

National and Cultural Identity in Mass Media: How to Resist Manipulations and Hate Speech

Oksana Sakhatska*

Department of Sustainable Technologies, Faculty of Tropical AgriSciences, Czech University of Life Sciences Prague, Kamýcká 129, 165 00 Prague

***Corresponding author:** Oksana Sakhatska, Department of Sustainable Technologies, Faculty of Tropical AgriSciences, Czech University of Life Sciences Prague, Kamýcká 129, 165 00 Prague.

Submitted: 16 April 2025 **Accepted:** 22 April 2025 **Published:** 29 April 2025

doi <https://doi.org/10.63620/MKJGPSCD.2025.1008>

Citation: Sakhatska, O. (2025). National and Cultural Identity in Mass Media: How to Resist Manipulations and Hate Speech. *J Glob Perspect Soc Cult Dev*, 1(2), 01-03.

Abstract

The mass media has become an arena where hate speech flourishes, not just as isolated incidents, but as a calculated means of manipulation. The research on this issue paints a disturbing picture—one where national and cultural identities are weaponized to foster division and animosity. Researchers have developed several methods to identify and mitigate the spread of such toxic content. Machine learning algorithms work by employing large datasets annotated for hate speech, these algorithms are trained to recognize patterns and features indicative of hateful language. The use of keyword detection involves creating exhaustive lists of terms and phrases frequently associated with hate speech. In the vast landscape of modern discourse, hate speech emerges as a malignant force, manifesting in various pernicious forms. Verbal aggression, intended to incite prejudicial hostility against specific groups based on race, religion, gender, sexual orientation, or other identifiable characteristics, is one such type. In an era dominated by mass media, national and cultural identities are increasingly susceptible to manipulation and hate speech. Recognizing the signs of manipulative tactics—such as sensationalist language, emotional exploitation, and selective reporting—enables individuals to navigate media landscapes with a more judicious mindset.

Keywords: Hate speech, Manipulation, Resistance to Manipulation

Introduction

Hate speech is usually directed against ‘others’ in society. It manifests itself through the ‘othering’ of minority groups, such as racial, ethnic, religious and cultural minorities, women and LGBTQI+ communities. In 1997, the Council of Europe issued a recommendation on hate speech, which defines it as ‘all forms of expression that spread, incite, encourage or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance’. The increased negativism in interpersonal communication largely reflects the trends in the media, which actively broadcast hate speech. But it is one thing to have less contact with news resources; it is quite another to have interpersonal communication that permeates everyday life.

How do you keep a sober mind and not explode during such conversations? How do you avoid getting hurt and hurting others? Communicative competence helps to remain constructive in such communication, namely, the ability to notice forms of verbal aggression in dialogues and respond to them adequately. Increased negativism in interpersonal communication largely reflects the trends in the media, which actively broadcast hate speech. But it is one thing to have less contact with news resources; it is quite another to have interpersonal communication that permeates everyday life. Hate speech is based on the opposition of two concepts: ‘we are the group’ and “they are the group”. The ‘they’ are representatives of a foreign community, whose differences are emphasized in every possible way in fa-

vour of the 'us'. The topic of discussion is often the disadvantage, inferiority of the 'outsiders' and the threats they pose.

Freedom of Speech and Online Platforms

There are many types of media and content on the Internet. Hate speech can be found on various platforms and online forums. As a rule, platforms known for more lenient terms of service or those that do not actively moderate content are likely to have more hate speech. Despite the general ban, hate speech is often found on major social media platforms such as Twitter and Facebook. While these companies prohibit hate speech and strive to remove it, they are not always able to do so. Sometimes this is because the content has not been reported or has not been detected by their software designed to find it. Sometimes it's because the platform's reporting channels are blocked and the content moderators haven't had time to get to it yet. Sometimes questionable content is part of a private or group discussion that is closed even to company moderators. Sometimes the content reported as hate speech does not fall within the definition of prohibited content, and sometimes moderators make bad decisions because they do not fully understand or follow company policy. Most of the public criticism of hate speech has been on the mainstream popular platforms, but many young people are moving to newer online spaces, especially those related to gaming, live streaming and image-sharing platforms.

Hate speech can also be found in message board-style chat rooms or forums, including what are known as 'controversial statements', as well as on more mainstream websites where it can slip through the cracks during gaming sessions, chats or forums that are not constantly moderated. Hate speech can be contained in videos, cartoons, drawings and even photographs. Image and video-based platforms can also contain hate speech. Moderation on these platforms can vary from user-created rules to no moderation at all. Message boards contain user-generated content, and some forums allow any type of content. Message boards are home to many memes and internet hoaxes, which can often contain hate speech.

