



Informing Science: the International Journal of an Emerging Transdiscipline

An Official Publication
of the Informing Science Institute
InformingScience.org

Inform.nu

Volume 26, 2023

DEFINE AND TACKLE HATE SPEECH: THE EXPERIENCE OF SOCIAL WORKERS IN ITALY

Amalia De Leo*	Department of Psychology, Catholic University of Sacred Heart, Milan, Italy	amalia.deleo@unicatt.it
Emanuele Russo	C.I.F.A. ETS, Torino, Italy	russo@cifaong.it

* Corresponding author

ABSTRACT

Aim/Purpose	The aim of this qualitative study is to explore social workers' representations of hate speech (HS), the effects it has on the community, and socio-educational actions aimed at combating it.
Background	Hate speech is any form of communication that promotes discrimination, hostility, or violence towards individuals or groups based on their identity. Although its spread is facilitated by particular characteristics of the online environment (such as anonymity and ubiquity), HS has pervasive consequences even in offline reality. In the last year, several community-based projects involving social workers have been implemented to address the problem. Professionals who work with the community play a crucial strategic role in the fight against HS. Therefore, it is imperative to begin by considering their perspective to gain a better understanding of HS and how it can be controlled.
Methodology	Following a psycho-sociological perspective, six focus groups were conducted with 42 social workers (19 females and 23 males) belonging to associations or organizations of a different nature, such as NGOs, local social promotion organizations, universities, private social organizations, whose mission included the theme of countering hate speech.
Contribution	There are no studies in the literature that consider the views of operators working to counter hate speech within communities. Our study contributes to deepening the knowledge of the phenomenon and identifying the most suitable strategies to combat it, starting from an approach that does not only focus on the online or offline dimension but on an integrated "onlife" approach. The

Accepting Editor Francesco Tommasi | Received: March 17, 2023 | Revised: May 5, May 12, May 17, 2023 | Accepted: May 18, 2023.

Cite as: De Leo, A., & Russo, E. (2023). Define and tackle hate speech: The experience of social workers in Italy. *Informing Science: The International Journal of an Emerging Transdiscipline*, 26, 115-134.

<https://doi.org/10.28945/5120>

(CC BY-NC 4.0) This article is licensed to you under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). When you copy and redistribute this paper in full or in part, you need to provide proper attribution to it to ensure that others can later locate this work (and to ensure that others do not accuse you of plagiarism). You may (and we encourage you to) adapt, remix, transform, and build upon the material for any non-commercial purposes. This license does not permit you to use this material for commercial purposes.

	study offers an outline of how hate speech affects the daily lives of the communities in the cities of Torino, Palermo, and Ancona. Additionally, it proposes a grassroots strategy to address hate speech.
Findings	The results suggest that strategies effective in countering hate speech in offline contexts may not be effective in online environments. The technological revolution brought about by social media has significantly expanded the potential audience while weakening traditional communities. Addressing hate speech in the present context requires efforts to rebuild fragmented communities, gaining a thorough understanding of how the new virtual public space operates, and prioritizing hate speech as a specific concern only after these initial steps.
Recommendations for Practitioners	The study identifies a complex issue with hate speech that lacks a clear approach or shared interpretation. Social media's unique algorithmic structure makes it difficult to find effective solutions that worked in offline contexts. Counteracting hate speech requires systemic strategies, involving both online and offline concerted actions, recognizing the social network as a key component. This requires practitioners, organizations, and institutions to collaborate and share expertise for successful interventions.
Recommendations for Researchers	Hate speech represents a violation of human rights and a threat to freedom of expression. The spread of hateful messages has a significant impact on society, as it can negatively influence social cohesion, diversity, and inclusion. Understanding the causes and consequences of hate speech can help develop effective strategies to prevent and counter it, which is a crucial challenge for both research and society as a whole. Studying hate speech should involve the use of interdisciplinary methodologies.
Impact on Society	The study contributes to increasing the knowledge of the need to adopt an integrated approach to counter hate speech at the community level. Unlike other European countries, Italy appears to be less equipped to effectively understand the phenomenon and coordinate a response, starting from the absence of a national institute to track hate speech-related phenomena.
Future Research	Future research should focus on comparative analysis at the European Union level to assess the ability of civil society in other countries to develop effective strategies against hate speech.
Keywords	hate speech, online hating, community, social workers, activism, social media, educational strategies

INTRODUCTION

Hate speech has become one of the main topics of debate for public opinion in the last years and has been the subject of an increasing number of projects and programs aimed at countering it. This article tackles the issue starting from the direct experience of 5 groups of social workers in three Italian cities (Turin, Ancona, and Palermo) involved in the educational project #Hatetrakers, who have gained a structured experience over time in countering hate speech through a community approach.

Hate speech needs in-depth analysis to be understood, as it is a very complex phenomenon that is difficult to grasp from the very point of its definition. Generally, we refer to the use of linguistic attacks based on race, ethnicity, religion, and sexual orientation (Chetty & Alathur, 2018; Smolla, 1990). According to Donzelli (2021), this type of hate speech is persecutory, discriminatory, and demeaning in nature, and is often directed towards historically oppressed groups. Although there is no universally accepted way to define hate speech, one of the most accepted definitions is that of the European Commission against Racism and Intolerance (ECRI) General Policy Recommendation N°15 on

Combating Hate Speech (2016), which states that hate speech is “an advocacy, 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”, color, descent, national or ethnic origin, age, disability, language, religion or belief, sex, gender, gender identity, sexual orientation, and other personal characteristics or status” (p. 3). Several typologies have been identified, such as online religious hate speech, racial hate speech, misogynistic hate speech, and even online political hate speech (Chetty & Alathur, 2018).

In the Italian context, a turning point in the national debate, which generated the necessity to focus on the phenomenon in order to understand the risks to society as a whole, can be identified with the publication in 2017 of the final report of the Commission on intolerance, xenophobia, racism, and hate phenomena chaired by the President of the Chamber of Deputies, in which it is explicitly requested to promote the collaboration between the various stakeholders in the fight against the phenomenon of discrimination and hate speech such as research institutes, teachers, judiciary, law enforcement, civil society associations (Jo Cox Committee of the Chamber of Deputies, 2017) (Before then, references to the problem of hate speech were sporadic. In the final report, reference is made, among others, to the European Commission’s General Policy Recommendation No. 15 of the Council of Europe’s European Commission against Racism and Intolerance (ECRI) of 21 March 2016, which identifies hate speech as “the instigation, promotion or incitement of denigration, hatred or defamation of a person or group of persons, or the subjecting of such a person or group to bullying, harassment, insults, negative stereotyping, stigmatization or threats and includes the justification of these various forms of expression on a variety of grounds, such as ‘race’, color, language, religion or belief, nationality or national or ethnic origin, as well as ancestry, age, disability, sex, gender identity, sexual orientation, and any other personal characteristic or situation” (Lillian, 2007; Tsesis, 2002). The Commission’s work has had the merit of raising, for the first time at an institutional level, the problematic nature of a social phenomenon that was certainly not born in recent times, but which, thanks to the advent of social media, has become increasingly present both in political debate and in interpersonal relations *tout court* (Paz et al., 2020).

