

Medical Misinformation on Social Media: How Section 230 Impeded Regulation of Online Misinformation

ABSTRACT. Although people have become increasingly reliant on social media for information, misinformation, especially medical misinformation, runs rampant. Companies like Facebook and Twitter have been given discretion on how they maintain their platforms, but the effect is less than desirable. Under their structures, engagement-driven information takes priority over factuality, contributing to a pandemic of misinformation. Despite this influence, social media platforms face no penalties for how users are affected. This is because Section 230 of the Communications Decency Act (CDA) protects interactive computer services like social media platforms from liability for misinformation created by users. During the COVID-19 pandemic, these issues were highlighted when false social media information hurt coordination between government and citizens and increased personal health risks. To protect general health, Section 230 needs to be amended to exempt no-liability protection from user-generated health misinformation, without precautions or user consent. Compliance with this law would require interactive computer services to label health related information for users and review information directly contradicting scientific consensus as compiled by governmental agencies. By exempting no-liability in medical misinformation, social media platforms and other interactive computer services would face incentives to limit the impact of cognitive biases and the spread of misinformation that harms.

AUTHOR. Kevin Zhou is an International Economics Undergraduate student and an upcoming Master's degree student at the School of Global Policy & Strategy at UCSD. He would like to thank James Holliday, Brendon Hahm, and the ULR Editorial Board for their time and advice on this article.

INTRODUCTION

In a letter to Richard Price in 1789, Thomas Jefferson said of education: “[W]herever the people are well informed they can be trusted with their own government; that whenever things get so far wrong as to attract their notice, they may be relied on to set them to rights.”¹ Those words reflect what Jefferson thought to be important to the institutions of democracy and maintaining the values of freedom. From his perspective, information is crucial to the balance between the government and the governed. After all, these ideas were built upon those from the American Revolution, which was represented by Jefferson and other Founding Fathers. The American identity came from a spread of information through newspapers and, later during the crafting of the Constitution, articles like the Federalist Papers. Hence, if people are “well informed,” then they are able to understand the political rhetoric and can properly represent themselves in a democratic government.

In the age of the internet, social media platforms have become major sources of information for the general public. A Pew Research Center study found that eighty-six percent of Americans get news from digital devices such as smartphones, tablets, or computers, and fifty-three percent of Americans often or sometimes get news from social media.² With more than half of Americans receiving news information from social media platforms, the impact of misinformation on social media can be considerable.

This article discusses the spread of misinformation on social media and why the causes of its spread are structural. First, social media platforms cause a cognitive effect called confirmation bias through structural social media algorithms, which can create echo chambers of information. Then, Section 230 of the Communications Decency Act (CDA) enables and expands these chambers by providing special immunity to social media platforms for the content published on their sites. Utilizing these background provisions, this article constructs legal prescriptions for the current systems, namely that Section 230 should be amended to take into consideration the role of social media systems in spreading and facilitating misinformation. These arguments are constructed with the goal of incentivizing social media companies to

¹ Letter from Thomas Jefferson to Richard Price (Jan. 8, 1789), *in* THE THOMAS JEFFERSON PAPERS, (Gen. Correspondence Ser. No. 1, 1651-1827).

² Elisa Shearer, *More than eight-in-ten Americans get news from digital devices*, PEW RESEARCH CENTER (Jan. 12, 2021), <https://www.pewresearch.org/fact-tank/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/>.

design online systems that better dampen misinformation and curb disinformation in order to reduce any kind of resulting harm, including poorly informed health decisions.

I. SOCIAL MEDIA ALGORITHMS

Social media platforms use a tool called “computer algorithms” to manage the order of the content that users see and optimize revenue generation. Access to these platforms is generally free because the business model is based not on maximizing transactions between the platform and the user but rather on maximizing engagement between users on the platform. Revenue generation algorithms on social media platforms collect data to help advertisers maximize the exposure of their advertisements to the users of the social media platform.³ The platform is unlike a conventional commercial platform where the business model is to sell the user a product in a transaction; instead, users give social media platforms engagement time, and social media platforms sell that engagement time to advertisers, who use that engagement to compel viewers to buy products and services.

