

# Combating Online Hate Speech through Critical Digital Literacy: Reflections from an Emancipatory Action Research with Roma Youths

Theodora Agapoglou (Corresponding author)

PhD Candidate

School of Philosophy & Education

Aristotle University of Thessaloniki, Greece

Nikolaos Mouratoglou

PhD Candidate

School of Philosophy & Education

Aristotle University of Thessaloniki, Greece

Konstantinos Tsioumis

Professor

School of Philosophy & Education

Aristotle University of Thessaloniki, Greece

Konstantinos Bikos

Professor

School of Philosophy & Education

Aristotle University of Thessaloniki, Greece

Received: April 11, 2021    Accepted: May 14, 2021    Published: May 21, 2021

doi:10.5296/ijld.v11i2.18524

URL: <https://doi.org/10.5296/ijld.v11i2.18524>

## Abstract

Despite the new and unprecedented opportunities for communication and networking in online settings, several challenges have arisen. One such challenge is the deviating behavior of social media users which is usually manifested through hate speech comments, often targeting minorities. For instance, the Roma minority constitutes one of the most discriminated groups, both online and offline. In this context, a significant lack of empirical research evidence can be observed regarding the exposure of minorities to online hate speech and the violation of their human rights. In this regard, the present paper illustrates the reflections of an emancipatory action research implemented with a group of Roma youths in Greece and discusses the impact of critical digital literacy activities on participants' awareness regarding online hate speech. Based on participants' views, the most empowering experience from this action research refers to the acquaintance of deliberating knowledge and skills in terms of producing non-violent counter-narratives.

**Keywords:** human rights, digital identity, empowerment, counter-narratives, critical digital citizens

## 1. Online Hate Speech: A “New” Form of Violating Minorities’ Civil Rights in the Cyber Era

In the modern digital era, characterized by the vast diffusion and exchange of information through advanced digital media and social networking platforms, democracy and citizens' participation are molded by new conditions, considering that the profile of the active emancipated citizen is discussed, redefined and ultimately adapted to social demands (Gilster, 1997; Lankshear et al., 1997; Koutsogiannis, 2011). The effective use of technology goes beyond the functional knowledge and ability to use software and hardware, encompassing the effective use of media in broader and diverse communication contexts (Bikos, 2012). Namely, the consumption and creation of digital content in various social media require complex social skills, as well as critical knowledge and values which are associated with digital citizenship, a concept that has emerged in policy discourse and academic literature to denote the norms of appropriate, responsible behavior of ‘netizens’ (Ribble et al., 2004; Frau-Meigs et al., 2017; Krutka & Carpenter, 2017), since human rights prove to be encroached in the cyberspace.

The increasing use of internet as a tool of disseminating hate and propaganda was identified by the European Agency of Fundamental Rights within the Framework Decision on Racism and Xenophobia (FRA, 2013). During the same year, the Youth Department of the Council of Europe launched the youth campaign “No Hate Speech Movement” to fight the dissemination of xenophobic and intolerant messages in the cyberspace (Lavchyan, 2013). Hate speech is the verbal expression of hostility, vilification, denigration, and stigmatization. It represents any form of expression that propagates, incites, promotes, or justifies forms of hatred based on intolerance, including expressions of aggressive nationalism and ethnocentrism, discrimination and hostility towards cultural minorities (de Latour et al., 2017). Although hate speech may be located at the least extreme end of the negative language continuum, it can still be harmful for individuals who are not resilient or capable of handling stressful

situations (Townsend, 2014).

Facilitating and distributing extremist beliefs and practices via social media should not be undervalued, while the boundaries between freedom of speech and others' human rights cannot be negotiated. To this respect, in 2016, the European Commission and major IT companies issued a public Code of Conduct targeting illegal online hate speech. Rules and community standards that prohibit hate speech were formulated, while artificial intelligence systems and experts undertook the role of reviewing the content that is reported to violate these standards (European Commission, 2016), with the aim of protecting individuals or groups that face online discrimination.

Although several European policies, initiatives and national laws protect minorities against hate speech (European Union Agency for Fundamental Rights, 2013; Amnesty International, 2020), Roma people still encounter severe obstacles regarding their social integration, remaining one of the most discriminated minority groups in Europe. This is also evident in Greece, despite the measures taken on the basis of the National Strategic Framework for Roma in 2011 (Hellenic Ministry of Labour and Social Security (2011) and the updated Action Plan for the National Roma Integration Strategy (Hellenic Ministry of Labour, Social Insurance and Social Solidarity, 2019). Since integrating minorities and protecting their rights lie at the core of democracy and are closely related with the concept of citizenship, international and European organizations stress the need of digital literate citizens with particular emphasis on vulnerable and minority groups (CDCPP, 2017). For instance, the acquisition of relevant knowledge and skills may facilitate the active participation of minority groups in social, political and economic life.

However, critical digital literacy, which combines digital literacy with critical literacy, may promote the democratic values and principles of citizenship by forming emancipated digital citizens who are able to protect themselves from harmful digital material and respect others. This process may include privacy issues, awareness regarding online presence and netiquette, the power of the media, conventions of digital products and ideological interests (Hinrichsen & Coombs, 2014). For example, being aware of mis- dis- and mal-information, individuals may question the credibility of various statements and information, examine whether their sources are reliable and reflect on any hidden "agendas" that promote certain viewpoints, stigmatize social groups, and even normalize or reinforce hatred and violence.