Hate speech is also found on image and video sharing platforms, some of which are extremely well-known and popular. The larger ones and even some of the smaller ones are moderated forums, but they still have problems with content moderation due to the lack of context associated with the images. The platforms contain a mixture of user-generated content and advertising. Unlike Facebook and Twitter, young people and online influencers are often found here. Finally, hate speech thrives on fringe platforms. New platforms are constantly emerging, and sometimes existing ones are shutting down. Many of these fringe platforms were created in response to content moderation and concerns about 'censorship' on mainstream platforms. Generally, any type of content is allowed, and many users are members of fringe groups or extremist audiences that produce and consume hate speech. These forums may operate in the United States or other countries, and even if the content is illegal, it can be difficult to get the platform to remove it. Why does hate speech occur on these platforms? Millions of posts are created and shared on social media every minute. The volume and scale of online content is so vast that human moderators are unable to manually enforce the platforms' terms of use. AI-based systems are still new and lack the contextual understanding to determine what constitutes

hate speech and what is acceptable political criticism, artistic expression, or unpopular opinion. Even under the best of circumstances, both human moderators and AI-based systems are prone to errors and misinterpretations. Not every hateful comment is easy to spot, understand and condemn. People who moderate content should be familiar with obscure language, memes, and the context of the comment.

Humour, irony and sarcasm should be re-evaluated in terms of their offensiveness. Incitement to hatred or violence is a central element of crimes in all countries. It is therefore useful to clarify what exactly constitutes incitement. It can be a call for someone to do something (you should...), a call to action (we should...), an encouragement (someone should...) or a stirring up of debate (wake up...). Statements containing such phrases can be classified as incitement.

Types of Hate Speech in the Media

1. Calls for violence. These are calls in connection with a specific situation, indicating the object of violence; proclamation of violence as an acceptable means in articles and reports, including direct calls for violence against a group. For example, in 2007, the Odesa-based newspaper *Nashe Delo* published an article entitled 'Kill the best of the goyim', signed by the pseudonym Yuri Streicher. The Jewish community of Odesa saw the pseudonym as a reference to Julius Streicher, the editor-in-chief of the anti-Semitic newspaper *Sturm und Drang*, who was executed by the Nuremberg Tribunal. In 2009, the court sentenced editor Igor Volin-Danilov to a 1.5-year suspended prison sentence for inciting ethnic hatred (Article 161 of the Criminal Code of Ukraine).

2. Direct incitement to discrimination, including in the form of general slogans.

For example, in 2009, during the flu epidemic, the leadership of the Transcarpathian police publicly addressed citizens with the following appeal: 'In connection with the complication of the epidemiological situation in the Transcarpathian region and the increased risk of disease, the management of the Main Department of the Ministry of Internal Affairs of Ukraine in the Transcarpathian region appeals to the population of the region to report every facet of meeting or communicating with foreign citizens, primarily those from South-East Asia and the Middle East'.

3. Veiled calls for violence and discrimination. That is, propaganda of 'positive', historical or contemporary, examples of violence or discrimination; expressions such as 'it would be good to do with ...', 'it's about time ...', etc. In 2016, when announcing the list of events to be held in the city, the mayor of Ivano-Frankivsk, Ruslan Martsinkiv, noted that they should be more actively advertised: 'Our city is much better than Lviv. We are a more patriotic, Christian city. And we will never allow what is happening in Lviv, marches of gays or anyone else,' Martsinkiv was quoted as saying by *Versiya* on 23 March 2016.

4. Forming of a negative image of an ethnic, religious or specific social group. Including those not related to specific eventual accusations, but rather conveyed in broader terms, the general meaning or tone of a text or text fragment. 'Probably, each of us has witnessed or even been a victim of Gypsies wandering around Lviv at least once. Judging by their number, the city cen-

tre, and especially Rynok Square, will soon have every chance of looking like a large gypsy camp, which you will not be able to simply avoid, and yet you will have to throw a hryvnia or two as a 'tie-up fee', - the Vgolos newspaper wrote on 15 July 2014.

5. Justification of historical cases of violence and discrimination. We are talking about expressions like 'the Turks massacred Armenians in 1915 in self-defence'.

6. Publications and statements that question generally accepted historical facts of violence and discrimination. For example, Holocaust denial or the statement that 'Crimean Tatars were deported for siding with Hitler'.

