Beyond Counter Speech: An Integrated Literacy Model to Mitigate Online Hate Speech against Roma in Greece

Nikolaos Mouratoglou, PhD (Corresponding author)

School of Philosophy & Education

Aristotle University of Thessaloniki, Greece

Theodora Agapoglou, PhD

School of Philosophy & Education

Aristotle University of Thessaloniki, Greece

Konstantinos Bikos, Professor

School of Philosophy & Education

Aristotle University of Thessaloniki, Greece

Konstantinos Tsioumis, Professor

School of Philosophy & Education

Aristotle University of Thessaloniki, Greece

Abstract

The rapid expansion of digital technologies has transformed modes of communication and social interaction, while also facilitating harmful phenomena such as online hate speech. This paper examines the dynamics, targets, and consequences of online hate speech, with a focus on its impact on the Roma minority in Greece. Drawing on existing literature and prior empirical research, the study identifies recurrent negative constructions of Roma in social media discourse, ranging from stereotypes of criminality and dependency to explicit calls for exclusion and elimination. The analysis highlights the profound psychological, social, and

civic consequences of such rhetoric, including anxiety, depression, self-ghettoisation, polarisation, and the normalisation of intolerance. The paper argues that conventional counter-speech strategies, while necessary, are insufficient to disrupt entrenched discriminatory narratives. Instead, it proposes a multidimensional framework designed to enhance critical engagement, foster empathy, and build digital resilience among both Roma and non-Roma populations. This framework emphasizes collaborative learning environments, culturally responsive pedagogy, and the development of competencies to manage and challenge harmful online discourses, as well as to promote inclusive civic engagement. By addressing both individual and structural dimensions, the study underscores the urgency of multi-level strategies to counteract the glocal hostility—where globalized hate speech tropes intersect with local socio-cultural prejudices.

Keywords: digital literacy, digital well-being, antigypsyism, psychoeducational interventions, digital citizenship.

1. Introduction

The widespread adoption of the internet, along with the diverse digital technologies, has fundamentally transformed human society and human interactions. The unprecedented access to and management of information, as well as the various modes (online, blended) of social interaction are concrete examples of powerful dynamics which transcend traditional limitations associated with geographical and temporal boundaries. The emergence of the internet in the latter part of the 20th century and the subsequent proliferation of digital devices, such as personal computers and smartphones, have played a pivotal role in this transformation (Castells, 2009). In addition, the seamless integration of digital technologies into everyday life has impacted the way individuals communicate, access information, and form social connections (DiMaggio, Hargittai, Neuman, & Robinson, 2001); technological advancements such as the development of smartphones, high-speed internet connectivity, and social media platforms may be regarded as crucial accelerators of this integration. Specifically, social media platforms like Facebook, X (formerly Twitter), and Instagram have revolutionised how people share information and connect with others (Kaplan & Haenlein, 2010). Mobile devices have further enhanced this connectivity, enabling users to access the internet and communicate from virtually anywhere, whereas high-speed internet has made it possible to transmit large amounts of data quickly, facilitating real-time interactions and the consumption of digital content (Rainie & Wellman, 2012).

The digital era has profoundly altered traditional modes of social interaction, leading to the proliferation of online communities and virtual relationships. Social media and other digital platforms have enabled people to maintain relationships and form new connections without the constraints of physical proximity (Hampton & Wellman, 2018). Such changes have given rise to new forms of social organisation and community building, as well as challenges related to the nature and quality of these interactions (Turkle, 2011). For instance, digital technologies have revolutionised the communication practices, facilitating real-time messaging, video conferencing, and social media interactions. These advancements have also made it possible for people to communicate in real-time across vast distances, significantly

enhancing the speed and efficiency of information exchange (Boyd & Ellison, 2007). Social media platforms, in particular, have played a crucial role in shaping interpersonal relationships and the dissemination of information. They provide a space for individuals to share personal updates, engage in public discourse, and mobilise for social causes, thereby influencing public opinion and social behavior. As Ellison, Vitak, Gray, and Lampe (2014) highlight social media may contribute to individuals' social and information resources.

When these individuals find others with shared interests, values or causes, they tend to create online communities (De Souza & Preece, 2004; Preece, 2004), an equivalent of a social group/collective in digital contexts. These communities may range from hobbyist forums to political discussion groups and play a significant role in social cohesion and identity formation (Wellman & Gulia, 1999). Online communities provide a sense of belonging and support for their members, fostering solidarity and collective action (Rheingold, 2000). The diverse nature of these communities highlights the varied ways in which individuals can connect and engage with others, forming networks that transcend traditional social structures. It might be further commented that online interactions might even transcend traditional social values, norms, and practices. For instance, this is particularly evident in online multiplayer gaming communities, in which "*constant negotiation of the values, norms and rules of the community contributes to the emergence of a shared understanding of the symbolic reality of the community and a shared social identity*" (Siitonen, 2007: 6). In addition, people may behave differently online compared to how they would in face-to-face interactions. This behavior, known as the 'online disinhibition effect' (Suler, 2004) often involves reduced inhibitions and increased willingness to express emotions, disclose personal information, or engage in behaviors that they might refrain from in offline settings, due to factors such as anonymity, invisibility, the asynchronous nature of communication, and the perceived distance from consequences.

In this context, several issues might arise including online privacy, cyberbullying, and the digital divide (Livingstone, 2008). Considering social media interactions, online privacy is threatened by the unauthorised use of personal information (Van Schaik et al., 2018), while cyberbullying can have severe psychological effects on young people (Kowalski et al., 2014) and minority groups (Llorent et al., 2016). Concerning the digital divide, it may negatively impact aspects of employment, education, health, and overall socio-economic development, further exacerbating social inequalities and exclusion (Molala & Makhubele, 2021). Building upon these issues, in conjunction with the ever-expanding use of the internet as a means of disseminating hate and propaganda (European Agency of Fundamental Rights, 2013), online hate speech emerges as a significant concern. Despite the several definitions of hate speech (see Perera et al., 2023), online hate speech is associated with distinctive features such as anonymity, invisibility, and instantaneousness (Brown, 2017), which may pose additional risks to the well-being of the targeted individuals or groups both in online and offline environments.