Awareness on online hate speech calls for a critical perspective since interactions in social media are always culturally situated. Therefore, the theoretical foundations of the present research, apart from those of Digital Citizenship mentioned above, derive also from Critical Intercultural Education (Ghosh, 2002; Kumashiro, 2001). The theory focuses on equality and social justice, promoting respect for diversity, as well as protecting human rights, facilitating social transformation, and leading to emancipation. Additionally, considering the mediating role of language in meaning construction, the innovative model of critical digital literacy (Hinrichsen & Coombs, 2014) constituted a promising framework for supporting vulnerable groups in becoming emancipated 'netizens'. This model draws on media education and critical theories and pedagogies, namely the model of critical literacy suggested by Freebody

and Luke (1990), also known as the “four-source model”. The critical digital literacy model offers a re-interpretation of the four resources by acknowledging the developmental stages of critical competence, adapted to digital environments. However, the model integrates an additional fifth “source”. Thus, the five sources are: “decoding” (practical and functional engagement), “meaning construction” (narrative complexity in digital environments), “use” (production and consumption of digital texts), “analysis” (critical judgment) and “persona” (identity issues and digital environments).

Since the volume of abusive digital content that threatens fundamental values of a democratic society is growing, an emerging need for mobilization, planning and campaigning with young people against online hate speech has been identified (see Grizzle et al., 2014). Even though several theories provide a fruitful background for describing and discussing such topics, nevertheless, there is a significant lack of empirical research evidence on the content and means of empowering vulnerable groups in online communities. To this response, we developed and implemented a critical digital literacy intervention program with Roma youths for tackling online hate speech through counter-narratives. Hence, the aim of the present paper is to describe the basic phases of a critical-emancipatory action research, to illustrate our reflections and discuss the empirical implications deriving from this endeavour.

## 2. Research Identity

The critical-emancipatory action research was considered as the most appropriate method, as it encompasses both theory and practice and is based on two principles that are inextricably linked to the research purpose: participation and change. In this case, change is related not only to the personal transformation and empowerment of the participants, but also to the transformation of the cyber space at a micro-level by facilitating the active participation of Roma youths in defending their rights in social media. Moreover, it is important to highlight that critical-emancipatory action research is an inherently democratic approach that encourages participants to control the process of the occurring changes (Elliott, 1991).

The action research started by identifying the problem that needs intervention and change. Although Roma youths are familiar with online media and use them intensively in their daily lives (Garmendia & Karrera, 2019), they often present an inability to critically analyse the causes of online hate speech and manage it effectively. Thus, they end up being or feeling inactive, resigned, marginalized, impulsive and/or aggressive. In 2018, the UN Human Rights Council acknowledged youth as the main target group of the fourth phase of the World Programme for Human Rights Education. The aim was to encourage efforts to promote human rights education and training in equality, non-discrimination, diversity, and inclusion by fostering civic awareness and participation, as well as building inclusive and peaceful societies (United Nations Youth Strategy, 2015). In this context, the present non-formal educational intervention, following the stages of the action research method, involved a group of Roma youths in activities that focused on critical digital literacy, non-violent communication, active participation, and human rights.

The action research took place in one of the most challenging districts of Thessaloniki, a local Roma community named Dendropotamos. The majority of non-Roma is usually unwilling to

visit the region because of the existing stereotypes and prejudices that are produced and reproduced both online and offline. As a result, there are few opportunities of interacting with other people. For this reason, before the action research, the research team carried out a long-term ethnographic research in order to become familiar (“cultural immersion”) with the social norms and cultural practices of the Roma community (Fetterman, 1998; Spradley, 1979; Thomas, 1993). An array of ethnographic methods and tools, such as observation, interviews, production of field notes and capture of audio data were employed. With the support of Fr. Athenagoras and the “Lighthouse of the World”, the participants attended 21 non-formal educational workshops that lasted approximately five hours each. Half of them took place in the “Light house of the World” in Dendropotamos, while the rest in the Department of Philosophy and Education of the Aristotle University of Thessaloniki.

The learning content of the activities included topics such as cultural identity negotiation, effective management of intercultural conflicts, human rights, diversity, respect, and democracy. To ensure participants’ active engagement, we deployed non-formal education techniques, such as learning energizers, simulators, and experiential activities. Additionally, all activities were based on digital media and tools like online gamified applications, timelines, digital comics, and digital storytelling tools. To this regard, digital literacy tools and practices, such as developing web quests, using conceptual maps and digital diaries played a significant role in promoting participants’ learning and reflection.

The action research consisted of three circles, each one including planning, acting, observing, reflecting on practice, re-planning and so on (Kemmis, McTaggart & Nixon, 2014). Recording the workshops was a key strategy to capture the critical moments and analyse them with the support of our critical friends. Participants’ diaries and focus group discussions were systematically analyzed during the debriefing sessions. The triangulation of data collection, namely the observation of the researchers/educators, outputs produced by the youths and the perspectives of the critical friends, enriched the validity and reliability of the action research.

