Hwang, L. (2025). Regulation of Misinformation in the Digital Age: First Amendment Rights and Government's Role in Social Media. *Critical Debates in Humanities, Science and Global Justice*, *5*(2), 103–113.

RESEARCH ARTICLES

Regulation of Misinformation in the Digital Age: First Amendment Rights and Government's Role in Social Media

Luke Hwang[a]

Critical Debates in Humanities, Science and Global Justice

Vol. 5, Issue 2, 2025

Misinformation has always existed, but the rise of social media has magnified its impact, threatening societal trust, relationships, and democratic systems. While the First Amendment of the United States Constitution protects free speech, including some forms of false expression, addressing misinformation without infringing on these constitutional rights remains a significant challenge. This study examines key precedents – United States v. Alvarez, Bantam Books v. Sullivan, National Rifle Ass'n v. Vullo, and Murthy v. Missouri – to explore the boundaries of government involvement with social media platforms in content moderation. While Alvarez highlighted the application of strict scrutiny to content-based regulations, Bantam Books and Vullo established that coercion, whether direct or indirect, violates First Amendment protections. Murthy further highlighted the difficulties plaintiffs face in proving government coercion and establishing legal standing requirements. This study proposes a new legal framework to address misinformation while safeguarding free speech. Key factors include prioritizing less restrictive alternatives to censorship, establishing clear boundaries on government involvement, ensuring transparency in content moderation, and adapting to emerging challenges from evolving technologies such as artificial intelligence (AI) and deepfakes. These measures aim to balance the need to address the harm caused by misinformation with upholding fundamental democratic values and constitutional rights. Combating misinformation requires a collective effort; while the government has a duty to protect free speech and avoid coercion, individuals, social media platforms, and government shall collaborate to foster an informed and open media environment to uphold the principles of free speech and democratic integrity.

INTRODUCTION

Fake news, along with other forms of misinformation, has been a pervasive part of human communication. Comparable to legitimate news, fake news has historically been widespread and has posed consistent risks to society. An early example is the 1835 "Great Moon Hoax" in the New York City's "The Sun," which falsely described lunar civilizations of man-bats, unicorns, and two-legged beavers (*Would You Have Believed the Great Moon Hoax?*, n.d.). The advent of the internet and social media has amplified the rapid

a Luke Hwang is a student at Korea International School (Pangyo)

dissemination of fake news, notably following the 2016 United States presidential election, where Collins Dictionary's Word of the Year was recognized as "fake news" (Grinberg et al., 2019; Meza, 2017). Fake news has become an increasingly important issue in today's media environment.

Experts argue that we are living in a post-truth era, where narratives can shape public opinion despite contradicting evidence (Moravčíková, 2020). For example, Dr. Carrie Madej's viral claim on YouTube in 2020, alleging that the COVID-19 vaccines were "designed to make us into genetically modified organisms," illustrated the challenge of combating misinformation during the pandemic (Carmichael & F., 2020). The spread of such claims has caused confusion and distrust, exacerbated by limited understanding and information about the new virus.

The surge of misinformation highlighted the essential need for accurate, evidence-based, and timely information during global health emergencies. Ensuring the open and free exchange of reliable information was vital for raising awareness about health risks and effectively managing them by fostering trust in and compliance with public health guidelines. Access to a variety of trustworthy information sources and perspectives was crucial for people to critically evaluate information, dispel rumors, and effectively combat misinformation.

MATERIALS AND METHODS

The First Amendment and Misinformation

The First Amendment of the United States Constitution prohibits government suppression or endorsement of specific ideas or messages (U.S. Const. amend. I). The United States Supreme Court (hereinafter, the "Supreme Court") has broadly interpreted this constitutional guarantee of free speech across all levels of governance and forms of communicative expression (*Chaplinsky v. New Hampshire*, 1942). In its decisions, the Supreme Court has held that prohibiting false speech would "chill" more valuable speech, potentially leading people to self-censor to avoid legal repercussions (*New York Times Co. v. Sullivan*, 1964). Therefore, the First Amendment ensures a "breathing space" for false statements and hyperbole that are "inevitable in free debate," implying that the government regulation of false ideas and false factual statements is constitutionally limited (*Sullivan*). Consequently, misinformation and false expression have generally been protected from government censorship under the First Amendment, unless they fall under specific exceptions to constitutional protection.