THE SOCIAL AND PSYCHOLOGICAL DIMENSIONS OF ONLINE HATE SPEECH

It should be emphasized that, when we speak of hate speech, we are not referring exclusively to emotion. The latter is defined, in fact, as an emotional expression of feelings and thoughts and aims to express a negative and hostile disposition, not necessarily permanent but very strong, towards someone or something (Malecki et al., 2021). Hate speech, and in particular online hatred, on the other hand, is defined “as the action of posting an explicitly negative evaluation of a person or object online primarily for the purpose of expressing one’s negative attitude towards that person, regardless of whether this causes actual harm to a concrete person” (Malecki et al., 2021). It is an emotion closer to anger whose goal is the destruction of the target rather than lasting discrimination (Ben-Ze’ev, 2008; Górska & Bilewicz, 2015; Tausch & Becker, 2013). In looking for the origin of hate speech in the psychological profile of the *hater*, some studies have found online haters would be characterized by poor self-identity, lack of self-awareness, lack of self-control, low self-esteem, and psychopathy (Bishop, 2013; Chao & Tao, 2012; Pace et al., 2021; Sorokowski et al., 2020). But it would not be correct to limit the problem of hate speech to the mere manifestation of antisocial and problematic personalities. According to Donzelli (2021), the origin of online hatred is linked to the growth of social media as a communication tool. Anonymity, temporal pervasiveness, indirect interactions, and the possibility of reaching anyone in any part of the hate globe (Hawdon et al., 2017) are the main characteristics of digital communication. The existence of a poorly or not at all controlled environment, where it is virtually possible to widen one’s basin of interlocution without exposing oneself personally, has facilitated the proliferation of racist, discriminatory, sexist, xenophobic, mi-

sogynist, homophobic ideologies (Costello et al., 2019) perpetrated mainly by political parties and extreme right-wing and/or religious fundamentalist groups. Indeed, the real social and problematic dimension of hate speech does not lie in the existence of the single isolated hater, even when serial, but in its organized dimension (Potok, 2015). Thanks to social media, organized groups of haters have a much greater capacity for incidence and pervasiveness than in the past, precisely because social media have the potential to reach an almost unlimited pool of people (Shahbaz, 2018). Moreover, because hate messages are characterized by strong emotionality they are spread with greater ease and speed by algorithms designed to encourage impulsive and non-rational reactions from users. Precisely because of the pervasiveness of the dissemination of these messages, we can say that the entire society is the target of online hate speech whose consequences transcend the network and invest real life (Douglas, 2007; Keipi et al., 2016). Online hatred, therefore, does not only affect the recipients of hate speech, who may manifest altered perceptions of issues, negative emotions, attitude changes, and even in changes in readers' commenting behavior (e.g., Flemming et al., 2017; Gervais, 2015; Rösner et al., 2016), but also the entire society which, often passively, witnesses the violence (Mullen & Smyth, 2004). Unlike cyberbullying or cyberstalking that targets individuals, online hate focuses on a collective (Räsänen et al., 2016). Some authors argue that increased exposure to hate language reduces sensitivity to it (Leets, 2002) and increases prejudice against the victims of the attack (Greenberg & Pyszczynski, 1985; Soral et al., 2018). It has also been pointed out that frequent exposure to hate speech may lead to the emergence of new norms of conduct that may perceive such language as morally justifiable and legitimate and thus lead to understanding such language as a norm rather than a criminal act (Bilewicz & Soral, 2020; Downs & Cowan, 2012). There are fears that such comments may decrease social cohesion and ultimately result in less prosocial behavior among citizens (Weber et al., 2020). Ultimately, exposure to hate materials has been linked to acts of violence (Foxman & Wolf, 2013; Kiilakoski & Oksanen, 2011).

STRATEGIES TO TACKLE HATE SPEECH: FROM REGULATORY ACTIONS TO BOTTOM-UP APPROACHES

To tackle this growing phenomenon, strategies have been implemented mainly on a regulatory/legal level (Bilewicz & Soral, 2020). In fact, it has been shown how a firm reaction by the relevant authorities can influence the greater or lesser spread of hate speech (Hawdon et al., 2017). It quickly became apparent that isolated actions, conducted solely at the regulatory/institutional level, are not able to structurally counter the spread of hate messages in the online world, nor mitigate their consequences outside the network.

Therefore, alongside these counter strategies, bottom-up experiences/projects involving local communities have been gaining ground in recent years. In particular, these are educational awareness-raising projects but also legal and psychological support projects that are carried out by civil society organizations such as associations, cooperatives, and international educational networks, which aim to strengthen society at different levels. Specifically, international and local non-governmental organizations prefer educational and awareness-raising approaches, mainly related to curricular Global Citizenship Education (see ParoleOstili, Amnesty International, C.I.F.A. ETS, ICEI), while some social enterprises (Chi odia paga) or law firms (Anti-Hater, Odiare ti costa) have focused more on legal assistance to victims. However, in most cases, these actions are carried out by organizations that do not have specific or historical know-how on hate speech but have decided to deal with it as an emerging problem. Moreover, the actors involved often belong to different organizational realities – sometimes with very different organizational cultures – and find themselves cooperating with each other without having in mind an unambiguous or shared representation of hate speech as it is strictly dependent on the area of expertise and the territorial context in which they operate.

For this reason, it is important to start from individual and organizational representations of hate speech and the impact it has on social contexts in order to identify the shared representations of the various social actors involved in counteracting it. Specifically, a qualitative study was implemented

within a psychosocial perspective (De Leo et al., 2022) that integrates psychological and social aspects in the analysis of human behaviors, interpersonal relationships, and group dynamics. This approach recognizes the mutual interaction between individuals and the social context in which they live, highlighting how psychological and social factors reciprocally influence each other with the aim of exploring:

- hate speech representation of operators working in educational and awareness-raising projects aimed at combating hate speech;
- the impact of hate speech on local communities; and
- strategies to combating hate speech from a socio-educational perspective.

MATERIALS AND METHODS

RESEARCH DESIGN OVERVIEW

This qualitative study utilizes Interpretative Phenomenological Analysis (IPA) to gain an in-depth understanding of social workers' lived experiences in addressing hate speech. IPA, as a research method, shares the epistemological assumptions of phenomenological and hermeneutic approaches (Smith et al., 1999; Smith & Osborn, 2003). It is a qualitative research method with specific rules and procedures, based on the phenomenological insights of Giorgi (1997). The primary aim of IPA is to describe and interpret the personal and idiographic dimensions of lived experience, as well as how individuals reconstruct and make sense of their stories. The interpretative phenomenological approach focuses on an idiographic level of analysis, examining specific and individual cases rather than general laws or principles. Furthermore, IPA research is interpretive in nature, following the principles of hermeneutics. The objective is to understand individuals' relationships with the world and the meaning they attribute to their experiences. This research approach involves a dual hermeneutic process: participants make sense of their world, and the researcher endeavors to understand the meaning participants assign to their world. Regarding the methodology, IPA utilizes interviews or focus groups to provide a structured yet flexible framework for the participants. This approach allows the researcher to avoid excessive intrusiveness, granting participants the freedom to share their stories. The researchers aim to immerse themselves in the participants' life world, considering them as the true experts on the research topic, capable of introducing novel perspectives and topics previously unexplored by the researchers. The goal is to achieve a comprehensive understanding of how participants interpret and make meaning of their experiences.

PROCEDURES

Participants were primarily recruited via convenience sampling through a network of the 'Hate Trackers' project, which aimed to implement educational initiatives for schools and local communities to counter hate speech. All project initiatives then converged on an *ad hoc* platform, www.hate-trackers.com, which has the task of presenting a multimedia and versatile approach to the problem, both in its online and offline manifestation. Before each Focus Group (FG), the professionals were explained the purpose of the research and how the work would be carried out. All participants gave written informed consent, agreeing to data being recorded, transcribed, and translated. Anonymity was explained and affirmed, as was the possibility of withdrawing from the study at any time. All FGs were conducted by a single researcher who traveled to the three surveyed cities and lasted approximately two and a half hours each. The researcher was trained in conducting this type of group interview. During the sessions, the moderator, by managing the speaking turns, ensured that all participants had the opportunity to speak. Data was collected based on principles of informed consent, following ethical principles of research in accordance with the Declaration of Helsinki guidelines. The research was approved by the Ethics Committee of C.I.F.A. ETS - Human Rights Education Office.