The qualitative data obtained from the observation, the focus groups, and the diaries illustrated participants’ transformations and reflected the gradual empowerment of each participant, at every step of the action research. The data were analyzed through Qualitative Content Analysis (Mayring, 2010), based on content construction, standard construction and typical construction of the interpretive and ideological-critical method (Bonidis, 2004). In this method, the focus lies on the internal structure of language, revealing hidden meanings “behind the words” and discussing them dialectically within the broader social, political, economic and cultural context (Holsti, 1969).

### **3. The Three Circles of the Action Research**

At the beginning, several “ice-breaking” and “team-building” activities were carried out, using various digital applications such as TikTok energizers and online games (online game show quiz, true or false, random cards etc.) that we specially designed to ensure a friendly, cooperative, playful and trustworthy atmosphere between the participants and us. Upon this introductory part, the Roma youths were invited to express their needs, expectations, and

concerns regarding the workshops, as well as describe their motivation and expected impact from this learning experience. Participants' statements were analyzed and used as a guide for designing the content and topics of the rest workshops.

The first circle included activities that focused on raising participants' awareness regarding their online identity and attitude as cyber citizens, as well as the presence of hate speech in cyber communities. The first activity included an online questionnaire inquiring participants' daily practices in relation to their digital presence (online behavior, digital devices, hours of navigation, social media channels etc.). The questionnaire not only provided important and useful data concerning their digital literacy practices, but also constituted a self-reflection opportunity as the participants became aware of the power that ICT, and mostly their mobile phones, have in their daily life. Participants visualized their data in diagrams via an online application and then discussed the answers of their team as a whole. It is also important to note that these data fed the activities that we carried out next, since each activity delved into a specific thematic question of the questionnaire.

Based on these diagrams, participants observed that they have a strong preference for interacting via Instagram, TikTok and Facebook, while they are unfamiliar with digital literacy tools and practices, such as developing web quests, using conceptual maps, digital diaries, timelines, digital comics and digital storytelling software. This conclusion provoked their reflection on a diverse range of topics including online communication, ethical aspects, privacy issues, ways of creating and sharing content, as well as examining the reliability of the information and its sources. Participants were split in groups and requested to search relevant information online and keep notes in a Padlet. Upon their WebQuests, participants shared the resources and new knowledge that they gained with the rest of the group during the plenary session.

The following activity invited them to reflect on the time they spend daily in social media, with the majority claiming that they do not navigate for more than three hours. The participants were invited to check the duration of their online presence in social media via their mobile phones. This proved to be a really enlightening activity, since they soon realized that they underestimated the duration, as one indicative statement reveals: "Wow...I thought that I was online three hours maximum per day. Actually, I was thinking that I spend even less time online and I thought that I exaggerated when I wrote down three hours in the questionnaire. I cannot believe it... The data of my phone have recorded five hours and seven minutes and it's not even night! Time flies when you are connected to social media...". The rest of the participants expressed similar views. This realization triggered their reflection on the power of the digital media in their daily life, as well as on the ways they use them. For instance, most of them stated that they chat with their friends, look for information and entertain themselves by watching videos, listening to music or "killing [their] time". They realized that, sometimes, they chat with their friends online even though they are neighbours and therefore they could easily meet in person. They also stated that they often play online games in which they interact with other people that they might not know. Our debriefing discussion focused on how both the internet and the various ICT tools have transformed traditional communities into cyber, expanded the communication channels, enlarged the

interaction with other people and modified human behavior and social norms.

At the same moment, though, we challenged them to reflect on the online presence of various social groups (including Roma), their experiences from interacting (or not) with those groups, as well as on the reasons that some groups are underrepresented in online settings. Via the WordClouds online application, they produced two interactive wordclouds using keywords of the two different contexts (online/offline) and compared their characteristics. For instance, they described the meaning that they attribute to the term “friend” in digital and *vis-à-vis* settings. We then encouraged them to focus on the nature of the online interactions. They reported that although they have lots of online friends, they do not interact with the majority of them. In addition, they realized that the usual bond for the creation of online groups is whether there is a common interest or not. Given this element, we initiated a discussion on the various groups that they have come across and focused on the topics of such groups. Soon enough, the participants acknowledged political, ideological, societal and lifestyle factors that contribute to the creation of online groups.

In another activity, the youths looked into their preference of watching photos and videos on social media rather than reading large posted texts. Based on this, we started a discussion on the different modes and conventions of digital texts (e.g., images, videos/stories, short messages, posts, infographics). Considering that the term “text” is a multimodal term (Kress, 2003) we wanted to ensure that all of the participants would be in position to discover the aim of the message, the target group, as well as any potential implicit interests. In addition, they were encouraged to reflect on the habit of reposting information, without checking its reliability and therefore become part of a vicious circle of disseminating fake news and disinformation. With regard to fake news and disinformation, participants made several references to their cultural group and criticized the way that people take into account false information, make negative comments and do not use reasonable arguments in their posts. Throughout this process they recognized the readers’ agency and role as a participant in meaning construction. They also identified that the content, style and purpose of the text is in dialogue with the readers’ prior experience and knowledge.

