Choosing Between the State and Social Media as the Arbiter of Decency and Truth

McKechnie, D. B.[*]

McKECHNIE, D. B.: Choosing Between the State and Social Media as the Arbiter of Decency and Truth. Právny obzor, *107*, 2024, special issue, pp. 4-20. **https://doi.org/10.31577/pravnyobzor.specialissue.2024.01**

This paper examines the contrasting approaches to regulating hate speech and misinformation in Europe and the U.S., with a focus on the role of social media. Guided by interpretations of Article 10 of the European Convention on Human Rights, European nations maintain latitude to restrict speech harmful to society, including hate speech and misinformation. Conversely, in the U.S., the U.S. Supreme Court's First Amendment jurisprudence places significant burdens on the State's ability to regulate hate speech and misinformation. While hate speech and falsities can cause both individual and social harm, there are deleterious impacts of empowering the State to regulate these ideas. When the State can eliminate hate speech and false ideas from public discourse, society's ability to challenge those ideas is diminished, resulting in indolent public discourse. Moreover, in democratic states, the majority will inevitably define hate speech and truth, and those definitions can change with control of the State. To ensure consistency and legitimacy as control of the State changes, an unfettered marketplace of ideas must be allowed to flourish. The importance of ensuring that unfettered marketplace of ideas has never been more important considering the rise of social media. When the State extends its regulation of hate speech and lies to social media platforms, it exerts control over the locus of the most diverse group of ideas in human history.

Key words: *Free Speech, Hate Speech, Regulation, Democracy, Social Media*

Introduction

Democracies face competing interests in a variety of ways. This is nowhere truer than balancing the right to free speech with the State's interest in protecting society. Certainly, in autocracies, there are no limits on suppressing speech when the State identifies ideas it believes will harm society. Democracies, however, respect the fundamental right of citizens to share and receive ideas, even those ideas that are distasteful, offensive, and dangerous. Whether facilitating the search for truth or enabling self-governance, the protection of free speech is vital to a well-functioning democratic state. Nevertheless, every democratic state places some limitations on the freedom of speech when the harm outweighs the benefits to society. This article explores where to strike the balance between protecting free speech and regulating hate speech and lies.

When examining the European and U.S. approaches to this dilemma, two options emerge. The State can either have more latitude or less latitude to regulate hate speech and lies. When considering the impacts of granting the State greater flexibility to regulate speech, two concerns arise. First, the greater the regulation of harmful speech, the greater the likelihood that individuals and society become lax in their ability to evaluate,

[*] Douglas B. McKechnie, Professor of Law, United States Air Force Academy.

challenge, and reconsider ideas. Second, as with any State decision, the regulation of hate speech and lies will simply reflect decisions by the governing majority. These concerns are particularly problematic regarding the identification and exclusion of hate speech and lies from the public debate on social media. To be sure, hate speech and lies can proliferate on social media. Nevertheless, the same characteristics that allow those ideas to proliferate make social media an unprecedented tool for communication. Social media platforms are the borderless town square where people, including those who dissent from the majority view, can most actively engage in the debate of public issues.

1. Regulating Hate Speech in Europe and the U.S.

The beliefs and ideas expressed on social media platforms are not exceptional when compared with those spoken in the corporeal world. Social media's virtual nature certainly erases the traditional social cues and norms, resulting in less inhibited communication of hateful ideas. Nevertheless, the public discourse in democracies, well before social media, has certainly included hate speech. Whether expressed directly to a person, or about a group of people, hate speech—speech that incites intolerance and disparages people based on their race, religion, sexual orientation, or some other characteristic—has routinely been a source of attempted state regulation.[1] The State's authority to regulate hate speech, however, is not consistent across democracies. The European Court of Human Rights has interpreted the European Convention on Human Rights as providing member states with wider latitude to regulate hate speech. The U.S. Supreme Court, conversely, has interpreted the U.S. Constitution in a way that removes all options for the State to regulate hate speech.

1.1 Hate Speech and the European Court of Human Rights

Hate speech, in the European and U.S. contexts, has a long history. The rise of the Nazi party in Germany in the early Twentieth Century, for example, relied heavily on the denigration of minorities.[2] This national experience with the impacts of dehumanization resulted in the Federal Republic of Germany codifying the recognition of human dignity in its constitution.[3] Indeed, Germany's Basic Law—its constitution—ratified after World War II in 1949, listed human dignity as not only the first protected right but also included it in a "forever clause," shielding it from the amendment process.[4] At the same time the "forever clause" shields human dignity from the amendment process, it also shields the right to free speech.[5] Notably, while there are no limitations on the right to personal

[1] Council of Europe. (1997). Recommendation No. R (97) 20 of the Committee of Ministers to Member States on "Hate Speech".

[2] Staudenmaier, Peter. "Racial Ideology between Fascist Italy and Nazi Germany: Julius Evola and the Aryan Myth, 1933–43." Journal of Contemporary History, vol. 55, no. 3, 2020, pp. 473–491.

[3] Doron Shulztiner & Guy E. Carmi, Human Dignity in National Constitutions: Functions, Promises, and Dangers, The American Journal of Comparative Law 461, 465 (2014).

[4] *Basic Law for the Federal Republic of Germany*, [as amended,] art. 1, 79, May 23, 1949.