PARTICIPANTS

During the research, 6 focus groups were conducted with groups of professionals working in the cities of Turin (2), Palermo (2), and Ancona (1). A total of 42 professionals were involved (19 women and 23 men) belonging to associations or organizations of a different nature, such as NGOs, local social promotion organizations, universities, and private social organizations, whose mission included the theme of countering hate speech, as well as local administration representatives. The inclusion criteria considered for the professionals were that they: (1) were professionals working on projects aimed at countering hate speech; (2) represented the variety of characteristics specific to the social operators present in the territory; and (3) voluntarily participated in the study. On the other hand, the main exclusion criterion was the inability to participate in the focus groups during the study period.

For this study, we used the focus group which is a qualitative research technique that involves a small group of individuals discussing and sharing their attitudes, perceptions, and experiences related to a specific topic (Krueger & Casey, 2009). The discussion is moderated by a trained professional who guides participants to express their thoughts and feelings about the topic (Morgan, 1996). The FG track includes the items of five main thematic areas (see Table 1).

Table 1. Focus group track

Thematic areas	Questions
1. Presentation and professional role	<i>a. Tell me about yourself and your professional role</i>
2. Representation of hate speech	<i>b. In your opinion, what is hate speech?</i>
3. Resources and critical issues in hate speech polluted environments	<i>c. Have you ever worked in hate speech polluted environments? d. What are the resources and critical issues you found out?</i>
4. Impacts of online hate speech on territorial communities.	<i>e. In your opinion, is there a direct link between online hate speech and territorial communities?</i>
5. Intervention strategy involving territorial communities to counter hate speech	<i>f. Can intervention strategies involving territorial communities be activated to address hate speech? g. Do you believe that they are more effective than other strategies?</i>

DATA ANALYSIS

After the data collection phase, the five focus groups were entirely transcribed and a paper-paper content analysis was anonymously conducted. Our goal was to understand experiences and explore the meaning-construction processes of social operators, while considering the relevant sociocultural context as part of the data interpretation process (Brocki & Wearden, 2006; Smith, 2004; Smith & Osborn, 2003; Smith et al., 1999).

The data analysis was conducted following the principles of IPA. First, the transcribed data were examined to identify emerging themes. Subsequently, the themes were analyzed in depth to understand their significance for the practitioners' experience. From the analysis of the materials, a number of cross-cutting and redundant themes emerged from which the categories we will report below were generated (Hsieh & Shannon, 2005).

Following the data analysis procedure proposed by Braun and Clarke (2006), we first read several times the transcriptions of the focus groups to gain a solid familiarity with the content and understand the general topic (Familiarising with data). In the second step, the data were coded into categories that emerged from the data itself (Coding). Subsequently, the macro themes that emerged from the coded categories were identified and organized (Identifying themes). Codes were collapsed and renamed if they were deemed to describe similar topics or if they could be described in a clearer

manner. At this point, we provided a definition and name for each theme (Defining and naming themes). Once the themes were identified, we analyzed them in more depth to understand the meaning and context in which they were mentioned in the data (Analysis of themes). At this point, the results of the analysis were interpreted to understand the experiences and opinions of the participants (Interpretation). In the last step, we verified the interpretation by comparing the results with the data collected and the experiences of the participants (Verification). Data analysis was conducted by three independent coders (the agreement for each of the pairs of judges – AD and CG; AD and MP; CG and MP – was calculated, and after that, we calculated the mean value). Using ComKappa software we verified the goodness (Cohen’s $K = 83\%$) of the inter-rater reliability (Cohen, 1960; Robinson & Bakeman, 1998). Cases of disagreement were considered and discussed until consensus was reached.

RESULTS

We generated five main themes through data analysis. These were: (1) The roots of hatred; (2) Hate speech as a form of communication that incites online hatred; (3) Pervasiveness of hate speech; (4) Link between territorial communities and online hate; (5) Actions to counter hate speech. We also generated subthemes and included quotes to support the themes (see Table 2).

Table 2. Superordinate and subordinate themes

Superordinate themes	Subordinate themes
1. The roots of hatred	1.1 <i>Normalization</i> 1.2 <i>Triggers (stereotypes, fear, racial/ cultural prejudice, lack of empathy, fragile hater identity)</i> 1.3 <i>Intersectionality</i>
2. Hate speech as a form of communication that incites online hatred	2.1 <i>The role of social media</i> 2.2 <i>Individual vs. organized hate speech</i> 2.2 <i>Absence of digital education</i>
3. Pervasiveness of hate speech	3.1 <i>Anonymity</i> 3.2 <i>Distance from the other</i> 3.3 <i>Polarization</i>
4. Link between territorial communities and online hate	4.1 <i>Online/ real life interpenetration</i> 4.2 <i>Hate speech as a disruptive agent of real communities</i> 4.3 <i>Focus on community ties and intermediality</i>
5. Actions to counter hate speech	5.1 <i>Individual level: victim protection and working with the hater</i> 5.2 <i>Community level: rebuilding the social fabric in order to network with institutions</i>

1. THE ROOTS OF HATRED

1.1. Normalization

The social attitudes and tensions that provoke feelings of hatred online have their roots in society and are generally no different from those that fuel hate speech in direct relational dynamics. What is worrying today is the normalization of hate speech in public debate. According to the research participants, we live in a socio-political context that tolerates and encourages hate speech and hate actions.

“And then in my opinion [hatred] is very much naturalized by the system. [...] The problem [...] is so naturalized in communication, in the system, and therefore in society, that in any case to counteract it on the net [...] is even more difficult”. (F, Palermo)

1.2. Triggers (stereotypes, fear, racial/cultural prejudice, lack of empathy, fragile hater identity)

Hate speech is fueled by negative stereotypes that lead people to consider certain groups or individuals as inferior, different, or less worthy of respect. In fact, even when they are positive and can therefore represent a certain advantage, stereotypes become harmful if they are rigidly applied to individuals and serve as a pretext to justify different treatment or behavior towards certain groups. According to our participants, it is precisely the excessive use of labels used to classify people, which would lead to a hardening of categories, that would foster discrimination.

"We forget that we are people... we are categories". (M, Turin)

Hate speech is, therefore, the expression of cultural, social, and/or religious prejudice advanced by an individual as a reflection of fear of loss of belonging to a group, to his or her group (of whites, of men, of healthy, of heterosexuals, ...). If hate speech is not countered, it can propagate in society and, as a spiral effect, exacerbate stereotypes, increasing the risk of further abuse and insults, and sometimes physical violence.

"The concept is simple: when a person either doesn't know another reality or is already prejudiced against the other reality, he decides to take the easy way out from his point of view and distance himself from that reality by saying that his own identity, his own reality is the right one, it's the perfect one and everything else is not OK. And so he starts to hate". (F, Turin)

Prejudices contribute to marking differences by amplifying the distance from the other. According to our participants, it is precisely non-knowledge and non-identification with the other that generate feelings of hatred and contempt.

"That is, many different things we can accept but some we cannot accept at all. And this can be a bit of a cause, because if this rejection or non-acceptance accumulates over time it can express itself in the form of hatred". (M, Turin)

Moreover, according to a practitioner, the basis of hate speech is a fundamental lack of empathy and the inability to put oneself in the other person's shoes, which leads to a lack of understanding of the effects that a certain behavior might have on the other person.

"I think that when hate speech comes into play, there is also a lack of empathy with respect to putting oneself in the other person's shoes and reflecting on what resonances it may have. On an emotional level, in the other person, what I express or and instead put, I intend to hurt and so I express myself". (F, Ancona)

Almost all research participants seem to agree that what haters have in common is a fragile and fragmented identity. The hatred these people manifest stems from a disturbance or discomfort that they pour out on social networks, which are often used as containers into which they pour their psychic malaise. It is likely that the haters have themselves been victims of hatred or that they experience conditions of a strong identity and relational fragility.