7. Statements about the inferiority (lack of culture, intellectual abilities, inability to work creatively) of a particular ethnic or religious, or a certain social group as such. For example, 'they are only suitable for street sweeping'.

8. Statements about the historical crimes of a particular ethnic or religious group as such. Such as 'they have always resorted to violence', 'they have always conspired against us'.

9. Statements about the criminality of a particular ethnic or religious group. For example, in Ukrainian media reports on the armed conflict in Donbas, Chechens were called 'Chechen killers sent by Kadyrov' (Korrespondent, 02.06.2014). Also, in June 2013, the Express media outlet published a news item entitled 'Illegal migrants are coming to Ukraine from everywhere', creating the impression that all foreigners in Ukraine do not have legal grounds for staying and are criminals. For some time, the chronicles of the Ministry of Internal Affairs of Ukraine published reports with headlines that indicated the ethnic origin of the person suspected of committing an offence or crime next to the words 'criminal', 'thief', 'swindler'. Journalists, using the official website of the Ministry of Internal Affairs as a source of information, repeated hate speech in their text, adding value judgements to the news. For example, in 2015, the Ivano-Frankivsk-based online media outlet Kurs ran the following headline: 'A 25-year-old Donetsk resident tried to sneak food from a Frankivsk supermarket', where the region of origin of the person was linked to the offence, which leads to discrimination against such a category of people as internally displaced persons.

10. Statements about the moral shortcomings of a particular ethnic or religious group. For example, 'Jews are greedy', 'Roma are liars'.

11. Accusations of the negative impact of a particular ethnic, religious or social group on society or the state. 'First black and Arab coffee shops, then streets and neighbourhoods. Assimilation, from which you cannot escape. But in assimilation, something (someone) dominates. In our case, it will not be us,' wrote the Ivano-Frankivsk-based publication Styk in March 2015.

12. Mentioning a certain group or its representatives in a derogatory or offensive context. Including in criminal chronicles or simply by mentioning an ethnonym. For example, during the Euromaidan, television often used such expressions as 'Banderites', 'fascists' and 'extremists' in relation to protesters.

13. Calls to prevent representatives of a particular ethnic or religious group from gaining a foothold in a region, district, city or village. For example, discussions about the inadmissibility of building a mosque in an 'Orthodox city'. 'The construction of a mosque will attract thousands of Muslims here. And if they are already behaving like the masters of the city, then in the future local Ukrainians will simply disappear due to mixed marriages. And if the government cares about the future of the nation, it should draw the appropriate conclusions,' the Grif newspaper wrote on 5 July 2011.

14. Quoting explicitly xenophobic statements and texts without a commentary that distinguishes between the position of the interviewee and the position of the journalist. As well as providing space in the newspaper for explicit xenophobic propaganda without editorial comment or other polemics.

15. Accusing the group of attempts to seize power or territorial expansion (literally, as opposed to calls to prevent them from gaining a foothold in the region). 'In recent years, 7 million able-bodied Ukrainians have left Ukraine in search of work, and about 7 million migrants from Africa, Asia and the Caucasus have come to Ukraine. We are no longer the masters of our land,' the Korrespondent.net website quoted Andriy Ilyenko, coordinator of a nationwide campaign against illegal migration, as saying in November 2011.

16. Denial of citizenship. That is, the reference to citizens as foreigners or 'non-citizens' because of their ethnic identification. 'Insinuations around this year's national selection of the Ukrainian 'Eurosinger' could not be avoided. Yuriy Syrotyuk, a member of the Svoboda party, questioned the appropriateness of choosing Gaitana as Ukraine's representative at this year's Eurovision Song Contest. In his opinion, the African-American Gaitana will not be able to adequately represent Ukraine, as she is not 'a person of our culture', - the Pohlyad newspaper wrote on 9 March 2012.

However, it was easy to check and find out that Gaitana was born in Kyiv and is a citizen of Ukraine, and therefore it is incorrect to call her an 'African-American'. She should be called simply a Ukrainian, or an Afro-Ukrainian, if she identifies herself as such. The Centre for Content Analysis also identified which media outlets most often resort to hate speech. The leader among them was Korrespondent.net (30 per cent of all hate speech messages). Hate speech is also widespread on 24tv.ua, Tsn.ua, Censor.net, Holos Ukrainy, Ukrinform, LigaBusinessInform, Gazeta.ua and Segodnya.ua. Most of the hate speech is directed at Russia, not at Donbas residents. Mostly, hate speech is used in quotes, less often in the journalist's opinion, but in two media outlets it prevails in the author's texts (Korrespondent.net, Liga), as they have powerful blog platforms.