The Supreme Court has developed several approaches to assess the constitutionality of regulations on speech and to determine when the government may regulate speech. Over time, the Supreme Court has narrowed its exceptions to constitutional protection, generally upholding regulations on false speech only within specific categories such as defamation, fraud, or false commercial speech (Brennon, 2022). However, the regulation of other types of false speech, including misinformation, has long been vague.

Methodology

This study examined the relationship between government censorship of misinformation and the First Amendment through a comprehensive review of case laws and legal documents. This study involved an in-depth review of significant Supreme Court cases, federal and state court decisions, and relevant statutes. Primary sources were identified through legal databases such as Justia U.S. Supreme Court Center and Justia Law; secondary sources, including law review articles, legal commentaries, and academic publications, were analyzed to provide context and interpretive perspectives. By critically evaluating how courts have balanced the government's interest in regulating misinformation with the First Amendment right to free speech, this study proposed a new framework for analyzing the complex interplay between free speech and governmental regulation in the context of misinformation.

RESULTS

Regulation on Content-Based Speech: Strict Scrutiny and Coercion Standard

In *Police Department of Chicago v. Mosely*, which addressed the constitutionality of a city's prohibition of non-union picketing outside a school building, the Supreme Court held that the First Amendment prohibits the government from restricting expression based on its message, ideas, subject matter, or content (*Police Dept. of Chicago v. Mosley*, 1972). Laws regulating speech based on its subject matter, topic, or viewpoint, or aiming to suppress or promote a particular message are considered content-based under the First Amendment, subject to strict scrutiny (Legal Information Institute, n.d.). Under strict scrutiny, such laws are presumptively unconstitutional unless the government can demonstrate that they are the "least restrictive means" of advancing a "compelling" governmental interest (*Sable Communications v. FCC*, 1989).

This strict scrutiny standard was highlighted in *United States v. Alvarez*, where the Supreme Court held the Stolen Valor Act, a federal law that criminalized false statements about receiving military decorations or medals, constitutional (*United States v. Alvarez*, 2012). The Supreme Court, however, was divided on the appropriate level of scrutiny for regulating false speech in this context (*Alvarez*). The majority opinion applied strict scrutiny, treating the federal law as a content-based regulation, while two (2) justices in a concurring opinion argued that an intermediate level of scrutiny would suffice if the law targeted "a subset of lies where specific harm is more likely to occur" (*Alvarez*).

In *Bantam Books, Inc. v. Sullivan*, the Supreme Court underscored the distinction between coercive and merely persuasive speech, establishing a principle that has gained new relevance in the social media era (*Bantam Books, Inc. v. Sullivan*, 1963). This principle bars government officials from coercing social media platforms to censor the speech of their users, similar to the First

Amendment's prohibition of the government coercing booksellers to restrict the circulation of objectionable materials (*Bantam Books*). In *Bantam Books,* the Supreme Court held that a state commission, lacking formal regulatory authority, violated the First Amendment when it "deliberately set about to achieve the suppression of publication" through "informal sanctions," including "threat of invoking legal sanctions and other means of coercion, persuasion, and intimidation" (*Bantam Books*).

In *National Rifle Ass'n of America v. Vullo*, the Supreme Court addressed a case where the New York state financial regulator pressured financial services companies to sever ties with clients such as the National Rifle Association of America (NRA). The Supreme Court held that "*Bantam Books* stands for the principle that a government official cannot directly or indirectly coerce a private party to punish or suppress disfavored speech on her behalf," and that to claim that "the government violated the First Amendment through coercion of a third party, a plaintiff must plausibly allege conduct that...could be reasonably understood to convey a threat of adverse government action in order to punish or suppress speech" (*National Rifle Association of America v. Vullo*, 2024). The *Vullo* decision was grounded in the precedent set by *Bantam Books*, which established that the First Amendment protections against government censorship cannot be circumvented by using a private intermediary by utilizing three (3) factors to distinguish between permissible persuasion and unconstitutional coercion: "(1) the authority of the government officials who are alleged to have engaged in coercion; (2) the nature of statements made by those officials; and (3) the reactions of the third party alleged to have been coerced" (*Vullo*).

On June 26, 2024, the Supreme Court delivered its decision in *Murthy v. Missouri*, a case where government officers were alleged of coordinating efforts to influence the moderation of online content in conflict with the government's views, particularly relating to the COVID-19 pandemic and other controversial topics. The Supreme Court ruled that the plaintiffs did not have the legal standing required to seek an injunction preventing government officers or agencies from engaging with social media platforms about content moderation. The Supreme Court emphasized that, in order to seek an injunction, plaintiffs are required to demonstrate a substantial risk of future harm traceable to government defendant (*Murthy v. Missouri*, 2024).