2. Targets of online hate speech: The case of the Roma minority in Greece

Online hate speech is characterised by derogatory language, stereotyping, and incitements to

violence and extremism. This form of digital aggression not only perpetuates harmful stereotypes but also reinforces discrimination, marginalisation and social exclusion. The anonymity afforded by the internet emboldens individuals to express hateful sentiments without fear of repercussion, thus amplifying the spread of such rhetoric (KhosraviNik & Esposito, 2018).

Online hate speech affects a wide range of individuals and groups, often targeting those who are already marginalised or vulnerable in society (for an overview see Silva et al., 2016). Ethnic and racial minorities are frequent victims, with groups such as African Americans, Jews, Muslims, and Asians experiencing significant levels of online hostility (Bliuc et al., 2018). Gender-based hate speech is also pervasive, particularly against women, who face threats, harassment, and derogatory comments that aim to undermine them (Jane, 2017; KhosraviNik & Esposito, 2018; Sjöholm, 2024). LGBTQ+ individuals are another major target (Brody et al., 2023), as well as refugees and migrants (Meza et al., 2019) and people with disabilities (Burch, 2017), who often encounter discriminatory rhetoric that contributes to their stigmatisation. Moreover, anti-hate activists who advocate for social justice may also become targets of hate speech (Gagliardone et al., 2015). Similarly, Obermaier and Schmuck (2022) observed six distinct groups of young people who encounter varying degrees of online hate speech targeting because of their gender, migration background, religion, or advocacy for LGBTQ+ rights. These categories also took into account how individuals perceive their own level of disadvantage compared to others in society (subjective deprivation), the levels of political engagement, and varying degrees of digital media literacy.

Concerning Greece, Voda (2023) reports that hate speech is directed at a variety of groups and communities, including migrants and refugees, the LGBTQ+ community, religious minorities, people with disabilities, women and Roma. The Roma community is an additional group that has been significantly targeted (see Agapoglou et al, 2021; Assimakopoulos et al, 2017; Breazu & Machin, 2022; Molnar, 2021; Townsend, 2014). Earlier work of the authors focused on inquiring the social representations for Roma people as illustrated in comments posted across dominant and less dominant social media in Greece (Mouratoglou et al., 2022a; Mouratoglou et al., 2022b). A key contribution of this research was the fact that it constituted the first attempt to portray the negative constructions that have been associated with the Roma community in Greece in online hate speech manifested in social media. Namely, based on qualitative analysis of public comments posted on Facebook and YouTube, 14 distinct categories emerged. A brief overview is included in Table 1.

Table 1. Negative constructions associated with Roma in Greek social media.

Category[1]	Description
1. Roma as a sanitary time bomb	Perceived as a health hazard due to poor hygiene practices. Comments often describe them as dirty and a threat to public health.
2. Roma as parasites	Perceived as parasites who exploit others and are often accused of not paying taxes and consequently not contributing to society.
3. Roma as sub-humans	Perceived as less than human, often comparing them to animals (mice, apes, worms and cockroaches) that need to be eradicated.
4. Roma as non-Greeks	Perceived as people from Türkiye, Mongolia, India, Bulgaria, Albania, Pakistan and Syria, countries that are perceived as inferior in the Greek mentality.
5. Roma as Greek language 'murderers'	Perceived as non-competent in using the Greek language, hence providing further justification for the previous category (Roma as non-Greeks).
6. Roma as ineducable group	Perceived as uninterested in education and thus uneducable yet seen as keen on taking advantage of educational benefits without a sincere desire to learn.
7. Roma as rejectors of social inclusion	Perceived as unwilling to integrate into society, preferring to remain isolated and reliant on social benefits, which fosters a perception of self-ghettoization.
8. Roma as recipients of high social benefits	Perceived as exploiting the welfare system, receiving excessive social benefits that discourage them from seeking employment and integrating into society.
9. Roma as a politically manipulated group	Perceived as being manipulated by politicians for votes, receiving benefits in exchange for political support, thus affecting the democratic process.
10. Roma as delinquents	Perceived as criminals involved in theft, drug dealing, and other illegal activities. They are often described as inherently

	destructive and dishonest.
11. Roma as a group to be subjected to birth-control	Perceived as a group with high birth rates, which should be reduced via sterilisation and other population control measures.
12. Roma as irresponsible parents	Perceived as parents who exploit their children for financial gain, neglecting their basic needs and well-being. Early marriages and incest are also cited.
13. Roma as victims of far-right ideology	Perceived as a group associated with ethnic cleansing, Auschwitz, crematoria, trains and Eugenics.
14. Roma as targets of mass-murder	Perceived as a group which needs to die. The difference with the previous category is the explicit references and requests for complete elimination.

3. Depicting the impact of online hate speech

Exposure to online hate speech is associated with diverse forms of psychological distress, including anxiety, depression, and trauma. Research conducted by Tynes et al. (2010) concluded that online racial discrimination significantly predicts depressive symptoms and anxiety in adolescents. Similarly, Schmitt et al. (2014) demonstrated that repeated exposure to hate speech leads to increased stress and decreased psychological well-being, namely self-esteem, depression, anxiety, psychological distress and life satisfaction.

Dreißigacker et al. (2024) confirm that victims of online hate speech experience heightened feelings of insecurity in offline environments. In the case of LGBTQ+ youth, Keighley (2021) found that online hate significantly impacts their emotional and behavioral responses including sadness, shame, and feelings of inferiority which in turn lead these young individuals to internalise the blame for their victimisation. In fact, Wachs et al (2022) have concluded that online hate speech victimisation is positively correlated with depressive symptoms, whereas Mouratoglou et al. (2022a) underline that Roma may be at higher risk of self-ghettoisation because they internalise feelings of inferiority imposed by the privileged non-Roma. Moreover, Cohen-Almagor (2022) highlighted the connections between hate speech and cyberbullying, citing several tragic cases where young individuals committed suicide as a result of such bullying. Regarding Roma, as Greenfields and Rogers (2020) highlighted that establishing a link between individuals' experiences of hate crime with their mental health, wellbeing, suicide and para-suicide is challenging considering the lack of objective recorded data. Nevertheless, the qualitative part of their study, which focused on the psychological effects of hate crime on Gypsy, Traveller and Roma (GTR) communities, revealed that the impact of hate speech falls under four main categories: a) mental health (depression/anxiety), b) withdrawal from, anger and suspicion of 'mainstream' society, c)

learning to 'live with it' and d) 'self-medication'/harmful behaviours/(emphasis on masculinities). Concerning Greece, the research conducted by Voda (2023) identifies patterns and presents case studies which indicate a significant link between hate speech and hate crimes.