In order to maximize user engagement, social media platforms manipulate personal content “feeds” in accordance to users’ interests. Those interests are tailored to the user through the use of personal user data. That includes “likes”, offline activity, online engagement topics, and other interactions the user has had on certain content.⁴ This has resulted in an enormous store of user data those social media platforms analyze to recommend content that is engaging to the user.⁵

Due to the quantity of the data, social media companies do not manually organize and interpret user data. Instead, they utilize computer algorithms, which are computer programs created to collect and analyze data using interpretation models to compile content recommendations that are optimized to maximize user engagement.⁶ This is a method that is able to predict, albeit with error, what content will keep users on the

³ Alfred Lua, *How the Instagram Algorithm Works in 2021: Everything You Need to Know*, BUFFER (2021), <https://buffer.com/library/instagram-feed-algorithm/>.

⁴ *Id.*

⁵ *Social Media and the Big Data Explosion*, FORBES (JUNE 28, 2012), <https://www.forbes.com/sites/onmarketing/2012/06/28/social-media-and-the-big-data-explosion/?sh=68e4ecb76a61>.

⁶ Sang Ah Kim, *Social Media Algorithms: Why You See What You See*, 2 GEO. L. TECH. REV. 147 (2017).

platform for as long as possible in a far more cost-effective way than employing human workers to manually collect and organize user data.

II. COGNITIVE BIASES ON SOCIAL MEDIA

Despite their efficiency, the nature of these algorithms results in biases in how people receive information. One research paper looked into a possible bias that comes as a result of social media algorithms: confirmation bias⁷, which is the tendency for people to believe any assertion if it is a confirmation of an existing thought.⁸ Specifically, the paper focused on how confirmation bias emerges in online social groups. By studying the dynamics of opinions spread in online social networks, researchers found that confirmation bias affects these social groups in ways different from offline, in-person social groups.⁹ The model that the researchers constructed had members of an online social network exposed to an external opinion, without being able to immediately reciprocate the flow of information. This is to mimic the effect of receiving information through online sources in contrast to in-person, offline interaction, where information shocks can be promptly addressed through conversation. Over time, the same information was shared by other members of the online social group, gradually changing the perception of the members through confirmation bias to reach a single information equilibrium. The end result is a singular perspective shared by members within the social group. Offline, an exchange of information happens through conversation, creating constant shocks to the information of each member of the social network, resulting in a lack of an equilibrium in information.¹⁰

This portrayal of online social networks, in application, displays how social media algorithms could display a single piece of information to multiple members of an online social network and would recommend similar content to users that share the similar interests. These members could then share this information under their own online social networks to unintentionally cultivate misinformation through confirmation bias. Other members that have already received this information from

⁷ Yanbing Mao et al., *On the Evolution of Public Opinion in the Presence of Confirmation BIAS*, IEEE Xplore (2018), <https://ieeexplore.ieee.org/document/8619824>.

⁸ Shahram Heshmat, *What Is Confirmation Bias?*, PSYCHOLOGY TODAY (Apr. 23 2015), <https://www.psychologytoday.com/us/blog/science-choice/201504/what-is-confirmation-bias>.

⁹ Yanbing Mao et al., *On the Evolution of Public Opinion in the Presence of Confirmation BIAS*, IEEE Xplore (2018), <https://ieeexplore.ieee.org/document/8619824>.

¹⁰ *Id.*

MEDICAL MISINFORMATION ON SOCIAL MEDIA: HOW SECTION 230 IMPEDED REGULATION OF ONLINE MISINFORMATION

their recommended social media feeds would see this information repeatedly, cognitively confirming the truthfulness of this information. If this piece of information is not true, then misinformation is successfully spread.