[5] Ibid.

dignity, the text of the German Basic Law contains limitations on the right to free speech. More specifically and relevantly, the right to free speech is limited by the "the right to personal honor."[6]

Similarly, the European Convention on Human Rights (ECHR) enshrines the freedoms of speech as a fundamental right, while at the same time placing limits on it. Like the U.S. Constitution's First Amendment, discussed below, Article 10 of the ECHR guarantees that "[e]veryone has the right to freedom of expression."[7] However, unlike the U.S. Constitution's First Amendment, Article 10 provides a list of interests that the State may rely upon to restrict the freedom of speech, including, most relevantly, restrictions "necessary in a democratic society, … public safety, … for the protection of health or morals, [and] for the protection of the reputation or rights of others "[8] Moreover, Article 17 of the ECHR provides that no state, group, or person may interfere with, or destroy, the freedoms set forth in the ECHR.[9] The U.S. Constitution is devoid of similar languge.[10] It is Articles 10 and Article 17, along with the European Court of Human Rights' (ECtHR) interpretation of those articles, that has defined the member State's authority to regulate hate speech.

When interpreting Article 10, the ECtHR has recognized that free speech and open debate are essential to a well-functioning democracy and the development of individuals.[11] Moreover, the ECtHR has determined that in democratic states, pluralism and tolerance demand that even disturbing speech that causes offense must be protected.[12] At the same time, the ECtHR has concluded that restrictions on free speech are permissible so long as they are proportionate to legitimate state interests in regulating speech—interests that include tolerance and protecting individual human dignity.[13] It is through the application of the latter principle that the ECtHR has routinely upheld member States's regulation of hate speech for its incitement to animosity, discrimination, and intolerance. Indeed, the ECtHR has suggested that this sort of speech contributes either nothing, or very little, to the debate on matters of public concern.

In discussing hate speech, the ECtHR has made a distinction between two types of speech based on its content. First, States rarely have the authority to regulate speech that constitutes political expression or speech that contributes to the discussion of matters of public interest.[14] On the other hand, when speech promotes or justifies xenophobia, hate, or other types of intolerance, the ECtHR has held that it is not normally protected by Article 10.[15] In determining whether alleged hate speech is unprotected, the ECtHR will take into

[6] Ibid., art. 5(2).
[7] ECHR, art. 10., U.S. Const. amend. I.
[8] Ibid.
[9] ECHR, art. 17.
[10] U.S. Const.
[11] Handyside v. United Kingdom, App. No. 5493/72, Eur. Ct. H.R. (1976), 49.
[12] Ibid.
[13] Ibid., Savva Terentyev v Russia, App. No. 10692/09, Eur. Ct. H.R. (2018), 65.
[14] Perinçek v. Switzerland, App. No. 27510/08 , Eur. Ct. H.R. (2015), 197.
[15] Ibid., 230.

account the political and social context of the speech.[16] The more volatile the circumstances, the less protection provided by Article 10.[17] The ECtHR will also consider whether the statements could be seen as advocating hatred or intolerance, either directly or indirectly.[18] In this context, if the advocacy is related to hatred or intolerance of an entire group, it is less likely to be protected.[19] Finally, the ECtHR will consider the mode of communication and its corresponding likelihood to lead to harmful consequences.[20] The more public and hostile the mode of communication, the less likely it will be protected.[21] By making this distinction between protected political speech and hate speech that meets the factors above, the ECtHR has determined that the latter has so little social value that it does not contribute to the debate on public issues, and is thus is not protected by Article 10.

Importantly, inciting hatred and intolerance, under the ECtHR's case law, is a distinct social harm and the incitement need not be articulated explicitly or directly. As for its status as a distinct social harm, the ECtHR has determined that incitement to hatred and intolerance may be criminalized even if it is not linked with a call to violence or another criminal act.[22] The harm caused by racist or otherwise defamatory speech about an identifiable group is, by itself, sufficient to remove the statements from Article 10 protection.[23] In addition, the ECtHR has not required members states to limit their criminalization of hate speech to only those provocations that are direct and explicit. Statements that tend to arouse a movement of opinion toward intolerance are beyond the reach of Article 10's protection.[24] Moreover, the statements need not be expressed clearly, with no room for confusion as to their intent.[25] Instead, members states may punish statements that only implicitly urge or call for discrimination.[26] Thus, not only has the ECtHR characterized hate speech as expression that is disconnected from the discussion of matters of public concern, it has provided members states latitude to criminalize hate speech even where it is unconnected to violent rhetoric and is not explicitly directed to inciting intolerance.

1.2. Hate Speech and the U.S. Supreme Court

Hate speech, as defined above, undoubtedly exists as part of the public discourse in the U.S. Like Article 10 of the ECHR, the First Amendment of the U.S. Constitution—Article 10's analog—protects the right to free speech.[27] However, unlike Article 10, the First Amendment's Free Speech Clause states that the State may "make no law … abridging the

[16] Ibid., 205.
[17] Ibid.
[18] Ibid., 206.
[19] Ibid.
[20] Ibid., 207.
[21] Ibid.
[22] Zemmour v. France, App. No. 63539/19, 54, Eur. Ct. H.R. (Dec. 20, 2022), 54
[23] Ibid.
[24] Ibid., 12, 44, and 64.
[25] Ibid., 13, 44, and 64.
[26] Ibid.
[27] U.S. Const. amend. I

freedom of speech, or of the press …" and contains no formal, textual exceptions.[28] Nevertheless, the U.S. Supreme Court has identified a few categories of speech that are unprotected by the First Amendment because of its content; hate speech is not so recognized.[29] Instead, hate speech, and the harm it causes to the individual and society, is contemplated and addressed by a variety of related U.S. Supreme Court cases. For example, the U.S. Supreme Court has erected nearly insurmountable constitutional burdens on the State's authority to punish speech that has the potential to lead to violence or is intended to cause emotional harm.[30] Perhaps more directly related to the concept of the State's authority to regulate hate speech, the U.S. Supreme Court has also rebuffed attempts to restrict speech that arouses resentment or alarm because of race, religion, gender, or the like.[31]