The following activity inquired participants’ perceptions regarding hate speech. The Roma youth watched a video that provided several examples of hate speech and shared their views on how they define hate speech, what characteristics it entails, who are the individuals/groups in target, why some groups are more susceptible to hate speech and which conditions favour the expression of hate speech. With our support, participants were encouraged to work in groups and undertake an online trawl to collect data and become aware of the complexities associated with online hate speech. Each group shared the data with the rest of the group using visualization tools (e.g., infographics, images, PowerPoint presentations, conceptual maps). After that, they contributed to building a Kahoot quiz by proposing questions about these topics, while after a break, all of them were invited to participate in the Kahoot game using their mobile devices.

The discussion on more concrete aspects of hate speech led us to proceed with the individuals’ reactions, as, according to their responses in the initial questionnaire, their most

popular reaction to hate speech was: “If I identify a negative statement or a post full of hate, which I do not approve it, my intention is to find out who wrote it and meet him/her, to speak in person”. It is also important to highlight that many youths did not exclude the possibility of using violence as a reaction. For example, they claimed: “It depends on the post, how violent I will become...” or “I may even beat him!”. However, it should be underlined that one participant argued that whenever he identifies a hate speech narrative, he usually reports this to the group administrators or the webmasters of the application. This statement was something unknown for the rest, so we invited that participant to share his knowledge with the whole group.

With the completion of the activities, we invited the group in a debriefing session to discuss any potential changes resulting from this process and encourage their reflection. At that moment, we had already analyzed the recordings of the workshops’ discussions and participants’ diaries that denoted a gradual transformation of the participants’ discourse. Many of the youths took the initiative to start a conversation based on specific online hate speech narratives that they identified at their free time while browsing in social media. The activities proved to have promoted their self-awareness regarding their online persona, since they paid more attention to the duration of their mobile use and to the ways they construct their identity through the posts, photos and videos that they uploaded and/or shared. Furthermore, they indicated that they are now more attentive to the posts they choose to repost, as well as to the people with which they connect and communicate. At this stage, the Roma youths seemed more curious about learning new ICT applications and how they can be integrated in their daily school/life practices. Moreover, they were capable of identifying hate speech and appeared to be concerned about the victims of hate speech. However, they did not make any clear reference to the cyber users’ responsibility to behave respectfully in online communities or to the risks that hate speech poses for democracy. It was obvious that they were not ready to conceive themselves as online “human rights defenders” (Keen, et al., 2020).

Therefore, the second circle of the action research focused on promoting a deeper consciousness in relation to human rights, the essence of democracy and individuals’ role as responsible citizens in the digital village. A simulation game in combination with the well-known activity “Take a step forward” (Brander, 2012) and an educational game designed by the research team named the “Wheel of Human Rights” promoted participants’ awareness on the necessity of ensuring human rights both offline and online. Another online educational game designed by the research team, named “Open the box”, included case studies according to which human rights may be challenged in online social communities. This activity invited youths to process the hate narratives in conjunction with the Declaration of Human Rights and reflect on the value system and attitudes of cyber citizens to ensure respectful and peaceful online interactions. This activity promoted their understanding on the necessity of respecting human rights both in offline and online settings, while at the same time facilitated their critical consciousness regarding the interplay between the digital and non-digital/conventional world.

In the “World Café” activity (Anderson, 2011), youths were requested to search for hate



speech narratives in Twitter, Facebook, Instagram, TikTok and YouTube and copy paste them in their group workspaces (Padlets). When all were ready, they were asked to identify which human rights are violated, explore the balance between the freedom of speech and hate speech, as well as reflect on the root causes of hate in each narrative. This activity enhanced their capacity of discerning the elements that contribute to meaning making and alternative (mis-)uses of digital content to reveal hidden intentions and understand the feelings and ideologies of the hate speech drivers. They were also encouraged to work in groups in an online activity named “Hate-meter” that was designed by the research team. Specifically, they were asked to select one of the narratives mentioned above, put themselves in the shoes of the victims of hate speech and explore how they felt when reading the hate speech narrative. Afterwards, they identified their emotional reactions based on an intensity scale of hate speech (Bahador, 2020) and noted down their thoughts evoked by these narratives. At the end, they were encouraged to think about the way(s) that they would react in real life. At this phase, the youths created posters for raising awareness on the existence and consequences of hate speech both for the individuals and the society in general. Through this activity they applied what they learned in earlier workshops, such as the role of the style, font size, colours and images, which should all be in dialogue with the purpose of the text.

The data analysis of the second circle revealed that the participants began to highlight the essence of their responsibility in respecting others in online communities. It was also evident that they kept in mind the reporting mechanisms concerning online hate speech either to social media administrators or to the observatory platform of the Council of Europe, strategies discussed at the previous circle. However, at that time they seemed being rather unaware of the ways that they could practically combat online hate speech. In other words, they needed hands-on strategies for addressing online hate speech, strengthen their arguments and develop empathy.