Furthermore, misinformation may have a more lingering effect than factual information. In a different study focusing on misinformation prevalent on the social media platform Facebook, researchers categorized kinds of information by “science news, conspiracy rumors, and trolling.”¹¹ This was then compared through functions of cascade values (the spread of information) and how long that information had been published. This study found that information labeled as scientific news spread quickly, but stabilized in terms of interest. Information regarded as conspiracy rumors were slower to assimilate, but grew in interest consistently over its lifespan.¹² Conspiracy rumors, which are a type of misinformation, would re-emerge as users access the information in the future. Although factual information gains more initial traction, misinformation lingers on the internet and needs to be addressed to minimize its harm.

However, echo chambers, in addition to computer algorithms, play a tangible role in perpetuating misinformation on social media. An “echo chamber” refers to an online social network where members of the same interest or belief tend to congregate.¹³ One study focused on Reddit, a social media platform where content is rated by “upvotes” and “downvotes” to measure interest. Reddit utilizes “subreddits”, smaller groups where users choose to post relevant content.¹⁴ The study chooses two subreddits: one that shared a common belief and one that discussed a single, general topic. The study found that within groups that are based on a similar belief, misinformation tends to spread more profusely between users within the group. This makes intuitive sense since there are few members that bring a contrasting opinion to the conversation when users in the group share a common belief. In contrast, the “subreddit” that focused on a single, general topic, which suggests conversations between contrasting beliefs about the same topic, displayed fewer examples of

¹¹ Michela Del Vicario et al., *The Spreading of Misinformation Online*, 113 *PROC. NAT’L ACAD. SCI.* 556 (2016).

¹² *Id.*

¹³ Franziska Zimmer, *Fake News in Social Media: Bad Algorithm or Biased Users?*, 7 *JOURNAL OF INFO. SCI. THEORY AND PRACTICE* ISSUE 2, 2019, 40.

¹⁴ Thomas J Law, *What is Reddit? The Ultimate Quickstart Guide for 2022*, Oberlo (Mar. 11, 2022), <https://www.oberlo.com/blog/what-is-reddit>.

misinformation. Comments classified by the study as misinformation were refuted by other users, effectively restricting misinformation.¹⁵

Social media users, like in the case of the Reddit study, may choose to engage with other users that display similarities in belief, creating echo chambers of information. Social media algorithms may also create echo chambers by exposing users to select perspectives that are interesting or engaging in order to maximize engagement time. Both types of echo chambers facilitate misinformation through confirmation bias. These structural systems are intentionally built into social media platforms and are an easy target for policy makers aiming to dampen the spread of misinformation.

III. IMMUNITY FOR SOCIAL MEDIA PLATFORMS

Social media platforms provide an opportunity for their users to publish relatively freely while the platforms themselves retain little liability for facilitating misinformation. This is because social media platforms have special immunity against liability for content published on their platforms. Under Section 230 of the Communication Decency Act (CDA) of 1996, they are able to publish most user content (with few exceptions) without being liable for the content or consequences of these posts.¹⁶ Section 230 has helped foster the growth of social media platforms as well as other online internet services by limiting the responsibility of interactive computer services (ICS) companies. Social media platforms fall within the category of interactive computer services, alongside websites for online services like ecommerce and many others. In the context of social media platforms, this would mean that no matter what content a user posts, the social media platform will not be held responsible for the consequences of these posts.

Section 230 has been successful in supporting interactive internet services during its developmental stage, but now faces criticisms for disincentivizing social media platforms from policing misinformation.¹⁷ The argument is that by protecting social media platforms from liability for what their users create, these platforms have less of

¹⁵ Franziska Zimmer, *Fake News in Social Media: Bad Algorithm or Biased Users?*, 7 JOURNAL OF INFO. SCI. THEORY AND PRACTICE ISSUE 2, 2019, 40.

¹⁶ VALERIE C BRANNON AND ERIC N HOLMES, CONG. RESEARCH SERV., R46751, SECTION 230: AN OVERVIEW (2021).

¹⁷ Daisuke Wakabayashi, *Legal Shield for Social Media is Targeted by Lawmakers*, THE NEW YORK TIMES (Dec. 15, 2020), <https://www.nytimes.com/2020/05/28/business/section-230-internet-speech.html>.