Certainly, hate speech can be disseminated as part of a call for violence or other unlawful activity. Yet the U.S. Supreme Court has taken a more general approach and reviewed the impact of the First Amendment on the State's regulation of any speech that advocates for unlawful activity, including violence.[32] Within the context of speech that may lead to unlawful activity or violence, the U.S. Supreme Court has made the distinction between two classes of speech; abstract discussion and advocacy.[33] The first class consists of the "abstract teaching . . . of the moral propriety or even moral necessity for a resort to force and violence …," and it may not be regulated as it is entirely protected by the First Amendment.[34] The second class is the advocacy of unlawful activity or the use of violence, and it loses First Amendment protection only in the narrowest, most dangerous situations.[35] The State may only punish this speech when it can show the speaker intentionally directed his advocacy to producing imminent lawlessness, and the unlawful activity is, in fact, likely to occur immediately.[36]

What is more, when a speaker advocates for unlawful activity, but his exhortations are not addressed to a specific person or group, or they promote unlawful activity at some unspecified time in the future, that speech is protected by the First Amendment.[37] The U.S. Supreme Court has implied that so long as advocacy does not in fact result in immediate, discernible unlawful activity, it is protected by the First Amendment.[38] Notably, the first case that gave rise to this line of U.S. Supreme Court jurisprudence arose out an attempt prosecute the leader of a Ku Klux Klan group that, among other things, used racist and anti-Semitic language to advocate for Black and Jewish Americans to be sent to Africa and Israel, respectively.[39] However, the U.S. Supreme Court addressed

[28] U.S. Const. amend. I
[29] R.A.V. v. City of St. Paul, 505 U.S. 377 (1992).
[30] Brandenburg v. Ohio, 395 U.S. 444 (1969); Snyder v. Phelps, 562 U.S. 443 (2011).
[31] R.A.V. v. City of St. Paul, 505 U.S. 377 (1992).
[32] Brandenburg v. Ohio, 395 U.S. 444 (1969).
[33] Ibid. at 447-48.
[34] Ibid. at 448.
[35] Ibid. at 447.
[36] Ibid.
[37] Hess v. Indiana, 414 U.S. 105, 108-09 (1973).
[38] NAACP v. Claiborne Hardware Co., 458 U.S. 886, 928 (1982).
[39] Brandenburg v. Ohio, 395 U.S. 444, 447 (1969).

the impact of the First Amendment on the advocacy of unlawful and violent conduct generally, not hate speech in particular. Thus, while the U.S. Supreme Court did not specifically consider the impact of the First Amendment on hate speech, it nevertheless articulated its legal test, conducted its analysis, and found that what would arguably amount to hate speech was indeed protected.

In addition, the U.S. Supreme Court has determined that the First Amendment wholly protects speech intended to inflict emotional distress, so long as the speech is discussing a matter of public concern.[40] Constitutional protection of speech on matters of public concern, the U.S. Supreme Court has reasoned, transcends the interest in protecting self-expression as an end unto itself.[41] Instead, speech on matters of public concern is the essence of self-governance, and must be protected no matter how hurtful, upsetting, offensive, or disagreeable.[42] Though the debate on matters of public concern often involves "vehement, caustic, and sometimes unpleasantly sharp attacks …." it is essential that it remain "uninhibited, robust, and wide-open" to avoid the risk of self-censorship.[43] Notably, in the case that the U.S. Supreme Court determined speech intended to inflict emotional distress was protected by the First Amendment— *Snyder v. Phelps*—the Court reviewed speech that parallels languge that is punishable under the ECtHR's jurisprudence.

In *Snyder v. Phelps*, the U.S. Supreme Court determined that the First Amendment protected protestors who held signs asserting that god hates homosexual people and homosexuality ruins nations.[44] The protestors also held signs that made sweeping allegations about Catholic priests raping boys.[45] Undoubtedly, the First Amendment's protection of the speech in *Snyder vs. Phelps* would extend to similar xenophobic and racist statements casting groups of people in a negative light based on their race, religion, or national origin. Moreover, presumably, the comments protected in *Snyder v. Phelps* would be unprotected by Articles 10 and 17 under the ECtHR's jurisprudence that has allowed the punishment of statements attacking Muslim immigrants and promoting intolerance of homosexual people.[46]

Finally, as perhaps the most closely aligned with an attempt to regulate of hate speech, the U.S. Supreme Court rejected a state's law that punished speech intended to provoke anger on the "basis of race, color, creed, religion or gender …."[47] It was undisputed that the state was attempting to prohibit speech that expressed hatred for those in the listed categories.[48] However, the U.S. Supreme Court determined that the state's classification and punishment of only some speech that expressed hatred, while leaving other speech untouched by the

[40] Snyder v. Phelps, 562 U.S. 443, 451 (2011).
[41] Ibid. at 452.
[42] Ibid. at 456.
[43] N.Y. Times Co. v. Sullivan, 376 U.S. 254, 270 (1964).
[44] Snyder v. Phelps, 562 U.S. 443, 454 (2011).
[45] Ibid.
[46] Le Pen v. France, App. No. 18788/09, Eur. Ct. H.R. (Apr. 20, 2010); Lilliendahl v. Iceland, App. No. 29297/18, Eur. Ct. H.R. (May 12, 2020).
[47] R. A. V. v. St. Paul, 505 U.S. 377, 380 (1992)
[48] Ibid. at 392-93.