Therefore, the third circle of the action research focused on non-violent communication strategies (Rosenberg, 2002; 2015) and the use of humour, as a mean of producing counter-narratives for combating online hate speech. Firstly, through an educational video they were presented with two alternative reactions, a violent and a non-violent, of the same speaker towards his interlocutor. Participants were encouraged to compare the communication stages and the consequences of each approach, illustrating the two processes in timelines by using an online application. Then, they watched a second video, discussed the principles of nonviolent communication in groups and explored the steps of non-violent communication. Through these introductory videos, the Roma youths were motivated to join an online activity based on de Bono’s six thinking hats (de Bono, 2016). The participants chose one discriminatory criterion reflecting the hater’s ideology, such as nationality, religion, ethnicity, gender, age bracket, or sexual orientation and explored the social media to locate hate speech narratives. Upon this process, they were encouraged to analyse each narrative by creating a counter-narrative based on the guiding questions that each one of the six hats entailed.

In specific, the ‘white hat’ invited them to make sure whether their arguments included in their counter-narratives fitted with the existing information they had. Additionally, they

reflected on the extra information needed/missing and on the ways that they can receive this information, as well as on the questions they needed to ask in order to intervene and combat hate speech. In simple words, at this stage, participants reflected on the data and information they have, identified any potential deficiencies and defined sources for acquiring additional reliable information to produce their counter-narratives. The “red hat” invited them to express their emotions and feelings and share their concerns and dislikes regarding the narrative they selected. The “green hat” promoted their creativity, by nudging them to think about the possibilities of reacting against the narrative, instead of remaining silent (Freire, 1985). It actually provided the space to explore their strategies while producing online counter-narratives by using the principles and stages of non-violent communication.

The “black hat” referred to identifying the risks, difficulties and problems that would possibly arouse after publishing their counter-narratives. It served as a “pre-reflection” on their proposed strategies. On the contrary, the “yellow hat” invited them to think the optimistic outcome and the positive consequences of their designed actions. Last but not least, the “blue hat” encouraged them to overview their thinking process reflecting on the proper ways for taking action, the transformations occurred to their planned strategies after the collection of all the information needed, the ways that the rest of the hats helped them to increase clarity about the topic and on the new knowledge that they have acquired so far. At the end of this process, they were capable of deciding the content of their counter-narratives by using appropriate tone-style, considering the audience addressed, choosing the medium and ensuring the advocacy of human rights and universal values.

Participants were also invited to play an online game designed by the research team in order to become acquainted with the characteristics of humour, using true or false questions. They were, then, encouraged to explore four types of humour, namely the affiliative, the aggressive, the self-enhancing and the self-defeating by filling in the Humour Styles Questionnaire (Martin et al., 2003). In the next activity, participants chose a hate speech comment and prepared a counter-narrative using their humour style. The most important aspect in this activity is that the hate narrative referred to the Roma minority. Although the narrative was extremely offensive, participants succeeded in managing their feelings, since they came up with inspiring responses using humour. At the final phase, the Roma youth produced a digital comic book, by creating scenes and dialogues based on the hate comments and the counter narratives that they produced.

The analysis of participants’ engagement during the third circle, as well as of the data deriving from their personal diaries indicated that the activities enabled them to employ analytical and critical skills. At that stage, they seemed more capable of translating their ideas, feelings and intentions into a digital form, becoming, though, more aware of the ethical, social and cultural aspects of their online interactions. For instance, they realized that their individual personas constitute “digital ambassadors” of their local community and their social group at large, suggesting that they should be more careful on their online behavior as they do not want to contribute to their ongoing stigmatization and marginalization. However, it is important to acknowledge that many of them expressed the need of actively safeguarding Roma’s reputation and human rights through counter-narratives that will be based on

reasonable arguments and humour that are often effective forms of tackling prejudices and stereotypes.

In our last meetings, participants were encouraged to practice the skills and knowledge acquired throughout the workshops in a simulation game designed by the research team. The aim was to check the youths' intention and means, while responding to online hate speech. The simulation game took place in Padlet and specifically in an interface that resembles with chat in social media applications. The youths were invited to interact with three non-Roma participants whose identities were not revealed to them. The majority of the participants acknowledged their personal responsibility for tackling hate speech while interacting with haters, as well as that if they use violent communication approaches, this will inevitably lead to violent responses, feeding into a vicious circle that will never end.

#### **4. Discussion**

Responding to the changing dynamics of the digital world, we attempted to engage Roma youths in the promotion of their critical consciousness regarding their social rights, which are often violated in the context of hate speech. Carrying out the action research during the Covid-19 pandemic was really challenging, considering the rising ideology that Roma are responsible for spreading the virus in their local societies (European Roma Rights Centre, 2020). In this regard, our aim was to empower the voices of the Roma youth in response to hateful comments by cultivating their critical digital skills and promoting their self-awareness. Even though critical digital literacy (Hinrichsen & Coombs, 2014) proved to be an effective framework for the cultivation of skills and competences for the democratic citizens of the 21st century, there is a general lack of empirical evidence, especially for non-privileged social groups. Being the first national attempt of integrating critical digital literacy in an intervention program and most importantly while working with a minority group such as the Roma, action research constituted the most suitable method for addressing the group's needs and expectations. This action research actually testified the power of theory in practice by producing new knowledge (Kemmis, McTaggart, & Nixon, 2014), which was then multiplied through the youths to their wider local community. Additionally, the Roma youths were not just participating, rather building the content and guiding the whole process, creating, in this way, a double-aimed educational program "with Roma and for Roma". Based on their concluding reflection, this process ensured their participation and motivation, along with the experiential and gamified activities that supported their learning such as digital storytelling. Our finding is in line with the conclusion of the Office of the High Commissioner Human Rights (2019) that identified storytelling as an appropriate and engaging methodology in human rights education for and by youth.