MEDICAL MISINFORMATION ON SOCIAL MEDIA: HOW SECTION 230 IMPEDED REGULATION OF ONLINE MISINFORMATION

an incentive to filter malicious or harmful content, including misinformation and violent rhetoric.

Indeed, according to Tim Kendall, ex-director of monetization at Facebook, Facebook intentionally uses section 230 immunities to promote content that will “provoke, shock, and enrage” its users to maximize attention.¹⁸ He states that the algorithm design is part of an “addictive business model”¹⁹ and, in his testimony to the House Committee on Energy and Commerce, describes the algorithms as a vehicle to “deliver . . . incendiary content . . . in the exact right way” in order to facilitate engagement and profits.²⁰ This proves two points: first, social media platforms like Facebook form revenue in a way that is directly correlated to engagement time, and, second, the use of incendiary content including “shocking images, graphic videos, and headlines that incite outrage”²¹ was internally recognized as a method to increase engagement times. Their algorithms trigger emotions to demand the most attention from Facebook’s users. Because social media companies lack the incentive to moderate content, Section 230 immunities have instead encouraged them to turn to promoting sensational content, regardless of accuracy. This is a huge cause of misinformation’s prevalence online, and an easy source for a solution.

IV. PRECEDENT FOR EXCEPTIONS

There are exceptions to this immunity, however, including criminal cases, intellectual property infringement cases, and sex trafficking cases that are subject to federal criminal law. In fact, the instance of anti-sex trafficking language in Section 230 was only recently signed into the exceptions’ clauses in 2018.²² This was done to combat online sex trafficking websites that have made hundreds of millions of dollars in revenue facilitating prostitution for more than a decade.²³

¹⁸ Elisa Shearer, *More than eight-in-ten Americans get news from digital devices*, PEW RESEARCH CENTER (Jan. 12, 2021), <https://www.pewresearch.org/fact-tank/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/>.

¹⁹ *Mainstreaming Extremism: Social Media’s Role in Radicalizing America: Before the H. Comm. On Energy & Commerce*, 116th CONG. (2020) (statement of Tim Kendall).

²⁰ *Id.*

²¹ *Id.*

²² Elizabeth Dias, *Trump Signs Bill Amid Momentum to Crack Down on Trafficking*, THE NEW YORK TIMES (Apr. 11, 2018), <https://www.nytimes.com/2018/04/11/us/backpage-sex-trafficking.html>.

²³ *Id.*

The specific statute exempting sex trafficking cases from Section 230 is the Allow States and Victims to Fight Online Sex Trafficking Act of 2017 (FOSTA). FOSTA amends Section 230 so that “[E]nforcement against providers and users of interactive computer services of Federal and State criminal and civil law relating to sexual exploitation of children or sex trafficking” is not prohibited.²⁴ In a June 2021 report by the United States Government Accountability Office (GAO), the impact of the FOSTA bill was analyzed and put into the context of major sex trafficking cases that were facilitated through online platforms. The report combines the impact of FOSTA and the federal seizure of backpage.com, “the largest online marketplace for buying and selling commercial sex”, which had happened five days before the enactment of FOSTA.²⁵ These two events happened in April 2018 and were followed by a massive relocation of platforms supporting online commercial sex.²⁶ By 2019, the Underground Commercial Sex Economy (UCSE) had become “fragmented”, with “several signs suggest[ing] . . . that demand has fallen and supply might be falling”.²⁷

However, concerns about the limitations of the FOSTA bill arise from its practical applications. Because the FOSTA bill is so chronologically close to the backpage.com seizure, the impact of the FOSTA bill is difficult to independently assess.²⁸ Furthermore, the Constitution’s Ex Post Facto Clause dictates that “only conduct engaged on or after [the enactment] is punishable by the law”.²⁹ Thus, as part of a new bill, the Section 230 amendment has had little application in court. One example is a Supreme Court of Texas opinion by Justice Blacklock in response to Facebook’s writ of mandamus petition to dismiss three separate claims for Facebook’s “negligence, negligent undertaking, gross negligence, and products liability” regarding sex trafficking on its platform.³⁰ Justice Blacklock recognized that, according to the

²⁴ Allow States and Victims to Fight Online Sex Trafficking Act of 2017, Pub. L. No. 1115-164, 132 Stat. 1253 (2018).