prohibition, was constitutionally impermissible.[49] By punishing only the disfavored speech that expressed hatred, and not punishing other speech that expressed hatred, the State was singling out and punishing disfavored speech based on its content and viewpoint.[50]

For example, the law punished hate speech based on one's race, while hate speech based on one's political affiliation was permissible.[51] Furthermore, the law punished hate speech based on one's gender, while speech condemning misogyny would not be punished.[52] These content and viewpoint based regulations, respectively, served to censor ideas the majority found distasteful while allowing speech that was, presumably, acceptable.[53] The State, however, lacks the authority to prohibit disfavored ideas from entering the public debate.[54] However laudable the State's intent, the U.S. Supreme Court determined that the First Amendment prohibited the censorship of speech based on the ideas expressed, particularly when censorship is intended to drive content or a view point out of the public discourse.[55]

2. Regulating Truth and Lies in Europe and the U.S.

Social media platforms provide venues for experts and novices alike to share knowledge and beliefs about a variety of subjects, including matters of public concern. These platforms can be an equalizing force by serving as a seamless method to distribute all manner of statements, from useful facts, to propaganda, to anodyne musings. Within each of these categories, one will inevitably find truths, half-truths, and falsities. Truths and lies, of course, have existed as long as humans have communicated, and long before social media existed. While social media companies did not invent truths and lies, the pervasive, permanently accessible nature of social media allows for truths and lies to be shared with an expansive and instantaneous distribution that was unthinkable before the rise of the internet. With the inevitable impact that lies can have on society, governments have an interest in prohibiting their dissemination, if not punishing it. However, any prohibition or punishment by necessity implicates the regulation of speech. As with hate speech, the European and U.S. understanding of the State's authority to regulate the speech leads to fundamentally different approaches to regulating lies in public discourse.

2.1 Truth and the ECtHR

The ECtHR has extended its interpretation of Articles 10 and 17 to allow member states, under certain conditions, to regulate false speech. Article 10 explicitly permits exceptions to free speech for protecting the reputation or rights of others, suggesting that

[49] Ibid. at 391.
[50] Ibid.
[51] Ibid.
[52] Ibid. at 391-92.
[53] Ibid. at 392.
[54] Ibid. at 387.
[55] Ibid. at 394.

reputational falsehoods fall outside its protections.[56] At the same time, falsities about historical facts are, perhaps, implicitly unprotected particularly when Articles 10 and 17 are read in conjunction.[57] Lastly, with its list of conditions on the freedom speech, Article 10 at least contemplates the possibility that false speech might be regulated by law if it undermines "democratic society, … national security, territorial integrity or public safety," causes "disorder or crime," or is necessary for the "the protection of health or morals …."[58]

Much of the ECtHR's jurisprudence surrounding the regulation of false speech focuses on member States' ability to regulate speech that harms the reputation of others. These claims most commonly take the form of defamation. Defamation, in its most common form, will rarely be protected by Article 10. However, several factors may countenance the protection of speech even if it harms one's reputation. These factors include whether the speech is relevant to a matter of public debate, whether the statement is a declaration of fact or opinion, and the context of the speech.[59] Where there is a confluence of defamatory speech and speech that casts entire groups in a negative light, the ECtHR is particularly likely to provide member states the latitude to punish the speaker.[60] In a similar way, while public officials are often expected to tolerate their reputations, public figures who allege defamation, might further allege that the defamation was intended to incite hatred, as a means to increase the likelihood that a court would find the defamation unprotected under Articles 10 and 17.[61]

In addition to false speech that harms the reputation of others, as a matter of policy and ECtHR jurisprudence, the denial of genocide and crimes against humanity is unprotected speech under the ECHR.[62] Put a different way, Article 10 does not restrain member States from prosecuting speakers for either denying historical events that constituted genocide and crimes against humanity, or asserting facts that are in conflict with historical events that constituted genocide and crimes against humanity.[63] The ECtHR has determined that this sort of speech pursues objects prohibited by Article 17 because it not only seeks to rehabilitate the perpetrators of the crimes, but also disrupts public order and suggests the victims themselves are lying.[64] Moreover, particularly within the context of denying the Jewish Holocaust, the denial of genocide and crimes against humanity constitutes racial defamation and anti-Semitism.[65] Arguably, the

[56] ECHR, art. 10.

[57] ECHR, arts. 10 and 17.

[58] ECHR, art. 10.

[59] Koltay, András. Freedom of Expression as the Foundation of Media Freedom. Wolters Kluwer, 2024, pp. 35–36.

[60] Atamanchuk v. Russia, no. 4493/11, Eur. Ct. H.R., (19 January 2016), 51-52.

[61] *Slovakia's President Sues Ex-Prime Minister for Defamation as Election Tensions Rise.* The Guardian, 14 Sept. 2023, https://www.theguardian.com/world/2023/sep/14/slovakias-president-sues-ex-prime-minister-for-defamation-as-election-tensions-rise.

[62] Framework Decision of the Counsil of European Union 2008 / 9413 / JHA; Perinçek v. Switzerland, App. No. 27510/08, Eur. Ct. H.R. (2015), 197.

[63] Garaudy v. France, App. No. 65831/01, Eur. Ct. H.R. (2003), p. 29.

[64] Ibid.

[65] Ibid.