Hate speech can also negatively impact individuals' sense of identity, heightening feelings of isolation, especially among minority groups such as immigrants (Wypych & Bilewicz, 2024). In this context, online hate speech can erode social cohesion by fostering an environment of hostility and division. For instance, Negrea-Busuioc et al. (2023) underline that the level of civic engagement (e.g., signing a petition or joining a protest to support Roma people's rights) is significantly impacted by online hate speech. According to Bilewicz and Soral (2020) exposure to hate speech leads to empathy being substituted by intergroup contempt, as well as to desensitisation, diminishing individuals' capacity to perceive the offensive nature of such language. Similarly, Dal Santo and D'Angelo (2022) highlight that radicalisation might result in individuals perceiving violence or discrimination against certain groups as normal. Therefore, the prevalent presence of online hate speech can lead to its normalisation, making such discourse seem acceptable or typical, especially when accompanied by quasi-humorous comments (Chovanec, 2021) or legitimised by politicians (Miškolci et al., 2018; Thompson & Woodger, 2020). In turn, this normalisation may desensitise individuals to aggression and violence, as shown in the research conducted by Govers et al. (2023); namely the researchers concluded that exposure to online hate speech can lead to radicalisation, where individuals adopt more extreme viewpoints and are more likely to engage in offline violent actions. Thus, online hate speech can also intensify group polarisation, leading to more extreme attitudes and behaviors within and among groups. Indicatively, Popenko (2022) stresses that Roma individuals in Ukraine face mental and physical violence, along with discrimination, as a result of online hate speech in social media. She also highlights that a significant cause of hate-motivated attacks and violations of their rights (including liberty, security, protection against discrimination, and privacy) is the authorities' failure to effectively regulate and control hostility.

4. Mitigating (the impact of) online hate speech

Considering this multifaceted impact, delimiting online hate speech requires comprehensive strategies that encompass legal measures, educational initiatives, and increased (critical) digital literacy to foster a more inclusive and respectful online environment for all marginalised groups.

One of the proposed strategies to delimit hate speech refers to effective regulation and policy, critical elements in mitigating (the impact of) online hate speech. Nevertheless, according to the European Agency of Fundamental Rights (2023), although most EU Member States have developed action plans and started implementing their national strategies (Note 1) to promote equality, inclusion and participation, there has been little or no progress in fighting antigypsyism, and ensuring equal access to education, employment, housing and health. Probably the main obstacle refers to the ongoing and complex challenge of balancing regulation with free speech rights, considering that "*policy makers need to take account of*

how power is exercised within a society between majority and minority social groups" (Bromell, 2022: 185).

Social media platforms partially act as digital/internet intermediaries (Note 2), which according to Citron and Norton (2011: 1440-1441) "*can foster digital citizenship by inculcating norms of respectful, vigorous engagement*". In this context, five years later the European Commission launched the EU Code of Conduct with the ultimate aim of preventing and countering the spread of illegal online hate speech (European Commission, 2016). Renowned social media platforms such as Facebook, Microsoft, Twitter and YouTube, and later on Instagram, Snapchat TikTok, LinkedIn and Rakuten Viber and Twitch joined the Code of Conduct. Since then, the implementation of the Code of Conduct has been monitored; according to the seventh monitoring there is a slower review of the reported incidents (as compared to 2021 and 2020), while the overall removal rate remains similar to 2021, but still lower than 2020 (European Commission, 2022). In addition, more advanced developments in natural language processing and text mining strategies (Fortuna & Nunes 2018), machine learning (Abro et al., 2020), artificial intelligence (Mehta & Passi, 2022) and deep learning (Aldreabi & Blackburn, 2023) have enabled the automated detection of hate speech. However, despite efforts to utilise such techniques for detection and monitoring, their performance remains unsatisfactory (Jahan & Oussalah, 2023). Such automations have also been implemented in counter speech, using automated bots to diminish instances of racist speech (Munger 2017; Tekiroğlu et al., 2020; Tekiroğlu et al., 2022).

According to Benesch (2014) counter speech (derived from the concept of counter narratives; see Solórzano & Yosso, 2002) describes the act of responding to hateful comments with the intention of influencing both the discourse and behavior of the person publishing/posting the comments. In 2013, the Council of Europe launched, and has since then been implementing, the 'No Hate Speech' (Note 3) youth campaign across 45 countries with the aim of empowering youth to combat online hate speech and promote human rights; similarly UNESCO published a study with the aim of raising awareness and discuss initiatives for countering online hate speech through media and information literacy (Gagliardone et al., 2015). Regarding Roma, non-violent counter narratives in conjunction with de Bono's six thinking hats have been used as a method-strategy for mitigating the effect of online hate speech during an emancipatory action research with Roma youth (Agapoglou et al, 2021). According to Mathew et al., (2018), counter speech is considered an effective method for addressing online hate speech while preserving freedom of speech. The research conducted by Garland et al. (2022) provides some initial insights in this regard; namely, having analysed 131,366 political discussions on Twitter, the researchers suggest that organised hate speech influences public discourse and in turn organised counter speech can be effective in reducing hateful rhetoric. Nevertheless, Schieb and Preuss (2016) provide an alternative perspective, stating that counter speech is delegated to dedicated users which in turn results in internet companies not developing new technologies or investing in manual intervention.

Well-informed and evidence-based interventions may also play vital roles in combating online hate speech. In Wachs et al.'s (2022) study, individuals targeted by online hate speech showed lower levels of depression when they demonstrated moderate to high levels of

resilience; resilience was conceptualised based on factors such as personal and social competence, structured coping methods, access to social support, and strong family bonds, in contrast to those with lower resilience levels. Similarly, Kaakinen (2018) suggests that offline social capital is linked to a reduced risk of becoming either a perpetrator or victim of online hate offenses. This implies that individuals who have robust and supportive offline social networks are less likely to engage in or be targeted by online hate speech. Moreover, strong connections to offline social networks mitigate the harmful consequences of victimisation resulting from online hate incidents. Therefore, strengthening resilience and offline social capital could be a crucial strategy in combating online hate speech.