In this decision, the Supreme Court solely addressed the issue of standing without considering the merits of the plaintiffs' claims, holding that: (1) the plaintiffs had failed to show "any discrete instance of content moderation caused by any of the challenged governmental actions"; (2) the plaintiffs could not establish a link between past restrictions on their speech by the platforms and the government defendants' communications, thus failing to prove that future censorship was likely; and (3) the plaintiffs failed to identify a cognizable injury under the First Amendment, as this requires a concrete and specific connection between the listener and the speaker (*Murthy*).

DISCUSSION

Competing Interests in Combating Misinformation

Under longstanding content moderation practices, social media platforms have implemented various measures to limit specific types of speech. These actions include adding warning labels to posts, removing posts, reducing the visibility of certain content, and suspending or banning users who repeatedly violate platform policies (Harrell & Jones, 2024). This approach was particularly evident during the 2020 U.S. presidential election when platforms like Facebook and Twitter (now X) enforced strict rules to mitigate the spread of fake news. However, during the 2024 U.S. presidential election, social media companies adopted a more lenient approach on content moderation compared to that of 2020, concerning researchers about the potential spread of misinformation and its risks, including violence. Over the years, X has become more tolerant of election-related falsehoods and has amplified pro-Trump narratives. On X, some users who regularly share election-related misinformation, AI-generated images, and baseless conspiracy theories have reportedly earned thousands of dollars from the platform (Spring, 2024). This occurred in the absence of clear guidelines or regulations to demonetize or suspend accounts that spread false information (Spring). This shift toward a more permissive approach on X has also influenced other platforms, such as YouTube, to ease their policies around misinformation (Allyn, 2024).

The Supreme Court's decision in *Bantam Books* to delineate the boundary at coercion appears reasonable, particularly in recognizing that government officials can appropriately engage with social media platforms about disseminated content. Such involvement can involve explaining how editorial decisions impact public health or safety. However, the court's role should be to closely scrutinize these interactions by weighing the competing interests at stake and assessing whether officials have attempted to coerce platforms into censoring content.

Misinformation has existed throughout history, but the COVID-19 pandemic highlighted the critical importance of safeguarding free speech to ensure that accurate information prevails over misinformation. However, instead of enhancing transparency in managing the surge of information, many countries curtailed freedom of speech when it was most needed (*Silenced and Misinformed: Freedom of Expression in Danger during Covid-19*, 2021). Governments' responses illustrated how global emergencies and societal unrest can be exploited to consolidate power, enact legislation undermining human rights, and suppress dissenting voices (Amnesty International).

Further complicating this issue, Meta CEO Mark Zuckerberg revealed in a letter to Congress that his company was pressured by the government to limit unfavorable speech online. Zuckerberg's letter came after *Murthy*, originally filed as *Missouri v. Biden*. This case involved state Attorney Generals from

Missouri and Louisiana seeking to prevent federal agents from unconstitutionally coercing social media platforms to remove or restrict protected First Amendment speech (McCullogh, 2024). In his letter, Zuckerberg disclosed that in 2021, "senior officials from the Biden Administration, including the White House, repeatedly pressured our teams for months to censor certain COVID-19 content, including humor and satire, and expressed a lot of frustration with our teams when we didn't agree." He expressed regret for not being more outspoken against this pressure (McCullogh).

During the oral arguments in March 2024 for *Murthy*, justices were skeptical of ruling to broadly limit the government's communication with social media platforms. They raised concerns about potentially impeding government officials' ability to discuss certain matters with these platforms. Ultimately, the Court ruled 6-3 that the plaintiffs did not provide enough evidence that the censorship by social media platforms was due to threats from the Biden Administration, thus lacking standing to sue (*Murthy*). This decision was made during a crucial time during the 2024 presidential election, highlighting the ongoing battle against misinformation. Unfortunately, the rise of new generative-AI tools and deepfakes only complicates this battle.

Under *Bantam Books* and *Vullo,* government coercion or inducement constitutes a First Amendment violation since the government's action can be directly traced through intermediaries to the speaker. In contrast, *Murthy* requires more substantial evidence, demanding a clear link that demonstrates a government's past coercion would certainly lead to further restrictions of communication in the future. The Supreme Court in *Murthy* found that, without the aforementioned details, the alleged future harm was no more than conjecture, insufficient to justify an injunction.