"Very often the haters were themselves victims of hatred". (M, Turin)

Fear has been identified as one of the key factors in understanding hate speech and actions. In a time of economic and social crisis, which strongly questions the certainties to which we are accustomed, the Western citizen feels a more pressing need to preserve his privileges for fear of losing them. The other (different from me) may question who we are, and this would lead to the assumption of defensive and/or closed attitudes.

"Racism is fear of the other, but it is fear above all that the other will take something away from us or question things we believe to be so". (M, Turin)

1.3. Intersectionality

To properly comprehend the origins of hate speech, it is essential to adopt an intersectional perspective. This approach suggests that conventional views of societal oppression, including racism, sexism, homophobia, transphobia, xenophobia, and other forms of intolerance, do not operate in isolation. Instead, these various types of exclusion are interrelated and together create a system of oppression that reflects the intersection of multiple forms of discrimination.

This framework can be used to understand how systematic injustice and social inequality occur on a multidimensional basis. In this sense, it is crucial to name hatred and get to the root of privilege and power relations.

“The hatred that a lesbian woman experience is different from the hatred that a heterosexual woman experiences. We have to have an intersectional perspective and how we do it in a way that is not just ‘let’s increase the possible boxes of discrimination that we experience’ but let’s go a little bit to the root of what are privileges and power roles”. (F, Turin)

2. HATE SPEECH AS A FORM OF COMMUNICATION THAT INCITES ONLINE HATRED

2.1. The role of social media

The problem of countering hate speech today intersects with the dilemmas and contradictions of the digital age. The propagation of hate speech has been facilitated by the development and spread of technological tools. Incitement to hatred has always existed, it is something implicit in human nature and online hatred is no different in essence from what is commonly attributed to offline hatred. What has changed today, according to our participants, is that with the spread of the net, a space is provided to create discriminatory messages that can spread on a large scale in a short time and with little effort.

“The Internet is a potentially incredible tool. Everyone can say what they want, anything they want, there is potentially no censorship, anyone can do a Facebook Live and denounce something. But, like all potentially enormous tools, it has enormous criticalities”. (M, Ancona)

2.2. Individual vs. organized hate speech

According to most participants, online hate speech can be categorized into two types: individual and organized. Individual hate speech, also known as “bottom-up” hate speech, is perpetuated by single individuals who express their strong personal discomfort through hate speech because they lack strategies to deal with it or do not have a support network. This type of hate speech is considered an indicator of the unease felt by individuals who are unable to comprehend the complexity of the contemporary world and resort to hatred.

“One thing that is not really talked about is that there are different types of hate speech: one thing is the hate speech of the - forgive me - but of the frustrated, shut-in loser who hates everyone, hates himself, and says ‘you must all die’. This is hate speech the basic one practically, the one maybe who doesn’t understand the contemporary world, who doesn’t accept it, who doesn’t understand why masculine men don’t dress like masculine men anymore, females don’t dress like females anymore and so on”. (F, Palermo)

“Very often an internal frustration is discharged on this medium” (M, Palermo)

On the other hand, organized hate speech is typically perpetrated by professional haters who are often connected to political factions and use hate speech as a tool to create consensus. Organized hate speech serves as a forerunner and organizer by providing and massively spreading keywords and slogans that then go viral on the web. This type of language appeals to the deepest and most primal fears, aiming to stir up people’s unease and channel it into political movements that can take advantage of it.

Define and Tackle Hate Speech

“On the other hand, there is hate speech, the organized kind, and even this should not be underestimated because it often acts as a forerunner and organizer, above all providing the key words, the key concepts. These kinds of buzzwords, slogans, which are often slogans that target our belly, our deepest, most ancestral fears, are in fact very often artfully orchestrated”. (F, Palermo)

“The real haters are those who make this speech: divide and rule”. (M, Turin)

Moreover, the two types of hate speech are often intertwined, as individuals who are continuously exposed to hate speech in online communities begin to adopt and spread the expressions and slogans of hate speech unconsciously.

“This thing I write - like, for example, that immigrants all come to rape our women - it's not that I necessarily thought it. Somebon, I've heard it often, I've read it often”. (F, Palermo)

According to one speaker, hate speech is also a consequence of the misuse of digital tools. While the internet can be a powerful tool for communication and information sharing, it can also have negative effects depending on how it is used, such as in the spread of hate speech through social media and online games.

“So, hate speech, in my opinion, is really the consequence of a whole series of misuse, of a lack of inspiration, and from a certain point of view, also by those who over the years have exploited this potential to create distances”. (F, Ancona)

2.3. Absence of digital education

The propagation and spread of hate messages via the web can thus be seen as the outcome of the absence or insufficient digital education. The sudden spread of the web, in fact, has not been accompanied by adequate digital education that would have allowed and would allow, warnings against the risks of misuse of the web and its platforms.

“So, another fundamental aspect, which I think is important, is the lack of digital education, because the Internet is a tool. Very powerful, but it has been given, it has been used without having the idea or the responsibility to say OK, I learn this tool first, maybe I ask myself how it works”. (F, Ancona)

3. Pervasiveness of Hate Speech

3.1. Anonymity

In the online dimension, the possibility of anonymity and the perception of not being in the real world contribute to the emergence and development of hate speech: users feel freer to express their opinion, not being fully aware of the negativity and value of their message. The dissemination of hate messages is more tolerated on the Internet than in the offline world because it is subject to fewer controls. Insulting or harassing people online is easier because people often express themselves under the cover of anonymity and feel this carries less risk.

“... in the sense that there is this lack of empathy, this anonymity”. (F, Turin)

3.2. Distance from the other

Undirected communication, i.e., mediated by a screen, creates a distance with the interlocutor that seems to legitimize and facilitate the use of incorrect language aimed at directly attacking the person. The screen creates a distance that penalizes identification with the other; the other is no longer considered as a physical person but as a mere user, an avatar.

“A linguistic trend that makes use of a screen, which is definitely the screen where you don't see yourself, don't let yourself be seen, etc. that exploits - as someone said - the fact of using dead content instead of live content. I mean, when you're live you behave differently”. (M, Turin)

“It is easier to hate online ... you when you write a comment don't think of the other person who is reading your comment as a physical person, but as a user, a FB profile, an Instagram profile”. (F, Turin)

“There is a kind of screen that gives the possibility to unleash aggression”. (F, Palermo)

In the pre-social media era, it was necessary to expose oneself in person (i.e., physically) in order to express one's opinion and this certainly implied greater awareness and self-control on the part of the person. In the online dimension, however, with the possibility of anonymity and the perception of not being in the 'real' world, users feel more legitimized to express their opinion, sometimes not being fully aware of the negative consequences of their message.

“Before there was more effort needed ... to go and shout in the square you had to wear a shirt”. (M, Palermo)

3.3. Polarization

More and more often people today identify (psychologically) with a group or a party, and there never seems to be, or rarely is, any real confrontation on the issues. One focuses not so much on what unites but on what divides us from others, who become enemies. Today, people very often do not vote for one party, but rather against the opposite party. Groups compete for resources, among which is also consensus.

“There is such an approach as ‘political cheering’; you first decide which side you are on ... If this thing is promoted by a party that is not mine, it doesn't belong to me, I don't like it”. (M, Ancona)

A relevant characteristic of social platforms is their ability to polarize ideas. People are more polarized the more they use very few if any, sources of information. If the news, in fact, tends not to confirm the group identity, the individual disregards it, or even ends up convincing himself that the source (e.g., a newspaper) publishes false news (thereby reinforcing his own polarization). One index of polarization, therefore, is the fact of following few if any sources of information. A citizen who, on the other hand, has a more varied 'media diet' tends to be less polarized since he receives various facets of the news by consulting different sources. In this perspective, one can, therefore, say that social networks become dangerous when there is an implication of organized groups that tend to enclose users in 'bubbles' through, for example, micro-targeting, which allows messages (of different kinds) to be addressed to extremely small categories of people, right down to the individual, in order to manipulate them.