ECtHR has suggested that there may be other historical facts that member States could prohibit speakers from contesting, in addition to the Holocaust.[66] The ECtHR has indicated the Holocaust was only an example, and it seemingly left open the possibility that debating the truth of other clearly established historical facts could legitimately be characterized by a member State as something apart from an effort to seek truth.[67]

Perhaps, however, the most consequential declarations regarding the right of member states to regulate falsities are not found in the ECtHR's jurisprudence. Instead, they can be found in the European Union's Digital Services Act (DSA).[68] The DSA, enacted in 2022, is an effort by the European Union to ensure a "a safe, predictable and trusted online environment" and requires certain online platforms to prohibit harmful activities.[69] These harmful activities include, among other things, the spreading of disinformation.[70] The spreading of disinformation, the DSA declares, poses societal risks including negative effects of public health and individuals' physical and mental well-being, as well as gender based violence.[71] The DSA also suggests disinformation can pose risks to public security, civil discourse, political participation, electoral processes, and equality.[72] The DSA does not define the term "disinformation."[73] However, the Act recognizes that member States were addressing, through their own legislation, the distribution of disinformation online and posited that a supranational approach to solving the problem was needed.[74] Thus, the DSA at a minimum contemplates, if not directly declares, that members States have the right to regulate some false information, beyond the defamatory speech and Holocaust denials discussed above.

2.2 Truth and the U.S. Supreme Court

In the early years of the U.S. Supreme Court's evaluation of the First Amendment's meaning, the Court penned what would become one of its most recognized axioms. When exploring the perimeter of constitutional protections afforded by the First Amendment, the U.S. Supreme Court posited that "[t]he most stringent protection of free speech would not protect a man in falsely shouting fire in a theatre and causing a panic.[75] Over the course of the next 100 years, however, the U.S. Supreme Court's understanding of the First Amendment and its application to false speech took on a more nuanced approach. Much of the U.S. Supreme Court's discussion of the right to free speech and untruth has arisen within the context of defamatory falsities.[76] The U.S. Supreme Court has recognized the

[66] Ibid.
[67] Ibid.
[68] European Parliament and Council of the European Union. Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act) and Amending Directive 2000/31/EC. (October 19, 2022).
[69] Ibid.
[70] Ibid. 9.
[71] Ibid. 83.
[72] Ibid. 82 and 95.
[73] See generally ibid.
[74] Ibid. 2.
[75] Schenck v. United States, 249 U.S. 47, 52 (1919).
[76] N.Y. Times Co. v. Sullivan, 376 U.S. 254, 269 (1964).

unprotected nature of falsehoods that harm another's reputation.[77] Yet when alleged defamation is directed at public officials and figures, the U.S. Supreme Court has struck the balance, overwhelmingly, in favor of free speech.[78] However, defamatory falsehoods are not the only untruths the U.S. Supreme Court examined in the context of free speech and First Amendment protections. The U.S. Supreme Court has also suggested that false statements are not, *ipso facto*, beyond the First Amendment's protection.[79]

The U.S. Supreme Court's most developed understanding of the First Amendment's impact on false statements has arisen out of questions surrounding defamation. In the U.S., with few exceptions, a defendant can raise truth as an absolute defense to a defamation allegation.[80] Thus, only false statements can give rise to defamation claims. Still, not all false and defamatory statements give rise to a legal claim. Instead, the U.S. Supreme Court has determined that public officials and public figures must prove, when suing for defamation, that the speaker acted with actual malice.[81] That is to say, public officials and public figures must prove that the speaker knew that the defamatory statement was false or recklessly disregarded whether it was false.[82] It is these decisions, and their progeny, that have given the U.S. Supreme Court the widest opportunity to explore the First Amendment's relationship with false statements.

Specifically, when considering the possibility that a high burden on public officials' and public figures' defamation claims will result in false statements going unpunished, the U.S. Supreme Court has expressed an oft repeated premise—the First Amendment was intended to "assure unfettered interchange of ideas for the bringing about of political and social changes desired by the people."[83] Indeed, First Amendment protection applies to all speech, and does not depend upon its "truth, popularity, or social utility …."[84] This understanding of the First Amendment and the constitutional protection of free speech arises from a commitment to protecting an "uninhibited, robust, and wide-open" public debate that may, at times, include false statements.[85] That is not to say that false statements are inherently valuable and worthy of constitutional protection, but instead that their constitutional protection is necessary for other reasons. The constitutional protection of false statements is a recognition that they are unavoidable in an unrestricted public debate and must be shielded by a constitution to ensure the freedom of expression is not chilled by the threat of legal punishment.[86] It is the concern for speakers' self-censorship that necessitates the protection of "some falsehoods in order to protect speech that matters."[87]

[77] Counterman v. Colorado, 600 U.S. 66, 73 (2023).
[78] N.Y. Times Co. v. Sullivan, 376 U.S. 254, 279-80 (1964).
[79] United States v. Alvarez, 567 U.S. 709, 718 (2012).
[80] 50 Am Jur 2d Libel and Slander § 252.
[81] N.Y. Times Co. v. Sullivan, 376 U.S. 254, 279-80 (1964).
[82] Ibid. at 281.
[83] Ibid. at 269.
[84] Ibid. at 271.
[85] Ibid. at 270.
[86] Ibid. at 271.
[87] Gertz v. Robert Welch, 418 U.S. 323, 341 (1974).