Throughout the workshops the youths acquired and developed critical digital literacy competences, in terms of knowledge, skills, values and attitudes (Barrett, et al., 2018). Namely, they became aware of the communication dynamics and the ethics of online interaction, while they integrated educational online applications in their daily activities, which undermined the exclusivity of social media applications. They also acknowledged that they became more capable of identifying, managing, evaluating and creating digital content,

fostering their digital literacy skills (Gilster, 1997). However, beyond their increased comfort in using digital media to promote their learning, the Roma youths became more sceptical about the way they use the digital media, they reflected on their online social interactions and realized that privacy, ethics and respect are not granted in online settings. Considering their online personas and digital footprint in combination with their cultural identity, they were able to reflect at a broader level between the online and offline social world and conclude that the online world is a mirror of the offline world (Dunbar, Arnaboldi, Conti & Passarella, 2015). In other words, they discussed hate speech in dialogue with the broader movement of antigypsyism and racism that they face. This comparison led them to recognize, reject and stand up against online hate speech by articulating a new alternative, non-violent discourse that aimed at conflict resolution (Rosenberg, 2002; 2015). In that sense, the Roma youths were empowered to uncover the conventions and stereotypes that are transferred in the online world by using rational arguments and agreed that taking action is key for breaking their silence (Freire, 1985). They realized the value of democracy and cooperation, as well as the need for openness and respect to Others' beliefs and worldviews (Barrett et al., 2018). Hence, digital citizenship and online hate speech were well-suited topics for the cultivation of critical digital literacy.

Acknowledging the limitations of this action research concerning the number of the participants involved and the duration of the intervention due to the pandemic, we propose that future initiatives should last longer and engage larger groups. Considering the extensive use of social media and ICT tools, along with the potential harmful content that is easily created and disseminated through social media, we would suggest that future action researchers could address other minoritized groups whose rights are also violated online. Last but not least, it is really crucial to engage people from privileged groups (e.g., non-Roma) in future action researches, since they can become agents and multipliers of social change, promote symmetrical intercultural interactions with vulnerable groups and become advocates of human rights.

## **5. Conclusion**

Considering that all digital citizens can be perceived as digital ambassadors of their culture in online communities, reflection on users' personas and individuals' digital practices may constitute a preventive measure for expressing or being exposed to hateful content. As social media and the extended use of the internet has offered several opportunities for promoting freedom of speech, at the same moment several challenges arose in relation to dis/mis/mal-information, propaganda and hate speech. Hence, the need for resilience, critical awareness and safe online navigation emerges as a significant necessity of our times, indicating that digital literacy practices should now inherently integrate critical components to respond both to social and digital encounters. The aforementioned necessity becomes even more urgent when interpreted in the light of antigypsyism and of the fragility of our democracies. Supporting minorities' empowerment through critical digital literacy interventions may contribute to questioning the existing online cultural inequalities and racism and deconstructing the negative social representations of their groups imposed by the privileged social groups. In this way, minorities may become online human rights defenders

by deploying counter-narratives to trigger intercultural interactions aiming at raising awareness and refining distorted beliefs, prejudices and eventually stereotypes.

The research findings indicate that the Greek Roma minority is recipient of online hate speech. However, such manifestations of negative discourse can be employed as authentic learning resources in non-formal activities, with the aim of empowering participants. This can be mainly achieved by providing opportunities for raising their critical awareness, strengthening their voices and promoting their purposeful action. Taking into account that violence leads to violence, a reflective and differentiated approach of non-violent counter narratives, as well as the use of humour proved to be rather efficient and deliberative strategies for combatting online hate speech. In this action research, active civic participation has been acknowledged as a key element of promoting Roma youths' democratic citizenship in terms of tackling racism and ensuring mutual understanding and respect, while digital citizenship extended to the online social sphere to highlight the respective significance of digital civic participation. All these aspects were synthesized and further contextualized based on their experiences as learning material by the Roma youths themselves, a process that promoted their active engagement and sustained their motivation throughout the critical digital literacy intervention.

In relation to the main strategy for combatting online hate speech, counter-narratives have been considered as an effective approach for responding to hate comments. In specific the Roma youths became familiar with the structure and the way of formulating counter-narratives. Concurrently, they became acquainted with the use of non-violent communication and humour in producing counter-narratives. Based on those two elements, the youths discovered that they could take action without, though, insulting or abusing other online users. In fact, some participants used concrete examples from their everyday life and made several references to customs of their cultural group to build their arguments. At the same moment, the youths attempted to share information regarding their cultural practices with the non-Roma haters aiming at raising the awareness of the wider online community. Hence, their counter-narratives reflected an emancipated praxis of digital civic engagement that created the space and the conditions for developing a respectful and non-violent dialogue.