Psychoeducational programmes aimed at promoting digital literacy and empathy may also empower individuals to recognise and counter hate speech. Yet, Bilewicz and Soral (2020) argue that while empathy and social norms could serve as effective barriers to the spread of hate speech, the presence of hate speech actually undermines these mechanisms, reducing their effectiveness. Costello et al. (2016) further support that norms are weaker online, considering that appropriateness, acceptability, and civility are significantly more fragile in online environments. Similarly, Álvarez-Benjumea and Winter (2018) concluded that adherence to norms in online conversations may actually be driven more by descriptive norms (e.g., perception about how others typically behave in a given context and being influenced by observing the common behaviors and actions of their peers) than by injunctive norms (e.g., perceptions of what behaviors are socially approved or disapproved of, based on societal or group expectations).

Moreover, dual-target programmes 'with and for the Roma' may support individuals to develop their identity and interact with the global community (Banks, 2012), as well as practices like the living libraries which might shift the asymmetrical intercultural interactions into symmetrical (Agapoglou & Kesidou, 2020). Moreover, Mouratoglou et al. (2022b) suggest that addressing key issues such as hate speech, human rights and citizenship education may contribute to mitigating online hate speech, while Voda (2023) emphasises the focus on the principle of human dignity and Cohen-Almagor (2022) on moral and social responsibility. In addition, critical digital literacy (Hinrichsen & Coombs, 2014) has been shown to benefit Roma youth in counteracting online hate speech narratives (Agapoglou et al, 2021). Information and media literacy can also offer insights and benefit the recipients of hate speech (Gagliardone et al., 2015), by focusing on critical thinking, netiquette, dis/mal/misinformation and propaganda (Mouratoglou et al., 2022b). A closely related (and broader) approach is reflected with the term of 'transliteracy' (Frau-Meigs, 2012; Thomas et al., 2007) which encompasses aspects of literacy (reading/writing), media literacy (platforms, media, tools) and information literacy (critical evaluation of information etc).

5. Discussion

The expanding use of the internet and digital technologies, coupled with technological advancements such as the internet of things, smart devices and artificial intelligence have (trans-)formed sociotechnological practices such as global connectivity, e-health, teleworking, e-learning, influencing individuals' behaviors and relationships, as well as the organisational

and societal structures. By transcending geographical and temporal limitations, these technologies have facilitated new modes of social interaction. In this context, social media platforms have allowed individuals to connect with each other, resulting in a sense of social engagement and belonging; yet they also led to a blurring of the boundaries between the physical and digital worlds. Inherent features such as the (perceived) anonymity, invisibility and instantaneousness (Brown, 2017) have impacted interpersonal interactions in a way which they now may not be subjected to traditional social structures, values, norms and practices. A prominent example of this is online hate speech which may pose significant risks to the well-being of targeted individuals or groups both online and offline. The study illustrates the pervasive and deeply ingrained nature of online hate speech and its substantial impact on marginalised groups, with a special focus on the Roma community in Greece.

Online hate speech may be viewed as a double-edged sword that leads to fragmentation at both individual and social levels, causing psychological distress, eroding trust within communities, and fostering a sense of isolation among targeted individuals. In terms of the psychological effects of online hate speech on targeted individuals, these entail heightened levels of anxiety, depression, and a diminished sense of self-worth (Schmitt et al., 2014; Tynes et al., 2010). The internalisation of these negative portrayals can lead to self-ghettoisation, where individuals withdraw from broader societal engagement, further isolating themselves and exacerbating their marginalisation. This withdrawal not only impacts the individuals' mental health but also diminishes their opportunities for social and economic participation, perpetuating a vicious cycle of reproducing stereotypes, discrimination, inequality, racism and injustice. In parallel, online hate speech contributes to fragmentation and polarisation within communities; it can strengthen the cohesion among like-minded racist community members based on the echo chamber effect (Criss et al., 2020) where individuals are exposed primarily to views that reinforce their own prejudices, deepens divisions and fosters hostility. This dual effect not only deepens social and digital divides but also entrenches hostile attitudes, making it harder to bridge gaps and foster understanding and acceptance.

Online hate speech is rooted in the ideological code of 'Us and Them' reflecting a perceived superiority-inferiority dichotomy that feeds into public discourse (Mouratoglou et al., 2022). As a result, it significantly affects an individual's life both online and offline, influencing their relationships with themselves and with others (De Ridder & Van Bauwel, 2015). In either scenario, the impact of online hate speech weakens the social fabric by exacerbating divisions and fostering an environment of intolerance and hostility. This, in turn, undermines collective efforts to address structural and institutionalised discrimination, given that the social fragmentation and increased polarisation hinder the development of inclusive policies and collaborative actions necessary for tackling systemic issues. Consequently, the potential for meaningful dialogue and cooperative problem-solving diminishes, leaving society ill-equipped to combat entrenched inequalities and injustices.

The study also highlights how hate speech, often legitimised by political figures or masked as humor, can lead to its normalisation. This process desensitises individuals, fostering radicalisation and more extreme attitudes and behaviors, including offline violence. The lack

of effective regulation and control by authorities over online hostility is a significant factor in the perpetuation and escalation of hate-motivated attacks. This regulatory failure results in ongoing violations of rights and increased incidences of violence. In this regard, the study underlines that policymakers and internet intermediaries may play a pivotal role in both the proliferation and mitigation of online hate speech, ensuring a respectful and safe environment. The European Commission's Code of Conduct, which several social media platforms have joined, is a step towards this goal. However, the effectiveness of such initiatives may potentially be strengthened with advanced technologies, such as artificial intelligence and machine learning for automatically detecting, removing and countering hate speech.

Moreover, the current review study highlights the role of evidence-based psychoeducational interventions to combat online hate speech. Empowering individuals through digital literacy and critical thinking skills can enable them to recognise and counteract hate speech effectively. In addition, empirical studies show that enhancing resilience (Wachs et al., 2022) and offline social capital (Kaakinen, 2018) can mitigate the adverse effects of hate speech. Based on the literature review, we propose the following elements that should be considered in such interventions by adopting an integrated approach of an expanded 'transliteracy' (Frau-Meigs, 2012; Thomas et al., 2007), coupled with necessary life skills which may function as a buffer against the psychological impact of online hate speech.

Table 2. Elements for psychoeducational interventions to proactively mitigate (the impact of) online hate speech.