While the U.S. Supreme Court's discussion of false speech and its constitutional implications arise primarily out of the Court's exploration of defamation, it has also addressed the State's direct attempt to punish intentional, indisputable lying.[88] Regarding the First Amendment and knowingly false statements, some lies, like those which defame or facilitate fraudulent, ill-gotten economic benefit, are not constitutionally protected.[89] The untrue nature of the statements, however, is not dispositive.[90] Lies alone do not lose constitutional protection simply because they are controverted by facts.[91] In fact, even when the speaker knows his statement is a lie, the First Amendment provides protection from the State's regulation.[92] It is a legally cognizable harm, coupled with the lie, that allows the State to regulate false statements.[93]

To permit the State to punish a lie, without more evidence of harm, would allow it to produce a list of topics about which false statements are punishable.[94] And that list of potential topics could be infinite.[95] The idea that the State may legally determine truth and punish falsity conflicts with the First Amendment's commitment to protecting speech and a properly functioning free society.[96] In a free society, the State lacks the authority to incarcerate liars; instead, "[t]he remedy for speech that is false is speech that is true."[97] Indeed, the suppression of false speech subverts the very purpose of free speech and relation to true ideas. First, it prohibits the truth from being vigorously examined, challenged, and verified in an open marketplace of ideas.[98] Second, if the State is empowered to arrange the terms of public discourse by limiting untruths, it only serves to make the exposure of those lies more difficult.[99] The First Amendment renders the State powerless to punish the expression of ideas related to matters of public concern, even where an overwhelming majority of citizens believe the ideas to be false and likely to lead to intolerable outcomes.[100]

3. The Dangers of Empowering the State to Regulate Hate Speech and Lies

The European and U.S. approaches to regulating hate speech and misinformation both pursue laudable goals. The European approach seeks to strike a balance. While Article 10 acknowledges the individual right to express and receive ideas, at the same

[88] United States v. Alvarez, 567 U.S. 709, 714-15 (2012).
[89] Ibid. at 719.
[90] Ibid.
[91] Ibid. at 719.
[92] Ibid. at 729-730.
[93] Ibid. at 719.
[94] Ibid. at 723.
[95] Ibid.
[96] Ibid. at 727.
[97] Ibid.
[98] Thomas I. Emerson, Toward a General Theory of the First Amendment, 72 Yale L.J. 877, 881-82 (1963).
[99] Ibid.
[100] Whitney v. California, 274 U.S. 357, 374 (1927) (Brandeis, J., concurring).

time, it recognizes that the State has the authority to protect the public good.[101] On the other hand, the First Amendment's protection of free speech, with no formal limitations, has been interpreted by the U.S. Supreme Court to elevate free speech above many other values and interests. These competing approaches are, seemingly, incompatible with in a cohesive jurisprudence. Reduced to their respective essence, either the State retains the presumptive authority to regulate hate speech and lies, or it does not. It cannot be gainsaid that hate speech and misinformation can cause harm to individuals and society. Nevertheless, the U.S. approach encourages an energetic and productive public discourse and spurns the State's capacity to declare certainty on matters of public concern. This is particularly true within the context of the State's attempt to regulate social media.

3.1. Where the State Regulates Speech Social Indolence Follows

Where the State is empowered to regulate hate speech and lies, its intervention breeds an indolent polity where public discourse is managed by the State instead of society. Removing hateful languge and lies from public discourse impacts both prospective speakers and listeners, and sanitizes the debate about matters of public concern. Knowing that their speech can open them to prosecution, prospective speakers who would otherwise participate in the public discourse and express intolerance or lies will refrain from speaking. This undoubtedly results in the otherwise harmful speech from entering public discourse. Chastened by the threat of punishment, however, those prospective speakers will be more likely to fully withdraw from public debate. The withdrawn, silenced speaker evades—or is denied—the opportunity of being confronted with contrary ideas that contest and challenge his intolerant or mistaken beliefs. There is no need to defend, and possibly reconsider, intolerant, racist beliefs in a vigorous, adversarial debate. The views the State would purge instead remain stagnant—unchanged in those who adhere to those views.

Perhaps more importantly, society loses the opportunity and benefit of considering unconventional, even detestable, ideas the speaker may have shared. The removal of viewpoints and ideas from the realm of acceptable topics for public debate stunts listeners' judgment and lulls them into a reliance on a paternalistic state. When the State identifies groups about which intolerant speech is impermissible, it inevitably leaves some groups unprotected and thus vulnerable to ostensibly permissible hate speech. These lines drawn by the State will simply reflect the majority's preferences—some groups may be subjected to hate speech that the State allows because the majority does not consider the group worthy of protection. Moreover, when the State identifies truth and prohibits the expression of the converse, society sees facts proclaimed by the State with reconsideration and debate no longer possible, welcomed, or seemingly necessary. Through confrontation with intolerant speech and lies, individuals within society evaluate a variety of competing views. They are forced to think critically about the various options presented by competing sides and, in doing so, develop the intellectual maturity necessary

[101] ECHR, art. 10.

in self-governing, free societies.[102] Indeed, through learning that fellow citizens share intolerant or false views, people are motivated to seek out and associate with like-minded individuals and participate in civil society.

These concerns about the impacts of interventionism are heightened by the prospect of regulating hate speech and falsities on social media. The internet and social media represent the fastest and most accessible platforms for sharing and receiving information in human history. Users are exposed to a vast array of topics and ideas, including those that challenge their existing beliefs and introduce new perspectives. Still other users may encounter topics and ideas they never knew existed. As a result, social media has become the most effective way to access a multitude of disparate voices, and it is only through unfettered access to disparate voices that society can effectuate—or gather opposition to—political and social change.[103] It is not for the State to play a role in regulating the virtual marketplace of ideas presented by social media.