“Social media are very prone to polarization, they are exploited to leverage on what are then gut issues, i.e. the fear that has characterized our lives a little in this year, the initial distrust of the other and therefore everything that is unknown, that is frightening, so it is much easier to say no or I don't like it, rather than maybe asking a question and listening to what the other person has to say?” (F, Turin)

4. LINK BETWEEN TERRITORIAL COMMUNITIES AND ONLINE HATRED

4.1. Online/real life interpenetration

Activity on the web is vast and continues to increase; it increasingly characterizes modern society, but it should not be perceived as a reality detached from real life, where the norms governing human behavior do not apply. The virtual existence of everyone is closely related to his or her 'real' identity. These two spaces of life are not entirely disconnected from each other: the virtual world has simply become an important part of the real world. For this reason, online hate speech often has consequences for our daily lives: people, feelings, experiences, and dynamics are the same, both online and offline. According to the research participants, there is a constant reflection of online life on offline life as the two worlds are not split but complementary. There is a constant slippage from the online

to the offline dimension. Therefore, it is important to be aware that the propagation of discriminatory attitudes online can legitimize the adoption of behavior in line with this attitude also in an offline dimension. In other words, hatred is self-feeding in an online/offline vicious circle.

“Once you cross a bank on one side, you can also cross it in the presence. I am almost 40 years old; I know that 25 years ago some things would never have been said out loud whereas now they are said. Maybe because in the meantime there has been a shift from online to presence”. (M, Palermo)

4.2. Hate speech as a disruptive agent of real communities

If left unchecked when spread online, hate speech also has negative consequences in the offline world, as it contributes to exacerbating racial tensions and other forms of discrimination and violence. Moreover, the potential ability of hate to spread rapidly in the virtual world increases its harmfulness.

“The online will gradually become more and more linked to territorial communities”. (M, Turin)

4.3. Focus on community ties and intermediality

Since there is a strong correlation between the two worlds, working with the territorial community can also have an impact on the online community. In fact, according to the operators interviewed, in order to combat HS, a work of reconstruction and reconstitution of ties and closeness in the territorial community is fundamental, in addition to online action.

“So I think that nowadays the online dimension and the reality in which we live physically influence each other. The difference is the target audience. With young people, in my opinion, you can also think, with appropriate tools, of starting from the online and then getting to the offline, let's call it that. With older categories of people, on the other hand, I would always continue from the territorial communities”. (F, Turin)

5. ACTIONS TO COUNTER HATE SPEECH

The participants in our research offer a rich and articulate reflection on what methods can be used to combat HS. In particular, the proposals provided are developed on two levels of operational action: one more related to working with individuals (e.g., the victim, but also the hater) and a more community/collective and network dimension involving various actors such as associations, institutions, and groups in the area.

5.1. Individual level: victim protection and working with the hater

Regarding the first (individual) level, a first and fundamental need that emerges is to protect and provide support to victims of hate speech. Unfortunately, the legislative framework has not evolved at the same pace as the use of the web and social media has spread in recent years, so the legal instruments for combating online hate that we have at our disposal today are often inadequate and insufficient. Victims, in particular, cannot rely on regulations that can effectively protect them. It is therefore necessary to introduce updates to existing regulations that allow for swift and agile action against hate speech and that offer special attention to the protection of hate speech victims.

“Already knowing that one can proceed rather than just suffer is already a tool that can be handed over to the victims”. (M, Palermo)

“Do you go to the police? Do you go to the carabinieri? Do you go to the police station? To the public prosecutor's office? We don't know”. (F, Palermo)

Alongside action on a legal-regulatory level, there is a need to set up listening areas, ‘access points’, which welcome victims and offer both psychological support and accompaniment to the victim throughout the complaint process.

“There could be spaces, yes at the level of let’s say counters, someone, although even the word counters is something to be used carefully, that’s why I used ‘available spaces’. That is, maybe an access point that says yes, OK, look for this thing you can do so for this thing you can report”. (F, Palermo)

The listening desks, or access points, are fundamental because they can also act as a link with other actors in the area, promoting the reintegration of the victim into the community fabric from which in many cases they had left.

“This could be a tool you know can help; turning to a community and thus feeling less alone when facing the problem”. (F, Palermo)

An original and innovative position adopted across the board by almost all our interviewees concerns the approach to haters. The hater is seen as a victim himself; in particular, haters are themselves victims of their own fear, their lack of culture, and their lack of empathy. They are fragile and in order not to show themselves this way, they identify with ideologies they consider strong. To better understand the reasons that lead them to perpetuate offensive or aggressive behavior, our interviewees suggest putting themselves in their shoes and trying to embrace their fragility.

“As for the rest, I would like to play devil’s advocate a little bit and propose an insinuation, in the sense that I would like to break a lance in favor of the haters. A bit like Emilia said, that is, in my opinion, everything starts from an upstream discomfort, in the sense that haters are very often themselves bullied by the role of the alpha male ... So bullying, mobbing, mobbing in the workplace, so there is that component there anyway”. (M, Turin)

In this sense, prevention and countermeasures should above all involve the psychological and socio-cultural dimension of people, in order to initiate a process of de-legitimization of the violence that increasingly often manifests itself unchecked. Therefore, it is so necessary to work on people’s emotional intelligence. Understanding one’s own emotions and knowing how to manage them encourages empathy towards others and the implementation of more proactive behavior.

“... it is important to work a lot on people’s emotional intelligence. Because the moment you know how to recognize your emotions and you know how to put yourself in the other person’s shoes, so it’s not just a question of empathy but also recognizing how you are. Already there you have done a piece of work”. (F, Turin)

An effective response to online hatred is through confrontation and dialogue with the proponent of hate speech. For example, one can reply to the author of discriminatory messages by responding in private to the dominant negative narrative by proposing an alternative view. It is as if on social platforms one loses the dimension of human contact that one then regains when trying to re-establish a more direct relationship with the person. According to one interviewee, even a simple switch from writing comments under a post, to a personalized interaction through chat, would be sufficient to re-establish contact with the other. It, therefore, seems essential to work on the interpersonal relationship that would favor identification with the other.

“If in the comments they expressed the maximum of hate speech, i.e., offense, it is as if there was no me on the other side. There was no human being but simply a bot, as if the comments I was replying to, that is, my replies were automatic. The moment someone came by to write to me privately in chat there was a face to face. So yes, I was the HateTrackers profile, but I was replying on a first name basis and you could see that there was someone there, my replies were much more contextualized to what I was receiving as questions and answers from the user. The tones changed, in the chat room the users from the aggressive tone they had in the comments went back to being, as it were - they recognized my status as a human being on the other side, so ‘ah ok, there is someone who is actually reading what I write, who is actually commenting on what I think’ and so they became tamer”. (F, Turin)

5.2. Community level: rebuilding the social fabric in order to network with institutions

A second level of action outlined concerns the possibility of acting at a more community level.

“Either the approach is communitarian or there is no possible approach to this”. (M, Palermo)

First and foremost, there is an urgent need to start again from the territorial communities, trying to work on the elements of continuity and transversal similarities in order to reconstitute the social bond of a community fabric that appears increasingly split and fragmented. This, however, requires a long time and constant work.

“From an operator’s point of view, I would focus on the territorial community, while the online community in an auxiliary, compensatory form as if the virtual one was the projection of the territorial one. For what reason? Because also based on some work done, the territorial community has many elements - being a geographically invariable territory - it is a fixed point. There are so many fixed points of reference. Because in order to do all the work - what I was saying, integration, mediation, cohesion - we need stable points of reference, otherwise we get lost along the way. The territory itself is a common good with everything that makes it up: history, language, culture, people. It becomes a common good for all migrant communities. In Turin I know there must be at least 200 different nationalities. But if they live in Turin, Turin is the same for everyone. So, we have more references, more arguments to interact, where on a virtual community it is difficult to have all these elements. But it can be compensatory, where the physical territorial space does not allow, we can lean on the compensatory virtual space. I would work in this form. Even virtual communities when creating discussions must have stable references in relation to the territorial one”. (M, Turin)

When there is insufficient engagement to be able to prepare actions, it is useful to start with the elements of continuity/similarity in order to eventually undermine even those differences that appear very deep-rooted and more difficult to question. In other words, the elements of connection act as a bridge and help to undermine even those differences that seem insurmountable.