²⁵ U.S. GOVERNMENT ACCOUNTABILITY OFFICE, GAO-21-385, SEX TRAFFICKING: ONLINE PLATFORMS AND FEDERAL PROSECUTIONS (2021).

²⁶ *Id.*

²⁷ Rob Spectre, *Beyond Backpage: Buying and Selling Sex In the United States One Year Later*, CHILDSAFE.AI (2020), https://static1.squarespace.com/static/5b853c2f266c07466413ac7d/t/5caf41bf6e9a7f33bdada158/1554989504655/Beyond+Backpage_+Buying+And+Selling+Sex+In+The+United+States+One+Year+Later.pdf.

²⁸ U.S. GOVERNMENT ACCOUNTABILITY OFFICE, GAO-21-385, SEX TRAFFICKING: ONLINE PLATFORMS AND FED. PROSECUTIONS (2021).

²⁹ The prohibition on state *ex post facto* legislation appears in Art. I, § 10, cl. 1.

³⁰ *In re Facebook, Inc. and Facebook, Inc. D/B/A Instagram*, Relators, No. 20-0434 (Tx. 2021).

majority of legal precedent, these claims would be dismissed under Section 230's effective no-liability for user-generated content, thus granting mandamus relief under common-law claims of negligence and liability. However, he further contended that section 230 does not "Create a lawless no-man's-land on the Internet" and cites FOSTA to deny mandamus relief against statutory human-trafficking claims.³¹ This means that the language of FOSTA allowed Facebook to be sued despite Section 230 protections and the overwhelming precedent of dismissing liability claims towards computer service providers. Statutory changes through the FOSTA bill have led to a change in legal precedent that argues in favor of increased liability from computer service providers.

V. CURBING MISINFORMATION THROUGH RESTRICTING ACCESS

Under Section 230 immunities, social media platforms limit the spread of misinformation through content and access restriction, such as deleting user-generated content and banning users' accounts. For example, repeated violations of Facebook community standards can result in the loss of the privilege to post and react to content on Facebook. These standards include the authenticity of content, safety, privacy concerns, and people's dignity.³² This means that content representing misinformation can be taken down and users spreading such content face restrictions in their use of the platform, thus limiting their ability to spread and affect other users with misinformation.

However, self-regulation can be severely inept. One example of self-regulation failing is described in an article published by the University of Chicago's *Chicago Journal of International Law*, detailing the relationship between social media platforms, specifically Facebook, and the Rohingya genocide in Myanmar.³³ Misinformation campaigns targeting Rohingyas used Facebook to spread hateful messages against the Rohingya people as part of a political agenda. In conjunction with the United Nation's Independent Investigative Mechanism for Myanmar, Facebook released an independent investigative commission towards the platform's effect on the Rohingya genocide and found that it was being utilized to "foment division and incite

³¹ *Id.*

³² *Facebook Community Standards*, META (Feb. 2, 2022), <https://transparency.fb.com/policies/community-standards/?source=https%3A%2F%2Fwww.facebook.com%2Fcommunitystandards%2F>.

³³ Neema Hakim, Comment, *How Social Media Companies Could Be Complicit in Incitement to Genocide*, 21 U. CHIC. L. CHICAGO JOURNAL OF INT'L LAW (2020).

offline violence”.³⁴ Facebook eventually responded, putting in place restrictions on content that incited violence. These restrictions were found to be effective in limiting the spread of misinformation exacerbating the Rohingya genocide.

Although the restrictions themselves proved to be effective, the Facebook example proved that social media platforms cannot be fully trusted to implement those restrictions. Facebook had received warnings of its effect on the Rohingya genocide through public news channels since 2013.³⁵ However, according to a United Nations commission for Myanmar in 2015, Facebook had not appropriately responded to those warnings. While social media companies such as Facebook, Twitter, and YouTube have developed self-regulatory algorithms to limit the spread of misinformation on their platforms, this is not enough to counter their ability to exacerbate conflict through misinformation.