Literacies	Skills addressed	Indicative content
Civic literacy	Critical thinking, empathy, resilience, assertiveness	Politics and policies, power dynamics, citizenship education, social responsibility and justice, democratic processes, human rights education, advocacy and activism.
Social literacy	Empathy, communication, self-awareness, conflict resolution, networking	Social norms, group dynamics, interpersonal relationships, respectful communication, cultural sensitivity, mediation, negotiation and conflict resolution, social support networks.
Intercultural literacy	Empathy, cultural sensitivity, identity awareness	Multiculturality, pluralism, intercultural communication, challenging stereotypes and biases, multiperspectivity.
Emotional literacy	Empathy, resilience, self-regulation, self-efficacy	Emotions, self-image, values, wellbeing, self-awareness, coping, self-care

		techniques,
Health literacy	Resilience, coping, stress management, online safety	Wellbeing (physical, digital, emotional, psychological, social), impact of online behavior on wellbeing and mental health, seeking help/support, stress management, growth mindset.
Legal literacy	Critical thinking, online safety, social responsibility	Legal rights, illegal behavior, reporting mechanisms, seeking legal support.
Ethical literacy	Social responsibility, critical thinking, empathy, integrity	Moral/ethical values, inclusivity, ethical dilemmas and decision making, understanding consequences of actions.
Technological literacy	Digital skills, problem-solving, online safety	Understanding technological tools, solving common tech issues, privacy, safety.
Media literacy	Empathy, critical thinking, social responsibility	Role of (social) media, media message analysis, bias identification, creating responsible content.
Information literacy	Critical thinking, resilience	Effective searching strategies, evaluating information credibility, ethical use of information.
Digital literacy	Social responsibility, online safety, digital	Netiquette, online safety, digital life, online rights and responsibilities, reporting mechanisms.
Critical digital literacy	Critical thinking, skepticism, digital skills, resilience	Evaluation of sources, identification of ideological interests and misinformation, understanding privacy settings, and managing digital footprints, online identity formation
AI literacy	Critical thinking, digital skills	Awareness about artificial intelligence and algorithms (e.g., personalisation), detection of AI in daily life (e.g., use of bots), using AI tools, ethical considerations.
Futures	Critical thinking, empathy, identity awareness, optimism,	Plurality, governance, power dynamics, ideologies, sociotechnical imaginaries,

literacy	growth mindset	meaning making, social change.

Considering the historical contextualisation—including social, political, and economic factors—the diverse targets, as well as the various forms and "layers" of enactment of online hate speech, it becomes evident how complex it is to fully understand and mitigate its effects. Online hate speech manifests in multiple forms, including direct threats, derogatory language, misinformation, and subtle microaggressions. These forms can operate on different layers, from individual interactions to coordinated campaigns by organised groups. At the same time, the evolution of societal norms, as well as of the political and economic conditions over time, is shaping the nature (and prevalence) of hate speech. In this context, identifying patterns of discrimination and marginalisation play a significant role in the current dynamics of online hate speech, as longstanding prejudices are often amplified in digital spaces. Therefore, an important consideration of any psychoeducational intervention would need to initiate from exploring the multifaceted, situated, nature of online hate speech and target the appropriate content and skills which will effectively equip, empower and safeguard individuals and marginalised groups.

Funding

None.

Informed Consent

Obtained.

Provenance and Peer Review

Not commissioned; externally double-blind peer reviewed.

Data Availability Statement

The data that support the findings of this study are available on request.

Competing Interests Statement

The authors declare that there are no competing or potential conflicts of interest.

References

Abro, S., Shaikh, S., Hussain, Z., Ali, Z., Khan, S., & Mujtaba, G. (2020). Automatic Hate Speech Detection using Machine Learning: A Comparative Study. *International Journal of Advanced Computer Science and Applications/International Journal of Advanced Computer Science & Applications*, *11*(8). https://doi.org/10.14569/ijacsa.2020.0110861

Agapoglou, T., & Kesidou, A. (2020). Identity Texts and Educational Drama as Pedagogical Tools for the Social Empowerment of Adolescents from Oppressed Minority Groups: Reflections from a Critical-Emancipatory Action Research. In A. Chatzidaki, P. Thomou, D. Kontogianni, Th. Michelakaki, & M. Argyroudi (Eds.), *Selected Papers from the 5th International Conference "Crossroads of Languages and Cultures: Languages and cultures at home and at school"* (186–195). Rethymno: EDIAMME. Retrieved June 15, 2024, from

https://clc5.ediamme.edc.uoc.gr/wp-content/uploads/2020/04/Selected-Papers_CLC5new.pdf

Agapoglou, T., Mouratoglou, N., Tsioumis, K. & Bikos, K. (2021). Combating online hate speech through critical digital literacy: Reflections from an emancipatory action research with Roma youths. *International Journal of Learning and Development, 11(2).* https://doi.org/10.5296/ijld.v11i2.18524

Aldreabi, E. &Blackburn, J. (2023). Enhancing Automated Hate Speech Detection: Addressing Islamophobia and Freedom of Speech in Online Discussions. *ASONAM '23: Proceedings of the 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 644–651. https://doi.org/10.1145/3625007.3627487

Álvarez-Benjumea, A., & Winter, F. (2018). Normative Change and Culture of Hate: An experiment in online environments. *European Sociological Review*, *34*(3), 223–237. https://doi.org/10.1093/esr/jcy005

Assimakopoulos, S. Baider, F. H. & Millar, S. (2017). *Online Hate Speech in the European Union: A Discourse-Analytic Perspective*. Springer Nature. http://library.oapen.org/handle/20.500.12657/27755

Benesch, S. (2014). *Countering Dangerous Speech to Prevent Mass Violence during Kenya's 2013 Elections*. Technical report, United States Institute of Peace. Retrieved June 15, 2024, from: https://www.ushmm.org/m/pdfs/20140212-benesch-kenya.pdf

Bilewicz, M., & Soral, W. (2020). Hate speech epidemic. The dynamic effects of derogatory language on intergroup relations and political radicalization. *Political Psychology*, *41*(S1), 3–33. https://doi.org/10.1111/pops.12670

Bliuc, A., Faulkner, N., Jakubowicz, A., & McGarty, C. (2018). Online networks of racial hate: A systematic review of 10 years of research on cyber-racism. *Computers in Human Behavior*, *87*, 75–86. https://doi.org/10.1016/j.chb.2018.05.026