“My closest friend who we grew up with, the one with I first talked about homosexuality. We were really young; we were 17-18 years old. He told me something very interesting: he said, you know, I must thank you because if I hadn’t - apart from affection, apart from friends - known you and your things ... I would probably grow up as an ordinary heterosexual. So that transition between me and him, man, that was important. For me as a welcoming thing but also for him because then he multiplied it. That’s why the personal relationship is fundamental”. (M, Turin)

To sum up, according to our participants, given the complexity of the phenomenon, there is a need to act on several levels (individual, group, and community) in order to prepare actions that are truly effective in combating hate speech. Moreover, one cannot disregard networking and synergic collaboration between the various actors in the area and a community approach that favors generative coexistence between the different communities if one wants to interact effectively with institutions to realize durable changes.

DISCUSSION

The participants in the study identified hate speech as an expression of social and/or religious cultural prejudice advanced by an individual as a reflection of the fear of loss of belonging to a group. If not countered, hate speech can propagate in society by exacerbating stereotypes, increasing the risk of abuse and insults, and leading to physical violence (Greenberg & Pyszczynski, 1985; Soral et al., 2018). *“The concept is simple: when a person either doesn’t know another reality or is already prejudiced against the other reality, he decides to take the easy way out from his point of view and distance himself from that reality by saying that his own identity, his own reality is the right one, it’s the perfect one and everything else is not OK. And so he starts to hate.”* Non-knowledge, non-identification accompanied by a lack of empathy and an inability to put oneself in one’s neighbor’s shoes would be strong drivers of hate and contempt generation. Along with this, as Malecki et al. (2021) also argue, explicit intent to hurt has been identified in hate speech. Regarding the identity of haters, as also emphasized by the above-mentioned authors, all the participants in the focus groups seem to agree that haters may possess a fragile and fragmented identity, suffering from major disorders and discomforts, perhaps themselves victims of hatred, which they pour out onto social networks in the form of hate speech (Barlow & Awan, 2016).

Fear of difference is identified as another key factor, a tool used by nationalist and populist propaganda in times of economic and social crisis (Modood, 2020). The participants identified two types of hate speech: a more individual type perpetuated by single individuals and a more structured/organized type perpetuated by organized groups. According to the focus group participants, hate speech produced by individuals would be an indication of strong personal discomfort, of *“those who do not understand the contemporary world, who do not accept it, who find it hard to accept diversity and accept it as such without pointing it out with discriminatory forms”*.

The other type of hate speech, according to the participants, is more political in nature. As such it would be more organized, structured, and used by social/political groups as a propaganda tool, aimed at manipulating discourse and creating consensus (Douglas, 2007). It has often been noted that the two types of hate speech are intertwined so that individuals begin to adopt certain expressions and ideas from discriminatory dialogues and slogans to which they are exposed. About the propagation of hate speech, it is argued that social media has played a central role in providing a strong impetus for propagation (Hawdon et al., 2017). What the online network has created is a space where hate speech can spread much faster on a large scale and without/almost no censorship given the poor regulation of social media (Farkas & Schou, 2018). Furthermore, in the online dimension, the possibility of anonymity and the perception of not being in the real world contribute to the emergence and development of hate speech: users feel freer to express their opinion, not being fully aware of the negativity and value of their message (Del Vicario et al., 2017; Lorenz et al., 2011). Moreover, in a context of non-direct, screen-mediated interactions, distances are created between interlocutors, and these seem to legitimize or at least facilitate the use of violent and discriminatory language (Bargh et al., 2002). Another identified feature of social media is polarization and media bubbles by algorithms and the action of organized groups that tend to create bubbles around users through the action of microtargeting to manipulate a group or individuals.

Finally, the participants argue that real and virtual identities are not two separate entities, but complementary. Therefore, the propagation of discriminatory attitudes online, as cited by other authors, can legitimize the adoption of similar behavior also in the offline dimension, self-feeding hatred in a vicious online/offline circle (Bilewicz & Soral, 2020; Downs & Cowan, 2012). A commonly thought solution, effective action against hate speech would be to work on rebuilding and reconstituting ties and closeness in and between territorial communities in order to bring change to the online community as well: *“Starting from the territorial community, working with people face to face you can somehow reach the online community as well. And change that part too”*. Proposals for actions to combat hate speech, according to the input received, should develop on two levels: one individual and therefore protection and support for victims but also for the perpetrators of hate speech themselves (who are also considered victims of hate speech), pushing for the improvement of the legislative framework on the issue, providing the activation of listening desks and the reintegration of the victim into the community fabric. Another issue, as far as haters are concerned, is to work on the psychological level of emotional intelligence. A second level on which it seems important to act is at a more community level, starting first from the territorial communities, working on the elements of continuity and transversal similarity in order to reconstitute the social bond of the community fabric.

PRACTICAL IMPLICATIONS

The picture that emerged from the FGs and the project, in general, reveals the existence of a problem that is by no means obvious, but toward which there is still no unambiguous approach or shared key of interpretation. There are several reasons for this multiplicity of readings: on the one hand, while hate speech is not a new issue at the societal level, the incidence and pervasiveness given by social networks is new. On the other hand, precisely because of the peculiar structure of the algorithms that regulate the dissemination of content on different platforms, it happens that potentially effective solutions in life contexts not mediated by technology have ambivalent results on social networks. A

paradigmatic example is the reaction to a hate message carried out directly on a social network. Certainly, countering it prevents online discourse from becoming the exclusive preserve of “hater” users or their acolytes; however, it risks generating escalations of aggression and attacks (shitstorms) that risk isolating the victim and allowing the hate message to spread even faster.

It follows that an analysis aimed at countering hate speech cannot be limited to the mere linguistic dimension. A relevant part, if not preeminent in some contexts, is the analysis of the tool itself, the social network, and the levers of the system that sees the social network as its cornerstone. In the absence of an in-depth understanding of how it is the social network system that makes use of hate speech, and not vice versa, it will not be possible, in the countering or educational actions put in place to counter the phenomenon, to go beyond a limited work of educating people about the emotionality and social meaning of words. Therefore, it is evident how hate speech, being configured as a social problem that transcends the place of its manifestation, must be addressed with systemic strategies involving concerted actions both online and offline. In other words, hate speech in its contemporary manifestation cannot disregard an *onlife* approach. In order to do so, it is necessary for practitioners and organizations to network and for there to be recognition and sharing of expertise, and for institutions to be involved throughout the process (De Leo et al., 2023).

LIMITATIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

The present research provides a preliminary analysis of the perception of the impact of hate speech on local communities and a first attempt to develop an integrated approach to addressing hate speech based on the representations of the social workers involved in the #hatetrackers project. The study is not free of limitations, as the choice of cities was based on the practical feasibility of the action by the implementing subject (C.I.F.A. ETS) and not on prior research verifying the greater representativeness of these urban areas compared to others at the national level. Regarding the selection of participants for the focus groups, it should be noted that the #hatetrackers project is of an educational nature; therefore, the professionals involved all belonged to the educational-recreational field and not to other potentially interested professional areas (lawyers, linguists, psychologists, etc.). Future research could involve these professionals to investigate the differences in combating hate speech and understand how these different approaches can integrate with each other. It should be emphasized that the staff involved was selected primarily for its optimal characteristics and function to the success of the educational project and not to compose a group that is as diverse and statistically representative as possible. However, with regard to this last point, it should be noted that in selecting participants, efforts were made to create groups that were as responsive as possible to the variety found in the territory. Another aspect that could be investigated in future research is the existence of specific discriminating variables in different territories (e.g., size of urban areas, experiences with hate speech prevention projects, census, etc.).