Though conflict may be an irregular occurrence, medical misinformation can have drastic personal effects daily. A 2021 study on Twitter examining the prevalence of misinformation in medical information articles showed that medical misinformation was extremely common on the platform. Smoking and drug-use related articles contained up to 87% misinformation, and vaccine-related articles contained 43% misinformation.³⁶ Overall, a large proportion of medical misinformation is present on Twitter, even after content moderation efforts and restrictions. This shows that current efforts to curb misinformation is far from enough to mitigate the damage of medical misinformation.

VI. LIMITATIONS OF RESTRICTING ACCESS

Further complicating matters, arguments against the legality of social media restrictions have been raised. A significant argument invokes the First Amendment right to freedom of speech. Due to the large range of influence that social media companies have developed, a restriction to the access of posting content on social media would be a deprivation of the freedom of speech.³⁷ Because of that restriction,

³⁴ Alex Warofka, *An Independent Assessment of the Human Rights Impact of Facebook in Myanmar*, META (Nov. 5, 2018), <https://about.fb.com/news/2018/11/myanmar-hria/>.

³⁵ *Id.*

³⁶ Franziska Zimmer, *Fake News in Social Media: Bad Algorithm or Biased Users?*, 7 JOURNAL OF INFO. SCI. THEORY AND PRACTICE ISSUE 2, 2019, 40.

³⁷ David L. Hudson, *In the Age of Social Media, Expand the Reach of the First Amendment*, 43 HUMAN RIGHTS MAGAZINE No. 4 (Nov. 18, 2018), https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/the-ongoing-challenge-to-define-free-speech/in-the-age-of-social-media-first-amendment/.

MEDICAL MISINFORMATION ON SOCIAL MEDIA: HOW SECTION 230 IMPEDED REGULATION OF ONLINE MISINFORMATION

the argument prescribes to the philosophy of market balance—that misinformation on social media should be regulated through a discussion of truths versus falsehoods by users on social media platforms, which would protect users’ first amendment rights on the internet.

Supreme Court Justices recognized in *Packingham v. North Carolina* in 2017 that social media platforms have become unambiguous public forums for communication.³⁸ Communication on social media is relatively effortless, accessible, and costless, compared to traditional public forums that require in-person communication. The summary explained that “[U]sers employ these websites to engage in . . . protected First Amendment activity”.³⁹ This reflects the previous contention that social media companies have an influence on people’s ability to carry out free speech.

An example of this argument being used to police speech limitations from media sources dates back to a previous technological development: radio wave broadcasts. The Fairness Doctrine was a 1949 communications policy by the Federal Communications Commission (FCC) that attempted to balance the representation of news topics on public broadcasts by mandating fair coverage and equal airtime for candidates of public offices.⁴⁰ Licensed broadcasters were required to abide by these FCC regulations.

The basis of the Fairness Doctrine was that media time is scarce and limited by the amount of available radio frequencies.⁴¹ The FCC feared that licensees would only display a single perspective to the public.⁴² Thus, broadcasting stations would be allowed to use airwave frequencies, a public resource, to broadcast their own programs, but would have to host other perspectives at the broadcasters’ expense.

The FCC’s policies regarding the Fairness Doctrine were upheld by the Court of Appeals for the Seventh Circuit in *Red Lion Broadcasting Co. v. FCC* as constitutional.⁴³ Red Lion Broadcasting was a radio station that “carried a program which constituted a personal attack” on Fred J. Cook, a writer. As per FCC regulations, Red Lion Broadcasting was to provide Cook with reply time on the

³⁸ *Packingham v. North Carolina*, 582 U.S. ___, at 2 (2017).

³⁹ *Id.*

⁴⁰ Matt Stefon, *Fairness Doctrine*, BRITANNICA (Mar. 17, 2021), <https://www.britannica.com/topic/Fairness-Doctrine>.