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, 13*(1), 210-230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Breazu, P., & Machin, D. (2022). Racism is not just hate speech: Ethnonationalist victimhood in YouTube comments about the Roma during Covid-19. *Language in Society*, *52*(3), 511–531. https://doi.org/10.1017/s0047404522000070

Brody, E., Greenhalgh, S. P., & Sajjad, M. (2023). Free speech or free to hate? Anti-LGBTQ+ Discourses in LGBTQ+-Affirming Spaces on Gab Social. *Journal of Homosexuality*, 1–26. https://doi.org/10.1080/00918369.2023.2218959

Bromell, D. (2022). Striking a fair balance when regulating free speech. In *Springer eBooks* (pp. 167–190). https://doi.org/10.1007/978-3-030-95550-2_8

Brown, A. (2017). What is so special about online (as compared to offline) hate speech? *Ethnicities*, *18*(3), 297–326. https://doi.org/10.1177/1468796817709846

Burch, L. (2017). 'You are a parasite on the productive classes': online disablist hate speech in austere times. *Disability & Society*, *33*(3), 392–415. https://doi.org/10.1080/09687599.2017.1411250

Castells, M. (2009). *The rise of the network society* (2nd ed.). Wiley-Blackwell.

Chovanec, J. (2021). 'Re-educating the Roma? You must be joking. . .': Racism and prejudice in online discussion forums. *Discourse & Society*, *32*(2), 156–174. https://doi.org/10.1177/0957926520970384

Citron, D. K., & Norton, H. (2011). Intermediaries and hate speech: Fostering digital citizenship for our information age. *Boston University Law Review, 91*(4), 1435-1484. https://digitalcommons.law.umaryland.edu/fac_pubs/1056/

Cohen-Almagor, R. (2022). Bullying, cyberbullying, and hate speech. *International Journal of Technoethics*, *13*(1), 1–17. https://doi.org/10.4018/ijt.291552

Costello, M., Hawdon, J., & Ratliff, T. N. (2016). Confronting online extremism. *Social Science Computer Review*, *35*(5), 587–605. https://doi.org/10.1177/0894439316666272

Council of Europe (2018). *Recommendation CM/Rec (2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries*. Retrieved June 15, 2024, from: https://search.coe.int/cm?i=0900001680790e14

Criss, S., Michaels, E. K., Solomon, K., Allen, A. M., & Nguyen, T. T. (2020). Twitter Fingers and Echo Chambers: Exploring expressions and experiences of online racism using Twitter. *Journal of Racial and Ethnic Health Disparities*, *8*(5), 1322–1331. https://doi.org/10.1007/s40615-020-00894-5

Dal Santo, E., & D'Angelo, E. (2022). Relationship of online hate, radicalization, and terrorism. In E. W. Dunbar (Ed.), *Indoctrination to Hate: Recruitment Techniques of Hate Groups and How to Stop Them*. ABC-CLIO.

De Ridder, S., & Van Bauwel, S. (2015). The discursive construction of gay teenagers in times of mediatization: youth's reflections on intimate storytelling, queer shame and realness in popular social media places. *Journal of Youth Studies*, *18*(6), 777–793. https://doi.org/https://doi.org/10.1080/13676261.2014.992306

De Souza, C. S., & Preece, J. (2004). A framework for analyzing and understanding online communities. *Interacting With Computers*, *16*(3), 579–610. https://doi.org/10.1016/j.intcom.2003.12.006

DiMaggio, P., Hargittai, E., Neuman, W. R., & Robinson, J. P. (2001). Social implications of the internet. *Annual Review of Sociology, 27*, 307-336. https://doi.org/10.1146/annurev.soc.27.1.307

Dreißigacker, A., Müller, P., Isenhardt, A., & Schemmel, J. (2024). Online hate speech victimization: consequences for victims' feelings of insecurity. *Crime Science*, *13*(1). https://doi.org/10.1186/s40163-024-00204-y

Ellison, N. B., Vitak, J., Gray, R., & Lampe, C. (2014). Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes. *Journal of Computer-Mediated Communication, 19*(4), 855-870. https://doi.org/10.1111/jcc4.12078

European Agency of Fundamental Rights (2013). *Opinion of the European Union agency for fundamental rights on the framework decision on racism and xenophobia - with special attention to the rights of victims of crime*. Retrieved June 15, 2024, from: https://fra.europa.eu/sites/default/files/fra-opinion-2-2013-framework-decision-racism-xenop hobia_en.pdf

European Commission (2016). *Code of conduct on countering illegal hate speech online*. Retrieved June 15, 2024, from: https://ec.europa.eu/info/policies/justice-and-fundamentalrights/combatting-discrimination/ra cism-and-xenophobia/eu-code-conduct-counteringillegal-hate-speech-online_en

European Commission (2022). *Countering illegal hate speech online. 7th evaluation of the Code of Conduct*. Retrieved June 15, 2024, from: https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/co mbatting-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-spe ech-online_en

European Commission (2022). Countering illegal hate speech online . Factsheet - 7th monitoring round of the Code of Conduct. Retrieved June 15, 2024, from: https://commission.europa.eu/document/download/5dcc2a40-785d-43f0-b806-f065386395de _en?filename=Factsheet%20-%207th%20monitoring%20round%20of%20the%20Code%20o f%20Conduct.pdf

European Union Agency for Fundamental Rights (2023). *Fundamental rights report – 2023*. Luxembourg: Publications Office of the European Union. https://data.europa.eu/doi/10.2811/45511

Fortuna, P., & Nunes, S. (2018). A Survey on Automatic Detection of Hate Speech in text. *ACM Computing Surveys, 51*(4), 1–30. https://doi.org/10.1145/3232676

Frau-Meigs, D. (2012). Transliteracy as the New Research Horizon for Media and Information Literacy. *Medijske Studije Media Studies,* 3(6). 14-27. https://hrcak.srce.hr/ojs/index.php/medijske-studije/article/view/6064

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering Online Hate Speech*. UNESCO Publishing. Retrieved June 15, 2024, from: https://unesdoc.unesco.org/ark:/48223/pf0000233231