CONCLUSION

The most significant problem with the spread of hate speech may initially appear to be relational; in reality, its most insidious component is the organized component as it threatens the internal cohesion of communities themselves. The political use of hate speech, amplified by the spread of social media, allows the realization of permanent alterations in the social and political life of a community that historically was only feasible by having massive economic resources and/or the near monopoly of mass media. Facebook, Twitter, Instagram, Snapchat, TikTok, and YouTube, through their commercial policies implemented through the operation of algorithms, have in fact played an important role in amplifying and mediating old and new forms of abuse, hatred, and discrimination (Matamoros-Fernández, 2017; Noble & Tynes, 2016; Patton et al., 2017). It is, however, the effect of a much broader systemic process. A possible interpretation may be that the global capitalist society, secularized and deprived of most of the social glues found in institutions such as the extended family and communities of reference (neighborhood, local associations, recreation centers, etc.), has accentuated

its fragmentation process with the advent and success of social media. Certainly, the pandemic from Covid-19 made the process systematic and finalized the identification of the individual as the basic unit of the social system. To put it differently, the modern individual is no longer considered the fundamental and inseparable element of a complex system based on human rights in a democratic society. Instead, the individual's identity is defined by their role as a consumer and the goods they consume, leading them to interact with institutions and corporations. The social media system, having stripped the individual of any affiliation or relationship of meaning, retroacts on communities by isolating individual constituents who have value only as they buy consumer goods, causes, attitudes, or as they provide information useful for the refinement of profiling processes in the hands of commercial companies and state security systems. Thus, there seem to be two parallel processes at work. On the one hand, that of the progressive isolation of the individual, who perceives himself as a consumer and is perceived as both consumer and product. On the other, that of a society mithridatized (addicted) to hate speech, accustomed to seeking not peaceful confrontations and acceptance of differences but zero-sum games in which, as a group or as an individual, one can be either submerged or saved. To face both, the community must regain its internal cohesion by perceiving itself no longer only as a local group but also as a virtual community. By adopting multidimensional strategies and solutions (Tommasi, 2023) that once again place the dignity of the individual at the center, it will be able to gradually neutralize the destructive effect of hate speech.

REFERENCES

- Bargh, J. A., McKenna, K. Y., & Fitzsimons, G. M. (2002). Can you see the real me? Activation and expression of the “true self” on the Internet. *Journal of Social Issues*, 58(1), 33–48. <https://doi.org/10.1111/1540-4560.00247>
- Barlow, C., & Awan, I. (2016). “You need to be sorted out with a knife”: The attempted online silencing of women and people of Muslim faith within academia. *Social Media + Society*, 2(4). <https://doi.org/10.1177/2056305116678896>
- Ben-Ze'ev, A. (2008). Hating the one you love. *Philosophia*, 36, 277–283. <https://doi.org/10.1007/s11406-007-9108-2>
- Bilewicz, M., & Soral, W. (2020). Hate speech epidemic: The dynamic effects of derogatory language on intergroup relations and political radicalization. *Political Psychology*, 41, 3–33. <https://doi.org/10.1111/pops.12670>
- Bishop, J. (2013). The effect of de-individuation of the internet troller on criminal procedure implementation: An interview with a hater. *International Journal of Cyber Criminology*, 7(1), 28–48.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brocki, J. M., & Wearden, A. J. (2006). A critical evaluation of the use of Interpretative Phenomenological Analysis (IPA) in health psychology. *Psychology & Health*, 21(1), 87–108. <https://doi.org/10.1080/14768320500230185>
- Chao, C.-H., & Tao, Y.-H. (2012). Human flesh search: A supplemental review. *Cyberpsychology, Behaviour, and Social Networking*, 15(7), 350–356. <https://doi.org/10.1089/cyber.2011.0481>
- Chetty, N., & Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression and Violent Behavior*, 40, 108–118. <https://doi.org/10.1016/j.avb.2018.05.003>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Costello, M., Rukus, J., & Hawdon, J. (2019). We don't like your type around here: Regional and residential differences in exposure to online hate material targeting sexuality. *Deviant Behavior*, 40(3), 385–401. <https://doi.org/10.1080/01639625.2018.1426266>

Define and Tackle Hate Speech

- De Leo, A., Cotrufo, P., & Gozzoli, C. (2022). The refugee experience of asylum seekers in Italy: A qualitative study on the intertwining of protective and risk factors. *Journal of Immigrant and Minority Health, 24*, 1224–1236. <https://doi.org/10.1007/s10903-021-01296-3>
- De Leo, A., D'Adamo, G., Morozzi, C., & Gozzoli, C. (2023). Taking care of forced migrants together: Strengths and weaknesses of interorganizational work from the perspective of social workers. *International Journal of Environmental Research and Public Health, 20*(2), 1371. <https://doi.org/10.3390/ijerph20021371>
- Del Vicario, M., Zollo, F., Caldarelli, G., Scala, A., & Quattrociocchi, W. (2017). Mapping social dynamics on Facebook: The Brexit debate. *Social Networks, 50*, 6–16. <https://doi.org/10.1016/j.socnet.2017.02.002>
- Donzelli, S. (2021). Countering harmful speech online: (In)effective strategies and the duty to counterspeak. *Phenomenology and Mind, 20*, 76–87. <https://doi.org/10.17454/pam-2007>
- Douglas, K. M. (2007). Psychology, discrimination and hate groups online. In A. Joinson, K. McKenna, U-D. Reips, & T. Postmes (Eds.), *The Oxford handbook of internet psychology* (pp. 155–164). Oxford University Press.
- Downs, D. M., & Cowan, G. (2012). Predicting the importance of freedom of speech and the perceived harm of hate speech. *Journal of Applied Social Psychology, 42*(6), 1353–1375. <https://doi.org/10.1111/j.1559-1816.2012.00902.x>
- European Commission against Racism and Intolerance. (2016). *ECRI General Policy Recommendation No. 15 on Combating Hate Speech*. <https://rm.coe.int/ecri-general-policy-recommendation-no-15-on-combating-hate-speech/16808b5b01>
- Farkas, J., & Schou, J. (2018). Fake news as a floating signifier: Hegemony, antagonism and the politics of falsehood. *Javnost – The Public, 25*(3), 298–314. <https://doi.org/10.1080/13183222.2018.1463047>
- Flemming, D., Cress, U., & Kimmerle, J. (2017). Processing the scientific tentativeness of medical research: An experimental study on the effects of research news and user comments in online media. *Science Communication, 39*(6), 745–770. <https://doi.org/10.1177/1075547017738091>
- Foxman, A. H., & Wolf, C. (2013). *Viral hate: Containing its spread on the Internet*. Macmillan.
- Gervais, B. T. (2015). Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology & Politics, 12*(2), 167–185. <https://doi.org/10.1080/19331681.2014.997416>
- Giorgi, A. (1997). The theory, practice, and evaluation of the phenomenological methods as a qualitative research procedure. *Journal of Phenomenological Psychology, 28*(2), 235–260. <https://doi.org/10.1163/156916297X00103>
- Górska, P., & Bilewicz, M. (2015). When “a group in itself” becomes “a group for itself”: Overcoming inhibitory effects of super-ordinate categorization on LGBTQ individuals. *Journal of Social Issues, 71*, 554–575. <https://doi.org/10.1111/josi.12128>
- Greenberg, J., & Pyszczynski, T. (1985). The effect of an overheard ethnic slur on evaluations of the target: How to spread a social disease. *Journal of Experimental Social Psychology, 21*(1), 61–72. [https://doi.org/10.1016/0022-1031\(85\)90006-X](https://doi.org/10.1016/0022-1031(85)90006-X)
- Hawdon, J., Oksanen, A., & Räsänen, P. (2017). Exposure to online hate in four nations: A cross-national consideration. *Deviant behavior, 38*(3), 254–266. <https://doi.org/10.1080/01639625.2016.1196985>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Jo Cox Committee of the Chamber of Deputies. (2017). *Final Report of the Commission on Intolerance, Xenophobia, Racism and Hate Phenomena*. <https://www.camera.it/leg17/1313>.
- Keipi, T., Näsi, M., Oksanen, A., & Räsänen, P. (2016). *Online hate and harmful content: Cross-national perspectives*. Routledge. <https://doi.org/10.4324/9781315628370>
- Kiilakoski, T., & Oksanen, A. (2011). Soundtrack of the school shootings: Cultural script, music and male rage. *Young, 19*(3), 247–269. <https://doi.org/10.1177/110330881101900301>