⁴¹ *Id.*

⁴² Audrey Perry, *Fairness Doctrine*, THE FIRST AMENDMENT ENCYCLOPEDIA (May 2017), <https://www.mtsu.edu/first-amendment/article/955/fairness-doctrine>.

⁴³ *Red Lion Broadcasting Co. v. FCC*, 395 U.S. 367 (1969).

broadcast so as to address his own perspective for the audience. Red Lion denied Cook reply time, claiming First Amendment rights to freedom of speech.

However, the Seventh Circuit upheld the legality and constitutionality of the FCC's regulations. *Red Lion* stated that the Fairness Doctrine helped to maintain the First Amendment rights of the public.⁴⁴ The logic followed that, because of the limited airwaves available for broadcast, the public held rights to broadcast time as well when it came to representation and the public's freedom of speech. The public's freedom of speech then ruled over the radio station's freedom of speech, and, thus, the FCC's regulations promoted freedom of speech rather than harmed it.

The Fairness Doctrine was later reviewed and receded as new technology allowed more broadcasting stations to be introduced, providing a variety of perspectives for American radio audiences.⁴⁵ Yet this example shows a potential scenario in which media companies can have control over their own platforms limited to serve the public good. More importantly, this shows that government moderation, specifically FCC intervention, is possible when systemic issues causing misinformation are present.

Quite the opposite of preventing social media companies from restricting access, current legal perspectives maintain the right for social media companies to restrict access of its users to content that goes against company guidelines. Private entities generally do not fall under the first amendment. However, the Supreme Court conducts a test to determine exceptions to that rule: whether the private entity provides services that are a "traditional, exclusive public function".⁴⁶ This means that private companies that provide services that would otherwise be exclusively provided by the government would fall under first amendment jurisdiction. This is the basis that allows social media companies to manage the content on their platforms.

In *Nyabwa v. Facebook*, the court found Facebook not guilty of a breach of the First Amendment because the First Amendment applies only to restrictions by the government, not private parties. Facebook, as a private company, does not fall under the jurisdiction of the First Amendment.⁴⁷ This decision also allows companies to continue to regulate their websites without fear of infringing on the right to freedom of speech of its users. Furthermore, social media platform usage is by no means a traditional service managed by the government. With the current legal landscape,

⁴⁴ *Id.*

⁴⁵ U.S. GOVERNMENT ACCOUNTABILITY OFFICE, GAO-21-385, SEX TRAFFICKING: ONLINE PLATFORMS AND FED. PROSECUTIONS (2021).

⁴⁶ *Manhattan Community Access Corp. v. Halleck*, 587 U.S. __ (2019).

⁴⁷ *Nyabwa v. Facebook*, CIVIL ACTION NO. 2:17-CV-24 (S.D. Tex. Jan. 26, 2018).

MEDICAL MISINFORMATION ON SOCIAL MEDIA: HOW SECTION 230 IMPEDED REGULATION OF ONLINE MISINFORMATION

companies like Facebook can control the extent to which they choose to curb misinformation on their sites without fear of intruding on first amendment rights of their users. However, as mentioned before, the ability to moderate content does not sufficiently protect users from the severe misinformation that exists.

VII. AMENDING SECTION 230

The prevalence of dangerous misinformation was highlighted during national emergencies like the COVID-19.⁴⁸ To mitigate its effect, countries around the world shut down their citizens' movements to wait for the development of a vaccine. The United States Food and Drug Administration (FDA) issued an Emergency Use Authorization for the first COVID-19 vaccine on December 11, 2020. Vaccinated people were predicted to have more than a ninety percent chance of immunity.⁴⁹ However, as of February 2022, less than sixty-five percent of Americans were fully vaccinated.⁵⁰ Like in the case of measles, which was believed to be eradicated in 2000 but observed a resurgence in the late 2000s, vaccine hesitancy likely contributed to this low vaccination rate.⁵¹ Misinformation about vaccines discouraged many Americans from taking the COVID-19 vaccine, resulting in deaths and disabled the workforce. Higher vaccination rates could have limited the effect of COVID-19 under the current scope, and combating misinformation would be an effective way of achieving this goal.