### **Acknowledgments**

This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning 2014-2020» in the context of the project “Act, Interact and Impact to counter Hate Speech: Cultivating Critical Digital Literacy Skills among Roma Youth” (MIS 5047859).



## References

- Amnesty International (2020). *Human rights in Europe. Review of 2019*. Retrieved from <https://www.amnesty.org/download/Documents/EUR0120982020ENGLISH.PDF>
- Anderson, L. (2011). Use the world café concept to create an interactive learning environment. *Education for Primary Care*, 22(5), 337-338. <https://doi.org/10.1080/14739879.2011.11494028>
- Bahador, B. (2020). *Classifying and Identifying the Intensity of Hate Speech*. Retrieved from <https://items.ssrc.org/disinformation-democracy-and-conflict-prevention/classifying-and-identifying-the-intensity-of-hate-speech/>
- Barrett, M., De Bivar Black, L., Byram, M., Faltýn, J., Gudmundson, L., Van't Land, H., Lenz, C., Mompoin-Gaillard, P., Popović, M., Rus, C., Sala, S., Voskresenskaya, N. & Zgaga, P. (2018). *Reference framework of competences for democratic culture: Volume 1: Context, concepts and model*. Strasbourg: Council of Europe Publishing. Retrieved from <https://rm.coe.int/background-to-the-framework-reference-framework-of-competences-for-dem/1680994101>
- Bikos, K. (2012). *Pedagogical issues raised by information and communication technologies* [In Greek]. Thessaloniki: Libra.
- Bonidis, K. (2004). *The content of the textbook as an object of research. Diachronic examination of the relevant research and methodological approaches* [In Greek]. Athens: Metaichmio.
- Brander, P. (2012). *Compass: Manual for human rights education with young people*. Strasbourg: Council of Europe.
- CDCPP-Steering Committee for Culture, Heritage and Landscape (2017). *Draft recommendation of the committee of ministers to member states on big data for culture, literacy and democracy: Follow-up to the 3<sup>rd</sup> Council of Europe exchange platform on the impact of digitization on culture*, Tallinn, September 2016. Retrieved from <https://rm.coe.int/draft-recommendation-on-big-data-for-culture-literacy-and-democracy-10/168075de34>
- de Bono, E. (2016). *Six thinking hats*. London: Penguin Life.
- de Latour, A., Perger, N., Salag, R., Tocchi, C., & Otero, P. V. (2017). *We can!: Taking action against hate speech through counter and alternative narratives*. Budapest: Council of Europe.
- Dunbar, R. I. M., Arnaboldi, V., Conti, M., & Passarella, A. (2015). The structure of online social networks mirrors those in the offline world. *Social Networks*, 43, 39-47. <https://doi.org/10.1016/j.socnet.2015.04.005>
- Elliott, J. (1991). *Action research for educational change*. Buckingham and Bristol, PA: Open University Press.

European Commission (2016). *Code of conduct on countering illegal hate speech online*. Retrieved from [https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online\\_en](https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en)

European Roma Rights Centre, ERRC (2020). *Roma rights in the time of Covid*. Retrieved from [http://www.errc.org/uploads/upload\\_en/file/5265\\_file1\\_roma-rights-in-the-time-of-covid.pdf](http://www.errc.org/uploads/upload_en/file/5265_file1_roma-rights-in-the-time-of-covid.pdf)

Fetterman, D. M. (1998). *Ethnography: Step by step*. Thousand Oaks: Sage Publications.

FRA-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 from [https://fra.europa.eu/sites/default/files/fra-opinion-2-2013-framework-decision-racism-xenophobia\\_en.pdf](https://fra.europa.eu/sites/default/files/fra-opinion-2-2013-framework-decision-racism-xenophobia_en.pdf)

Frau-Meigs, D., O'Neill, B., Soriani, A., & Tomé, V. (2017). *Digital Citizenship Education: Volume 1: Overview and new perspectives*. Strasbourg: Council of Europe. Retrieved from <https://book.coe.int/en/human-rights-democratic-citizenship-and-interculturalism/7452-digital-citizenship-education-volume-1-overview-and-new-perspectives.html>

Freebody, P. & Luke, A. (1990). Literacies programs: Debates and demands in cultural context. *Prospect: An Australian Journal of TESOL*, 5(3), 7-16. <https://eprints.qut.edu.au/49099/>

Freire, P. (1985). *The politics of education: Culture, power and liberation*. Cambridge, MA: Bergin and Garvey. <https://doi.org/10.1007/978-1-349-17771-4>

Garmendia, M. & Karrera, I. (2019). ICT use and digital inclusion among Roma/Gitano adolescents. *Media and Communication*, 7(1), 22-31. <http://dx.doi.org/10.17645/mac.v7i1.1624>

Ghosh, R. (2002). *Redefining multicultural education*. Scarborough: Nelson Thomson Learning.

Gilster, P. (1997). *Digital literacy*. New York: Wiley Computer Pub.

Grizzle, A., Moore, P., Dezuanni, M., Asthana, S., Wilson, C., Banda, F., & Onumah, C. (2014). *Media and information literacy: Policy and strategy guidelines*. Paris: Unesco.