Garland, J., Ghazi-Zahedi, K., Young, J., Hébert-Dufresne, L., & Galesic, M. (2022). Impact and dynamics of hate and counter speech online. *EPJ Data Science*, *11*(1). https://doi.org/10.1140/epjds/s13688-021-00314-6

Govers, J., Feldman, P., Dant, A., & Patros, P. (2023). Down the rabbit hole: detecting online

extremism, radicalisation, and politicised hate speech. *ACM Computing Surveys*, *55*(14s), 1–35. https://doi.org/10.1145/3583067

Greenfields, M. & Rogers, C. (2020). *Hate: "As regular as rain" A pilot research project into the psychological effects of hate crime on Gypsy, Traveller and Roma (GTR) communities.* Retrieved June 15, 2024, from: https://bnu.repository.guildhe.ac.uk/id/eprint/18142/1/18142_Greenfields_M%20Rogers_C.pdf

Hampton, K. N., & Wellman, B. (2018). Lost and saved... again: The moral panic about the loss of community takes hold of social media. *Contemporary Sociology, 47*(6), 643-651. https://doi.org/10.1177/0094306118805414

Hellenic Ministry of Labour and Social Affairs (2023). *National strategy and action plan for the social inclusion of Roma 2021 – 2030 1st update*. Retrieved June 15, 2024, from: https://egroma.gov.gr/wp-content/uploads/2023/06/EL_%CE%95%CE%A3%CE%9A%CE%95-%CE%A1%CE%BF%CE%BC%CE%B1-2021_2030_FINAL_1st-REVISION_2023-05.04.2023.pdf

Jahan, M. S., & Oussalah, M. (2023). A systematic review of hate speech automatic detection using natural language processing. *Neurocomputing*, *546*, 126232. https://doi.org/10.1016/j.neucom.2023.126232

Jane, E. A. (2017). *Misogyny Online: A Short (and Brutish) History*. SAGE Publications.

Kaakinen, M. A. I. (2018). *Disconnected Online: A social psychological examination of online hate*. [Doctoral Thesis, University of Tampere]. Tampereen yliopistopaino. http://urn.fi/URN:ISBN:978-952-03-0767-7

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59-68. https://doi.org/10.1016/j.bushor.2009.09.003

Keighley, R. (2021). Hate Hurts: Exploring the impact of online hate on LGBTQ+ young people. *Women & Criminal Justice*, *32*(1–2), 29–48. https://doi.org/10.1080/08974454.2021.1988034

KhosraviNik, M., & Esposito, E. (2018). Online hate, digital discourse and critique: Exploring digitally-mediated discursive practices of gender-based hostility. *Lodz Papers in Pragmatics*, *14*(1), 45–68. https://doi.org/10.1515/lpp-2018-0003

Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin, 140*(4), 1073-1137. https://doi.org/10.1037/a0035618

Livingstone, S. (2008). Taking risky opportunities in youthful content creation: Teenagers' use of social networking sites for intimacy, privacy, and self-expression. *New Media & Society, 10*(3), 393-411. https://doi.org/10.1177/1461444808089415

Llorent, V. J., Ortega-Ruiz, R., & Zych, I. (2016). Bullying and Cyberbullying in Minorities: Are They More Vulnerable than the Majority Group? *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01507

Mathew, B., Saha, P., Tharad, H., Rajgaria, S., Singhania, P., Maity, S. K., Goyal, P., & Mukherje, A. (2018). Thou shalt not hate: Countering Online Hate Speech. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1808.04409

Mehta, H., & Passi, K. (2022). Social media hate speech detection using Explainable Artificial Intelligence (XAI). *Algorithms*, *15*(8), 291. https://doi.org/10.3390/a15080291

Meza, R. M., Vincze, H., & Mogos, A. (2019). Targets of online hate speech in context. *Intersections*, *4*(4). https://doi.org/10.17356/ieejsp.v4i4.503

Miškolci, J., Kováčová, L., & Rigová, E. (2018). Countering Hate Speech on Facebook: The Case of the Roma Minority in Slovakia. *Social Science Computer Review, 38*(2), 128–146. https://doi.org/10.1177%2F0894439318791786

Molala, T. S., & Makhubele, J. C. (2021). The connection between digital divide and social exclusion: Implications for social work. *Humanities & Social Sciences Reviews*, *9*(4), 194–201. https://doi.org/10.18510/hssr.2021.9427

Molnar, L. (2021). The Law of the Jungle. The Online Hate-speech against the Roma in Romania. *Victims & Offenders*, *16*(8), 1108–1129. https://doi.org/10.1080/15564886.2021.1895391

Mouratoglou, N., Agapoglou, T., Bikos, K. & Tsioumis, K. (2022a). Exploring the Social Representations of Roma in Online Hate Speech: findings from a preliminary study. In Th. Bratitsis (ed.), *Proceedings of the 12th Panhellenic and International Conference "ICT in Education"* (pp. 385-392). Florina, May 14-16, 2021. ISBN: 978-618-83186-5-6. Retrieved June 15, 2024, from: https://eproceedings.epublishing.ekt.gr/index.php/cetpe/article/view/3771

Mouratoglou, N., Agapoglou, T., Bikos, K. & Tsioumis, K. (2022b). Hindering democratic citizenship fulfilment: the case of online hate speech against Roma in Greece. In Katsillis, M. I. & Krzywosz-Rynkiewicz, B. (Eds.), *Young People's Citizenship and Education: Building Collaborations in Communities* (pp. 99 - 114). Children's Identity and Citizenship European Association. ISBN: 978-1-3999-4393-2. Retrieved June 15, 2024, from: https://www.cicea.eu/pdfs/Katsillis_&_Krzywosz-Rynkiewicz_eds_2022_Building_Collaborations_in_Communities.pdf#page=105.09

Negrea-Busuioc, E., Buturoiu, R., Oprea, D.-A. & Boţan, M. (2023). Hate speech in the Romanian online media and its impact on people's civic engagement with the Roma minority. *Revista Romana de Sociologie, 34*(1/2), 21-42.