To be sure, racist, xenophobic, and false ideas flourish on social media. However, the State must be powerless to coercively eliminate these ideas.[104] While social media users may believe racist, xenophobic, and false ideas the State hopes to remove from the public discourse, the possibility that the ideas could or might lead to harmful behavior in the future is not sufficient reason for prohibiting them.[105] Self-governance demands that social media, perhaps the most revolutionary means by which to share ideas, be unfettered so long as there is even a moment to "expose through discussion … falsehood and fallacies …."[106] The active political and social participation necessary to expose falsehoods and fallacies will atrophy if offending speech is silenced. It is up to those within the virtual marketplace to test, affirm, and refute ideas. Indeed, the values inherent in a robust, wide-open marketplace of ideas are no less important because one's voice resonates farther and faster than it could have when pen and paper were one's only means of communication.[107]

3.2. Majority Rule and Regulating Speech

While majority rule is not the only component of a democratic state, it is a foundational principle for decision-making in democracies. Whether through a simple or supermajority, the policies adopted in a well-functioning democracy generally reflect the interests and preferences of the majority. Key political decisions—such as tax policy and foreign relations—are shaped by the electoral process and the selection of representatives to govern on behalf of the electorate. As citizens evaluate the successes and failures of their government, they exercise their power to retain or replace elected officials. Consequently,

[102] 303 Creative LLC v. Elenis, 600 U.S. 570, 584 (2023) ("By allowing all views to flourish … we may test and improve our own thinking both as individuals and as a Nation.").

[103] N.Y. Times Co. v. Sullivan, 376 U.S. 254, 269 (1964).

[104] 303 Creative LLC v. Elenis, 600 U.S. 570, 598 (2023).

[105] Ashcroft v. Free Speech Coal., 535 U.S. 234, 253 (2002).

[106] Whitney v. California, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring).

[107] 303 Creative LLC v. Elenis, 600 U.S. 570, 587 (2023).

as people's interests and priorities evolve, so too do their representatives, the laws, and the State's policies. It is these natural and necessary democratic fluctuations that reveal why the State is unfit to regulate hate speech and lies.

When regulating hate speech, the State must make distinctions among the population. More specifically, the State categorizes the population, for example, by sexual orientation or race, and establishes prohibitions on hate speech when speech advocates for intolerance for those in a protected category. While undoubtedly a worthy goal, the fundamentals of a democratic state suggest that hate speech prohibition could be used to silence minority opinions, even if abhorrent, about matters of public concern regarding the protected category of people. What is equally possible is that what may be considered a protected category of people by one iteration of government will not be considered a protected category by the next government. In this way, hate speech prosecutions can simply become a tool for changing majorities to target detractors. Furthermore, there will inevitably be minorities within society that never gain enough political power to invoke hate speech protection from the majority. Although these minorities might be entitled to hate speech protection, their small size or lack of influence can lead the State to overlook them as a group deserving protection. Democratic states are adept at balancing the costs and benefits of political choices and changing course once the electorate has assessed their decisions. Determining who must be protected from contemptuous speech, however, is not within the State's remit "simply because society finds the idea[s] [themselves] offensive …."[108] Choices regarding how to battle hate speech are better left to the marketplace of ideas.

If a democratic state lacks the competence to determine when hate speech should be prosecuted, it also lacks the competence to declare truth and punish its opposite. The very concept of truth itself, particularly regarding matters of public concern, does not lend itself to discovery by the State. What the State declares to be true in a democracy is simply a reflection of what the current majority perceives to be true. Like the reorientation of protection for those subjected to hate speech, as democratic control of the State changes so too can the State's declaration of truth. To be sure, it is in direct opposition to the very concept of truth that it can change as democratic control of the government changes. For example, while France's Gayssot Act prohibits denial of the Holocaust, a majority could rescind the Act.[109] In revoking the Act, the facts of the Holocaust would not become less true. Similarly, while one EU member state may prohibit denial of the Holocaust and another dose not, these discrepancies do nothing to subvert the historical accuracy of the Holocaust. Declaring and imposing truth by majority-rule is not validation of that truth. Instead, "the best test of truth is the power of the thought to get itself accepted in the competition of the market …."[110]

With the power of the State to support it, the majority can easily promote its version of truth. This is particularly so where the State exerts significant control over major

[108] Snyder v. Phelps, 562 U.S. 443, 458 (2011).

[109] French Republic. Loi n° 90-615 du 13 juillet 1990. Journal Officiel de la République Française, 13 July 1990.

[110] Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

public media outlets.[111] Through its oversight and management of public television and radio, legislative decisions can empower the State to orchestrate public media. Public media will then advocate for the majority's version of the truth and eliminate coverage of the opposite, if not refuting it entirely. Political maneuvering to empower the State to use public media to propagate its truths, while at the same time using its police powers to investigate those who controvert that truth, ensures that the debate on matters of public concern is, in fact, no debate at all. Free speech protections are rarely needed for the majority's understanding of truth. Instead, free speech protections are most needed for minority viewpoints that oppose the majority's truth.

There is a particular concern regarding proposals to regulate false statements on social media. Undoubtedly, social media companies can make editorial decisions, elevate speakers, and remove users. Through content moderation, they have significant influence over the information their users see. Social media platforms are, nonetheless, online communities for sharing ideas where users are exposed to an almost limitless variety of ideas and viewpoints. Moreover, social media transcends regional and national boundaries. Where the State exerts control over those ideas that cannot be challenged within society, social media platforms offer users an opportunity to seek contrary views and test the accuracy of the State's version of truth. To be sure, the State is free to choose the viewpoints it wishes to express, even on social media.[112] However, collectively, social media platforms must remain sources to access not only the State's understanding of truth, but also the dissenter's understanding of truth. While false statements will inevitably exist in an unregulated social media environment, "[s]ome degree of abuse is inseparable from the proper use of everything; and in no instance is this more true, than in [the publication of ideas]."[113] The alternative, however, is even more troubling—allowing the State to determine when an idea is true and then pressuring speakers to withhold their opposing viewpoints for fear of punishment.[114]