Hellenic Ministry of Labour and Social Security (2011). *National strategic framework for Roma*. Retrieved from [https://ec.europa.eu/info/sites/info/files/greece\\_national\\_strategy\\_en.pdf](https://ec.europa.eu/info/sites/info/files/greece_national_strategy_en.pdf)

Hellenic Ministry of Labour, Social Insurance and Social Solidarity (2019). *National operational action plan for the social integration of Roma* [In Greek]. Special secretariat for the social integration of Roma. Retrieved from <https://egroma.gov.gr/wp-content/uploads/2019/06/LOW-%CE%95%CE%A0%CE%99%CE>

%A7%CE%95%CE%99%CE%A1%CE%97%CE%A3%CE%99%CE%91%CE%9A%CE%9F\_SXEDIO\_DRASHS\_ROMA.pdf

Hinrichsen, J., & Coombs, A. (2014). The five resources of critical digital literacy: A framework for curriculum integration. *Research in Learning Technology, 21*, 21334. <https://doi.org/10.3402/rlt.v21.21334>

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading/Mass: Addison-Wesley.

Keen, E., Georgescu, M., & Gomes, R. (2020). *Bookmarks: A manual for combating hate speech online through human rights education*. Strasbourg: Council of Europe.

Kemmis, St., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. London: Springer. <https://doi.org/10.1007/978-981-4560-67-2>

Koutsogiannis, D. (2011). *Adolescents' digital literacy practices and identities* [In Greek]. Thessaloniki: Centre for the Greek Language.

Kress, G. (2003). *Literacy in the new media age*. London: Routledge. <https://doi.org/10.4324/9780203299234>

Krutka, D. G., & Carpenter, J. P. (2017). Digital citizenship in the curriculum. *Educational Leadership, 75*(3), 50-55. Retrieved from <http://www.ascd.org/publications/educational-leadership/nov17/vol75/num03/Digital-Citizenship-in-the-Curriculum.aspx>

Kumashiro, K. (2001). Toward a theory of anti-oppressive education. *Review of Educational Research, 70*(1), 25-53. <https://doi.org/10.3102%2F00346543070001025>

Lankshear, C., Gee, P., Knobel, M., & Searle, Ch. (1997). *Changing literacies*. Buckingham. Philadelphia: Open University Press.

Lavchyan, Z. (2013). *The conference of the no hate speech movement campaign. Mobilising, planning and networking for campaigning effectively with young people against hate speech online*. 7-9 November, European Youth Centre Strasbourg. Retrieved from <https://rm.coe.int/report-the-conference-documentation-book-nhsm-7-to-9-nov-2013/168079725b>

Martin, R., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the humor styles questionnaire. *Journal of Research in Personality, 37*(1), 48-75. [https://doi.org/10.1016/S0092-6566\(02\)00534-2](https://doi.org/10.1016/S0092-6566(02)00534-2)

Mayring, P. (2010). *Qualitative inhaltsanalyse*. Retrieved from <https://www.beltz.de/fileadmin/beltz/leseproben/978-3-407-25730-7.pdf>

Office of the High Commissioner Human Rights (OHCHR) (2019). *Seven key elements on building human rights-based narratives on migrants and migration*. Retrieved from



<https://www.ohchr.org/Documents/Issues/Migration/SevenKeyElements.pdf>

Ribble, M. S., Bailey, G. D., & Ross, T. W. (2004). Digital citizenship: Addressing appropriate technology behavior. *Learning & Leading with Technology*, 32(1), 6-12. Retrieved from <https://files.eric.ed.gov/fulltext/EJ695788.pdf>

Rosenberg, M. B. (2002). *Nonviolent communication: A language of compassion*. Encinitas, CA: Puddledancer Press.

Rosenberg, M. B. (2015). *Living nonviolent communication: Practical tools to connect and communicate skillfully in every situation*. Encinitas, CA: Puddledancer Press.

Spradley, J. P. (1979). *The ethnographic interview*. New York: Holt, Rinehart and Winston.

Thomas, J. (1993). *Doing critical ethnography*. Newbury Park, CA: Sage Publications. <https://doi.org/10.4135/9781412983945>

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>

UNESCO. (2015). *Global citizenship education: Topics and learning objectives*. Paris: UNESCO. Retrieved from [https://static1.squarespace.com/static/52ec6057e4b0f081ced7f862/t/5cd113acb208fc49d0f04ac4/1557205956233/unesco\\_gced+Topics+and+Learning+Objectives.pdf](https://static1.squarespace.com/static/52ec6057e4b0f081ced7f862/t/5cd113acb208fc49d0f04ac4/1557205956233/unesco_gced+Topics+and+Learning+Objectives.pdf)

United Nations Youth Strategy. (2015). *Youth 2030: Working with and for young people*. Retrieved from [https://www.un.org/youthenvoy/wp-content/uploads/2018/09/18-00080\\_UN-Youth-Strategy\\_Web.pdf](https://www.un.org/youthenvoy/wp-content/uploads/2018/09/18-00080_UN-Youth-Strategy_Web.pdf)

### Copyright Disclaimer

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).