Obermaier, M., & Schmuck, D. (2022). Youths as targets: factors of online hate speech victimization among adolescents and young adults. *Journal of Computer-mediated Communication*, *27*(4). https://doi.org/10.1093/jcmc/zmac012

Perera, S., Meedin, N., Caldera, M., Perera, I., & Ahangama, S. (2023). A comparative study of the characteristics of hate speech propagators and their behaviours over Twitter social media platform. *Heliyon*, *9*(8), e19097. https://doi.org/10.1016/j.heliyon.2023.e19097

Popenko, E. (2022). *The impact of the legal ambiguity of "hate speech" on the Roma community in Ukraine*. Central European University. eTD Collection. Retrieved June 15, 2024, from: https://www.etd.ceu.edu/2022/popenko_elvira.pdf

Preece, J. (2004). Online Communities: researching sociability and usability in hard to reach populations. *AJIS. Australasian Journal of Information Systems/AJIS. Australian Journal of Information Systems/Australian Journal of Information Systems*, *11*(2). https://doi.org/10.3127/ajis.v11i2.132

Rainie, L., & Wellman, B. (2012). *Networked: The new social operating system*. MIT Press.

Rheingold, H. (2000). *The virtual community: Homesteading on the electronic frontier*. MIT Press.

Schieb, C. &Preuss, M. (2016). Governing Hate Speech by Means of Counter Speech on Facebook. *66th ICA Annual Conference*, Fukuoka, Japan. Retrieved June 15, 2024, from: https://www.researchgate.net/publication/303497937_Governing_hate_speech_by_means_of__counterspeech_on_Facebook

Schmitt, M. T., Branscombe, N. R., Postmes, T., & Garcia, A. (2014). The consequences of perceived discrimination for psychological well-being: A meta-analytic review. *Psychological Bulletin*, *140*(4), 921–948. https://doi.org/10.1037/a0035754

Siitonen, M. (2007). *Social interaction in online multiplayer communities*. Jyväskylä studies in humanities 74. Retrieved June 15, 2024, from: https://jyx.jyu.fi/bitstream/handle/123456789/13444/9789513929312.pdf

Silva, L., Mondal, M., Correa, D., Benevenuto, F., & Weber, I. (2016). Analyzing the targets of hate in online social media. *Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM)*. https://doi.org/10.48550/arxiv.1603.07709

Sjöholm, M. (2024). Regulation of online gender-based hate speech and international human rights law: Current status and challenges. *Zoom-in, 104*, 3-16. Retrieved June 15, 2024, from: https://www.qil-qdi.org/regulation-of-online-gender-based-hate-speech-and-international-human-rights-law-current-status-and-challenges/

Solórzano, D. G., & Yosso, T. J. (2002). Critical Race Methodology: Counter-Storytelling as an Analytical Framework for Education Research. *Qualitative Inquiry*, *8*(1), 23–44. https://doi.org/10.1177/107780040200800103

Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior, 7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Tekiroğlu, S.S., Bonaldi, H., Fanton, M., & Guerini, M. (2022). Using Pre-Trained Language Models for Producing Counter Narratives Against Hate Speech: a Comparative

Study. *Findings*. https://doi.org/10.48550/arXiv.2204.01440

Tekiroğlu, S.S., Chung, Y.L & Guerini, M. (2020). Generating Counter Narratives against Online Hate Speech: Data and Strategies. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1177–1190, Online. Association for Computational Linguistics. https://doi.org/10.18653/v1%2F2020.acl-main.110

Thomas, S., Joseph, C., Laccetti, J., Mason, B., Mills, S., Perril, S., & Pullinger, K. (2007). Transliteracy: crossing divides. *First Monday*. Retrieved June 15, 2024, from: https://firstmonday.org/ojs/index.php/fm/article/download/2060/1908

Thompson, N., & Woodger, D. (2020). "I hope the river floods": online hate speech towards Gypsy, Roma and Traveller communities. *British Journal of Community Justice, 16*(1), 41–63. https://research.gold.ac.uk/id/eprint/27941/7/BJCJ_Thompson_and_Woodger_2020.pdf

Townsend, E. (2014). Hate speech or genocidal discourse? An examination of anti-Roma sentiment in contemporary Europe. Journal of Multidisciplinary International Studies, 11(1), 1-23. http://dx.doi.org/10.5130/portal.v11i1.3287

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.

Tynes, B. M., Giang, M. T., Williams, D. R., & Thompson, G. N. (2008). Online Racial Discrimination and Psychological adjustment among adolescents. *Journal of Adolescent Health*, *43*(6), 565–569. https://doi.org/10.1016/j.jadohealth.2008.08.021

Van Schaik, P., Jansen, J., Onibokun, J., Camp, J., & Kusev, P. (2018). Security and privacy in online social networking: Risk perceptions and precautionary behaviour. *Computers in Human Behavior*, *78*, 283–297. https://doi.org/10.1016/j.chb.2017.10.007

Voda, V. (2023). *The phenomenon of hate speech in public discourse: the case of Greece. Is hate speech correlated* with *hate crimes in Greece?* KU Leuven, European Master's Programme in Human Rights and Democratisation Retrieved June 15, 2024, from: https://repository.gchumanrights.org/items/f249c6b5-3ea4-4392-8669-e111072c53cb/full

Wachs, S., Gámez-Guadix, M., & Wright, M. F. (2022). Online Hate speech Victimization and depressive Symptoms among Adolescents: The Protective Role of Resilience. *Cyberpsychology, Behavior and Social Networking*, *25*(7), 416–423. https://doi.org/10.1089/cyber.2022.0009

Wellman, B., & Gulia, M. (1999). Net surfers don't ride alone: Virtual communities as communities. In B. Wellman (Ed.), *Networks in the global village* (pp. 331-366). Westview Press.

Wypych, M., & Bilewicz, M. (2024). Psychological toll of hate speech: The role of acculturation stress in the effects of exposure to ethnic slurs on mental health among Ukrainian immigrants in Poland. *Cultural Diversity & Ethnic Minority Psychology, 30*(1), 35–44. https://doi.org/10.1037/cdp0000522

Note 1. The most recent respective policy measure in Greece refers to the first update of the national strategy and action plan for the social inclusion of Roma 2021 – 2030 (Hellenic Ministry of Labour and Social Affairs, 2023).

Note 2. According to the Council of Europe (2018), the term 'internet intermediaries' refers to a broad and diverse group of entities that facilitate online interactions between individuals and organisations by providing various functions and services.

Note 3. https://www.coe.int/en/web/no-hate-campaign#:~:text=The%20No%20Hate%20 Speech%20Movement%20is%20a%20youth%20campaign%20led,national%20campaigns% 20in%2045%20countries