Conclusion

Hate speech and lies are undoubtedly harmful to individuals and society. Certainly, the State could endeavor to root out and punish every incitement to intolerance, every falsehood. This Ministry of Decency and Truth, however, would present its own problem. It would undermine the fundamental right to formulate beliefs, share ideas, and debate the merits of even the most abhorrent views, without interference from the State. Because it is impossible to cleanse society of intolerance and false ideas, it is better that those ideas be shared publicly where they can be evaluated and disputed. Indeed, though certainly imperfect, social media platforms have become the most effective and efficient

[111] Żuk, Paweł. "One Leader, One Party, One Truth: Public Television Under the Rule of the Populist Right in Poland in the Pre-Election Period in 2019." Javnost - The Public, vol. 27, no. 3, 2020, pp. 287–307.

[112] Rust v. Sullivan, 500 U.S. 173, 194 (1991).

[113] Madison, James. 1799. Report on the Virginia Resolutions. Founders Online, National Archives. https://founders.archives.gov/documents/Madison/01-17-02-0202 (accessed December 8, 2024).

[114] Moody v. NetChoice, LLC, 144 S. Ct. 2383, 2402-2403 (2024).

means of engaging with the widest possible audience and most diverse set of ideas. It is only through this process of engaging even the most detestable ideas that individuals and democracies remain involved not only in the scientific process of evaluating and reevaluating ideas, but ultimately, the difficult task of self-governance.

Bibliography

Abrams v. United States, 250 U.S. 616 (1919) (Holmes, J., dissenting).
Ashcroft v. Free Speech Coal., 535 U.S. 234, 253 (2002).
Atamanchuk v. Russia, no. 4493/11, Eur. Ct. H.R., (January 91, 2016).
Basic Law for the Federal Republic of Germany, [as amended,] May 23, 1949.
Brandenburg v. Ohio, 395 U.S. 444 (1969).
Council of Europe. (1997). Recommendation No. R (97) 20 of the Committee of Ministers to Member States on "Hate Speech".
Counterman v. Colorado, 600 U.S. 66 (2023).
Doron Shulztiner & Guy E. Carmi, Human Dignity in National Constitutions: Functions, Promises, and Dangers, The American Journal of Comparative Law 461 (2014).
ECHR, art. 10.
ECHR, art. 17.
European Parliament and Council of the European Union. Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act) and Amending Directive 2000/31/EC. (October 19, 2022).
Framework Decision of the Counsil of European Union 2008 / 9413 / JHA.
French Republic. Loi n° 90-615 du 13 juillet 1990. Journal Officiel de la République Française, 13 July 1990.
Garaudy v. France, App. No. 65831/01, Eur. Ct. H.R. (2003).
Gertz v. Robert Welch, 418 U.S. 323 (1974).
Handyside v. United Kingdom, App. No. 5493/72, Eur. Ct. H.R. (1976).
Hess v. Indiana, 414 U.S. 105 (1973).
Koltay, András. Freedom of Expression as the Foundation of Media Freedom. Wolters Kluwer, 2024.
Le Pen v. France, App. No. 18788/09, Eur. Ct. H.R. (Apr. 20, 2010).
Lilliendahl v. Iceland, App. No. 29297/18, Eur. Ct. H.R. (May 12, 2020).
Madison, James. 1799. Report on the Virginia Resolutions. Founders Online, National Archives. https://founders.archives.gov/documents/Madison/01-17-02-0202 (accessed December 8, 2024).
Moody v. NetChoice, LLC, 144 S. Ct. 2383 (2024).
NAACP v. Claiborne Hardware Co., 458 U.S. 886 (1982).
N.Y. Times Co. v. Sullivan, 376 U.S. 254 (1964).
Perinçek v. Switzerland, App. No. 27510/08, Eur. Ct. H.R. (2015).
R.A.V. v. City of St. Paul, 505 U.S. 377 (1992).
Rust v. Sullivan, 500 U.S. 173 (1991).
Savva Terentyev v Russia, App. No. 10692/09, Eur. Ct. H.R. (2018).
Schenck v. United States, 249 U.S. 47 (1919).
Slovakia's President Sues Ex-Prime Minister for Defamation as Election Tensions Rise. The Guardian, 14 Sept. 2023, https://www.theguardian.com/world/2023/sep/14/slovakias-president-sues-ex-prime-minister-for-defamation-as-election-tensions-rise.
Snyder v. Phelps, 562 U.S. 443 (2011).
Staudenmaier, Peter. "Racial Ideology between Fascist Italy and Nazi Germany: Julius Evola and the Aryan Myth, 1933–43." Journal of Contemporary History, vol. 55, no. 3, 2020.
Thomas I. Emerson, Toward a General Theory of the First Amendment, 72 Yale L.J. 877 (1963).
U.S. Const.
U.S. Const. amend. I.

United States v. Alvarez, 567 U.S. 709 (2012).
Whitney v. California, 274 U.S. 357 (1927) (Brandeis, J., concurring).
Zemmour v. France, App. No. 63539/19, Eur. Ct. H.R. (Dec. 20, 2022).
Żuk, Paweł. "One Leader, One Party, One Truth: Public Television Under the Rule of the Populist Right in Poland in the Pre-Election Period in 2019." Javnost - The Public, vol. 27, no. 3, 2020.
50 Am Jur 2d Libel and Slander § 252.
303 Creative LLC v. Elenis, 600 U.S. 570 (2023).