There has been criticism of the decision in *Murthy*. The Supreme Court ruled that the plaintiffs lacked standing to challenge the alleged coercion, basing its ruling solely on procedural grounds. Unlike *Bantam Books* and *Vullo*, *Murthy* did not provide substantive First Amendment guidance. Instead, it focused on whether the plaintiffs could demonstrate ongoing harm or direct government coercion, leaving the constitutional issues unaddressed. As Justice Alito noted in his dissent, this decision could create a model that allows future government officials to indirectly influence public discourse by setting a high bar for plaintiffs to prove explicit coercion and immediate harm (Hurwitz, 2024).

In a world where social media platforms serve as gatekeepers to public discourse, their editorial decisions on which speech to publish and the extent of government influence on those decisions carry significant weight. Given social media platforms' First Amendment rights, any governmental effort to combat misinformation by simply compelling these platforms to remove content would not only infringe upon the First Amendment rights of users who posted the content but also violate the platforms' own editorial autonomy.

Therefore, to address these concerns, establishing an appropriate standard framework for regulating false speech requires a balancing test among protecting individuals' rights to free speech, preserving platforms' editorial rights, and enabling the government to address misinformation for the public good. It is imperative for the courts to develop a legal framework that protects all users' free speech, upholds the social media platforms' editorial autonomy, and discerns non-coercive government interactions or engagements with these platforms.

Key Factors for a New Framework

The First Amendment grants the fundamental right to free speech, and the rapid growth of social media has solidified it as a primary platform for expression. However, the rapid expansion of online expression has also raised concerns about the potential for government suppression of protected speech, especially regarding censorship on social medial platforms. Developing a new legal framework to address these issues requires careful consideration of multiple factors to effectively manage misinformation while safeguarding the First Amendment rights.

1. **Balancing Competing Interests**: Finding a precise balance among the interests of users, social media platforms, and government is essential. Imposing overly permissive standards for government intervention risks stifling free communication and violating the First Amendment right, while a lack of action against misinformation risks harming public safety and undermining democratic discourse.

2. **Prioritizing Less Restrictive Alternatives**: Even when the government regulation of speech is deemed permissible, it should prioritize less speech-restrictive and non-coercive measures. If no less restrictive options are available, any government-imposed regulations must be narrowly tailored to serve a compelling public interest, subject to strict scrutiny. Content moderation strategies, such as labeling misinformation, fact-checking, or reducing the visibility of harmful content, should be preferred over outright removal or bans.

3. **Ensuring Transparency and Accountability**: Social media platforms should clearly establish and disclose their content moderation policies, criteria for addressing misinformation, and the decision-making processes behind such content moderation. Similarly, government interactions with platforms must be conducted transparently, with detailed

documentation to prevent undue influence or coercion. Regular audits or oversight by independent bodies can help maintain accountability for both platforms and government.

4. **Defining Clear Boundaries on Government Influence**: Clear and explicit boundaries shall be defined to limit government involvement in content moderation on social media platforms. Advisory roles to combat misinformation shall avoid coercive tactics, threats, or implied penalties aimed at pressuring platforms to censor content. Based on *Murthy*, claims of government overreach shall require clear evidence of coercion directly causing specific harm.

5. **Clarifying Standing Requirements and Evidentiary Standards**: Inspired by *Murthy*, legal challenges against government influence shall meet standing requirements, demonstrating: (a) a specific, concrete injury caused by government action; (b) a direct link between government pressures and the platform's content moderation; and (c) an imminent risk of future harm beyond past actions or conjecture.

6. **Adapting to Emerging Technologies**: The framework shall also account for the challenges posed by rapidly evolving technologies, such as AI and deepfakes, which can significantly amplify the spread of misinformation. Regulatory measures should evolve alongside technological advancements while preserving the First Amendment protections and encouraging innovation.

By incorporating these key factors into a comprehensive legal framework, society can effectively address the challenges posed by misinformation while upholding the fundamental principles of the First Amendment and preserving the editorial autonomy of social media platforms. Transparency, accountability, and respect for constitutional rights are essential to navigating the complexities of misinformation in a way that upholds and strengthens democratic values.

While the government does not have the authority to suppress misinformation without adhering to strict scrutiny standard, it retains the ability to participate in public discourse by sharing its perspectives and even criticizing dissenting views. When done without coercion or suppression, the government can actively encourage and persuade social media platforms to amplify content aligned with public interest or government-endorsed perspectives. In turn, social media platforms can benefit from constructive guidance and insights from the government, helping to foster a more informed and balanced public dialogue.