- Krueger, R., & Casey, M. (2009). *Focus groups: A practical guide for applied research*. Sage.
- Leets, L. (2002). Experiencing hate speech: Perceptions and responses to antisemitism and antigay speech. *Journal of Social Issues, 58*, 341–361. <https://doi.org/10.1111/1540-4560.00264>
- Lillian, D. (2007). A thorn by any other name: Sexist discourse as hate speech. *Discourse & Society, 18*(6), 719–740. <https://doi.org/10.1177/0957926507082193>
- Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences, 108*(22), 9020–9025. <https://doi.org/10.1073/pnas.1008636108>
- Malecki, W. P., Kowal, M., Dobrowolska, M., & Sorokowski, P. (2021). Defining online hating and online haters. *Frontiers in Psychology, 12*. <https://doi.org/10.3389/fpsyg.2021.744614>
- Matamoros-Fernández, A. (2017). Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication & Society, 20*(6), 930–946. <https://doi.org/10.1080/1369118X.2017.1293130>
- Modood, T. (2020). Islamophobia and normative sociology. *Journal of the British Academy, 8*(1), 29–49. <https://doi.org/10.5871/jba/008.029>
- Morgan, D. L. (1996). *Focus groups as qualitative research*. Sage.
- Mullen, B., & Smyth, J. M. (2004). Immigrant suicide rates as a function of ethnophobias: Hate speech predicts death. *Psychosomatic Medicine, 66*(3), 343–348. <https://doi.org/10.1097/00006842-200405000-00009>
- Noble, S. U., & Tynes, B. M. (Eds.). (2016). *The intersectional internet: Race, sex, class, and culture online*. Peter Lang.
- Pace, U., D'Urso, G., & Zappulla, C. (2021). Hating among adolescents: Common contributions of cognitive distortions and maladaptive personality traits. *Current Psychology, 40*, 3326–3331. <https://doi.org/10.1007/s12144-019-00278-x>
- Patton, D. U., Brunton, D.-W., Dixon, A., Miller, R. J., Leonard, P., & Hackman, R. (2017). Stop and frisk online: Theorizing everyday racism in digital policing in the use of social media for identification of criminal conduct and associations. *Social Media + Society, 3*(3). <https://doi.org/10.1177/2056305117733344>
- Paz, M., Montero-Diaz, J., & Moreno-Delgado, A. (2020). Hate speech: A systematized review. *SAGE Open, 10*(4). <https://doi.org/10.1177/2158244020973022>
- Potok, M. (2015). The year in hate and extremism. *Intelligence Report*. <https://www.splcenter.org/fighting-hate/intelligence-report/2015/year-hate-and-extremism-0>
- Räsänen, P., Hawdon, J., Holkeri, E., Keipi, T., Näsi, M., & Oksanen, A. (2016). Targets of online hate: Examining determinants of victimization among young Finnish Facebook users. *Violence and Victims, 31*(4), 708–725. <https://doi.org/10.1891/0886-6708.VV-D-14-00079>
- Robinson, B. F., & Bakeman, R. (1998). ComKappa: A Windows' 95 program for calculating kappa and related statistics. *Behavior Research Methods, Instruments, & Computers, 30*, 731–732. <https://doi.org/10.3758/BF03209495>
- Rösner, L., Winter, S., & Krämer, N. C. (2016). Dangerous minds? Effects of uncivil online comments on aggressive cognitions, emotions, and behavior. *Computers in Human Behavior, 58*, 461–470. <https://doi.org/10.1016/j.chb.2016.01.022>
- Shahbaz, A. (2018). *The rise of Digital Authoritarianism: Fake news, data collection, and the challenge to democracy*. <https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism>
- Smith, J. A. (2004). Reflecting on the development of interpretative phenomenological analysis and its contribution to qualitative research in psychology. *Qualitative Research in Psychology, 1*(1), 39–54.
- Smith, J. A., Jarman, M., & Osborn, M. (1999). Doing interpretative phenomenological analysis. In M. Murray, & K. Chamberlain (Eds.), *Qualitative health psychology: Theories and methods* (pp. 218–241). Sage. <https://doi.org/10.4135/9781446217870.n14>

Define and Tackle Hate Speech

- Smith, J. A., & Osborn, M. (2003). Interpretative phenomenological analysis. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods* (pp. 51–80). Sage.
- Smolla, R. A. (1990). Academic freedom, hate speech, and the idea of a university. *Law and Contemporary Problems*, 53(3), 195–205. <https://doi.org/10.2307/1191797>
- Soral, W., Bilewicz, M., & Winiewski, M. (2018). Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior*, 44(2), 136–146. <https://doi.org/10.1002/ab.21737>
- Sorokowski, P., Kowal, M., Zdybek, P., & Oleszkiewicz, A. (2020). Are online haters psychopaths? Psychological predictors of online hating behavior. *Frontiers in Psychology*, 11, 553. <https://doi.org/10.3389/fpsyg.2020.00553>
- Tausch, N., & Becker, J. C. (2013). Emotional reactions to success and failure of collective action as predictors of future action intentions: A longitudinal investigation in the context of student protests in Germany. *British Journal of Social Psychology*, 52, 525–542. <https://doi.org/10.1111/j.2044-8309.2012.02109.x>
- Tommasi, F. (2023). Applied psychology and informing science: Introduction to the developing special series. *Informing Science: The International Journal of an Emerging Transdiscipline*, 26, 1–3. <https://doi.org/10.28945/5087>
- Tsesis, A. (2002). *Destructive messages: How hate speech paves the way for harmful social movements*. New York University Press.
- Weber, M., Viehmann, C., Ziegele, M., & Schemer, C. (2020). Online hate does not stay online – How implicit and explicit attitudes mediate the effect of civil negativity and hate in user comments on prosocial behavior. *Computers in Human Behavior*, 104, 106192. <https://doi.org/10.1016/j.chb.2019.106192>

AUTHORS



Amalia De Leo, Ph.D, is a Postdoctoral Researcher in Work and Organizational Psychology at the Department of Psychology, Catholic University of Milan, Italy. She holds various academic positions, including adjunct professor of Work and Organizational Psychology at the Faculty of Political and Social Sciences of the Catholic University of Brescia. For several years, she has been collaborating as a researcher on national and international research projects at the Alta Scuola di Psicologia Agostino Gemelli Psychology School and at the Centro di Ricerca sullo Sviluppo di Comunità e i Processi di Convivenza on topics related to inter-institutional collaboration, with a particular focus on forced migrations, organizational competences and cultures in a transcultural context, and quality of services.



Emanuele Russo holds a Ph.D. in International Relations and Political Science from the University of Turin. He is the Coordinator of the Global Campaign for Education in Italy and the Chairperson of Amnesty International Italy. He works as expert in Global Citizenship Education at C.I.F.A. ETS, coordinating nation and European wide projects to counter hate speech.