are viewed as the responsibility of everyone, regardless of personal experience of inequalities or injustices. In short, the need to create an institutional framework to protect people in vulnerable situations and to institutionalise responses to hate speech, which presently falls on the shoulders of the third sector or users themselves. Finally, it is important to point out that as much as the Internet can be a breeding ground for hate speech, it should also be seen as a space for reflection and transformation - a place where new movements can advocate for equality, an end to discrimination and condemnation, and the development of appropriate responses.

**NO MORE
~~HATERS~~**

BIBLIOGRAPHY

~~BIBLIOGRAPHY~~

Bayer, J., & Bárd, P. (2020). Hate speech and hate crime in the EU and the evaluation of online content regulation approaches. In *Policy Department for Citizens' Rights and Constitutional Affairs*. [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU\(2020\)655135_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/655135/IPOL_STU(2020)655135_EN.pdf)

Bazzaco, E., García-Juanatey, A., Lejardi, J., Palacios, A., & Tarragona, L. (2017). ¿Es odio? *Manual práctico para reconocer y actuar frente a discursos y delitos de odio*. Institut del Drets Humans de Catalunya y SOS Racisme Catalunya. https://www.idhc.org/arxius/recerca/1517393506-ES_ODIO_Manual_practico_vF.pdf

Bustos Martínez, L., De Santiago Ortega, P. P., Martínez Miró, M. Á., & Rengifo Hidalgo, M. S. (2019). Discursos de odio: una epidemia que se propaga en la red. Estado de la cuestión sobre el racismo y la xenofobia en las redes sociales. *Mediaciones Sociales*, 18, 25–42. <https://doi.org/10.5209/meso.64527>

DiFranzo, D., & Gloria-Garcia, K. (2017). Filter bubbles and fake news. *XRDS: Crossroads, The ACM Magazine for Students*, 23(3), 32–35. <https://doi.org/10.1145/3055153>

Fernández-Oruña, J. C., Jiménez, F. S., Sánchez, D. H., Ferrés, C. M., Villazala, T. F., Moreno, F. M., Anta, M. Y. S. A., García, M. R., Pérez, V. G., Orozco, A. M., & Martín, M. Á. G. (2019). *Informe de la evolución de los delitos de odio*. https://www.m-culture.go.th/mculture_th/download/king9/Glossary_about_HM_King_Bhumibol_Adulyadej's_Funeral.pdf

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(Specialissue1), 298–320. <https://doi.org/10.1093/poq/nfw006>

Fundación Secretariado Gitano. (2017). *Serie Materiales de Trabajo no 67: Guía para combatir el discurso de odio*.

Garmendia, M., Iglesias, E., Casado del Río, M. Á., & Mascheroni, G. (2016). *Net Children Go Mobile: Riesgos y oportunidades en internet y el uso de dispositivos móviles entre menores españoles (2010-2015)*. <http://www.ehu.es/documents/1370232/0/Informe+NET+CHILDREN+RED.ES+España+2016.pdf>

Gómez Martín, V. (2020). Odio en la Red. Una revisión crítica de la reciente jurisprudencia sobre Ciberterrorismo y Ciberodio. *Revista de Derecho Penal y Criminología*, 20(20), 411. <https://doi.org/10.5944/rdpc.20.2018.26458>

Gordo, Á., & Megías, I. (2006). Jóvenes y cultura Messenger. *Tecnología de la información y la comunicación en la sociedad interactiva*. Fad. INJUVE. <https://doi.org/10.5281/zenodo.3670899>

Hine, C. (2000). *Etnografía virtual*. Universitat Oberta de Catalunya.

Jubany, O., & Roiha, M. (2018). *Las palabras son armas. Discurso de odio en la red* (Issue February 2018). Edicions de la Universitat de Barcelona.

Richardson-Self, L. (2018). Woman-Hating: On Misogyny, Sexism, and Hate Speech. *Hypatia*, 33(2), 256–272. <https://doi.org/10.1111/hypa.12398>

Sanmartín, A., Ballesteros, J. C., Calderón, D., & Kuric, S. (2020). *De puertas adentro y pantallas afuera. Jóvenes en confinamiento*. Centro Reina Sofía sobre Adolescencia y Juventud, Fad. <https://doi.org/10.5281/zenodo.4054836>

Terol-Bolinches, R., & Alonso-López, N. (2020). La Prensa Española En La Era De La Posverdad : Verificación De Datos Para Combatir Las Fake News the Spanish Press in the Post-Truth Era : *Revista Prisma Social*, 31, 304–327.

**YOUNG PEOPLE ONLINE:
BREAKING THE CHAINS OF HATE SPEECH**

SUMMARY

**NO MORE
~~HATERS~~**

Break the hate chain!

Centro
Reina Sofía
sobre adolescencia
y juventud

fad

M MALDITA.ES

With the support of:

Google.org