CONCLUSION

Misinformation poses significant risks to relationships, social cohesion, and society's political systems. However, the enforcement of truth should not fall solely under the government's authority, nor should the responsibility for upholding truth rest entirely with the government. Safeguarding constitutional freedom of speech is the government's role, while it remains our collective duty to seek and uphold truth through such freedom. An effective legal framework shall strike a balance between combating misinformation and preserving free expression, ensuring that government actions remain transparent, non-coercive, and respectful of democratic principles.

Published: April 30, 2025 EDT.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at http://creativecommons.org/licenses/by/4.0 and legal code at http://creativecommons.org/licenses/by/4.0/legalcode for more information.

References

Allyn, B. (2024, November 4). *Social media's lax rules around misinformation worry election watchers*. NPR. https://www.npr.org/2024/11/04/nx-s1-5175792/social-medias-lax-rules-around-misinformation-worry-election-watchers

Bantam Books, Inc. v. Sullivan, 372 U.S. 58 Justia Law ___ (1963). https://supreme.justia.com/cases/federal/us/372/58/

Brennon, V. (2022). *False Speech and the First Amendment: Constitutional Limits on Regulating Misinformation*. Congressional Research Service. https://crsreports.congress.gov/product/pdf/IF/IF12180

Carmichael, J. G. & F. (2020, July 26). *False and misleading claims about the hunt for a vaccine*. BBC News. https://www.bbc.com/news/53525002

Chaplinsky v. New Hampshire, 315 U.S. 568 Justia Law ___ (1942). https://supreme.justia.com/cases/federal/us/315/568/

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*(6425), 374–378. https://doi.org/10.1126/science.aau2706

Harrell, A., & Jones, B. (2024, June 26). *Supreme Court Decides Murthy v. Missouri*. Faegre Drinker. https://www.faegredrinker.com/en/insights/publications/2024/6/supreme-court-decides-murthy-v-missouri

Hurwitz, J. (Gus). (2024, August 6). *After Murthy v. Missouri, Diffuse Jawboning Remains Murky*. The Regulatory Review. https://www.theregreview.org/2024/08/06/hurwitz-after-murthy-v-missouri-diffuse-jawboning-remains-murky/

McCullogh, C. (2024, August 27). *Mark Zuckerberg says Meta was "pressured" by Biden administration to censor Covid-related content in 2021*. CNN. https://edition.cnn.com/2024/08/27/business/mark-zuckerberg-meta-biden-censor-covid-2021/index.html

Meza, S. (2017, November 2). *"Fake news" has been named the word of the year by Collins Dictionary*. Newsweek. https://www.newsweek.com/fake-news-word-year-collins-dictionary-699740

Moravčíková, E. (2020). Media Manipulation And Propaganda In The Post-Truth Era. *Media Literacy and Academic Research* |, *3*(2). https://www.mlar.sk/wp-content/uploads/2020/12/2_Moravcikova.pdf

Murthy v. Missouri, 144 S. Ct. 1972 Justia Law ___ (603 U.S. ___ 2004). https://supreme.justia.com/cases/federal/us/603/23-411/

National Rifle Association of America v. Vullo, 144 S. Ct. 1316 602 U.S. ___ ___ (Justia Law 2024). https://supreme.justia.com/cases/federal/us/602/22-842/

New York Times Co., v. Sullivan, 376 U.S. 254 Justia Law ___ (1964). https://supreme.justia.com/cases/federal/us/376/254/

Overview of Content-Based and Content-Neutral Regulation of Speech. (n.d.). LII / Legal Information Institute. https://www.law.cornell.edu/constitution-conan/amendment-1/overview-of-content-based-and-content-neutral-regulation-of-speech

Police Dept. of Chicago v. Mosley, 408 U.S. 92 Justia Law ___ (1972). https://supreme.justia.com/cases/federal/us/408/92/

Sable Communications v. FCC, 492 U.S. 115 Justia Law ___ (1989). https://supreme.justia.com/cases/federal/us/492/115/

Silenced and misinformed: Freedom of expression in danger during Covid-19. (2021, October 19). Amnesty International. https://www.amnesty.org/en/documents/pol30/4751/2021/en/

Spring, M. (2024, October 29). *How X users earn thousands from US election misinformation and AI images*. BBC. https://www.bbc.com/news/articles/cx2dpj485nno

United States v. Alvarez, 567 U.S. 709, Justia Law ___ (2012). https://supreme.justia.com/cases/federal/us/567/709/

Would you have believed the Great Moon Hoax? (n.d.). Earthsky.Org. https://earthsky.org/human-world/great-moon-hoax-august-25-1835