In order to combat such misinformation, the incentives of social media companies need to change. If social media platforms face a need to limit misinformation, then they can design algorithms that quell the spread of misinformation. To incentivize social media platforms to change their algorithms that are spreading misinformation, holding them responsible for spreading such misinformation is the most effective way. Like when Section 230 was amended to include exceptions for sex-trafficking cases, Section 230 should again be amended to incentivize moderation of online content. This can be

⁴⁸ *SARS-CoV-2 Transmission*, CENTERS FOR DISEASE CONTROL AND PREVENTION (May 7, 2021), <https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/sars-cov-2-transmission.html>.

⁴⁹ *FDA Takes Key Action in Fight Against COVID-19 By Issuing Emergency Use Authorization for First COVID-19 Vaccine*, U.S. FOOD & DRUG ADMIN., (Dec. 11, 2020), <https://www.fda.gov/news-events/press-announcements/fda-takes-key-action-fight-against-covid-19-issuing-emergency-use-authorization-first-covid-19>.

⁵⁰ *Coronavirus (COVID-19) Vaccinations*, OUR WORLD IN DATA, (Feb. 25, 2022), <https://ourworldindata.org/covid-vaccinations?country=USA>.

⁵¹ Christian Akem Dimala et al., *Factors Associated with Measles Resurgence in the United States in the Post-Elimination Era*, SCI. REPORTS (Jan 8, 2021), <https://www.nature.com/articles/s41598-020-80214-3#Sec10>.

done by adding a new exemption to Section 230(c)(1): exempt no-liability from the spread of health misinformation, without reasonable precautions or user consent, that directly causes or otherwise incites bodily harm.

In this context, health misinformation would be defined as “[A]ny health-related claim of fact that is false based on current scientific consensus”.⁵² Like during the Fairness Doctrine, the standards for those claims would be determined by the FCC, in conjunction with relevant government agencies such as the Center for Disease Control (CDC) and Food and Drug Administration (FDA).

This argument comes from two main points: first, the increased control social media companies have over what content is displayed, and second, the original intent of Section 230 being a precautionary incentive for social media companies to self-regulate. In the last few years, Twitter and other social media platforms have switched from personal “feeds” based on chronology to personalized feeds created by deep-learning algorithms using users’ data.⁵³ Through this algorithm, social media platforms have more control over what content becomes interacted with more often, as well as the categories of users’ posts that get more exposure. By having more control over the content, social media companies should also have more responsibility for the content they promote. Thus, amending misinformation and its consequences into Section 230’s exceptions’ clause is a legislative change that plays to the ability of social media platforms. Their ability to facilitate misinformation means that they should have an increased responsibility for the spread of misinformation.

Furthermore, although Section 230 has become successful in facilitating the expansion of online services, the original intent of Section 230 was to “modernize the existing protections against obscene, lewd, indecent, or harassing uses of a telephone.”⁵⁴ Although Section 230 is most cited for protections against liability for user-generated content, its real intent was to protect companies as they actively police their own platforms. Leading up to Section 230 was *Stratton Oakmont v. Prodigy*, a New York Supreme Court case that indicted Prodigy’s “Money Talk” bulletin board as responsible for defamation of Stratton Oakmont, Inc. due to user-generated content subject to defamation claims being published on the site. The court found Prodigy

⁵² Sylvia Chou et al., *Where We Go From Here: Health Misinformation on Social Media*, 110 AM. JOURNAL OF PUBLIC HEALTH SUPPL. 3, 273-275 (2020).

⁵³ Nahema Marchal et al., *Junk News During the EU Parliamentary Elections: Lessons from a Seven-Language Study of Twitter and Facebook*, OXFORD INTERNET INST. (2019), <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/05/EU-Data-Memo.pdf>.

⁵⁴ U.S. GOVERNMENT ACCOUNTABILITY OFFICE, GAO-21-385, SEX TRAFFICKING: ONLINE PLATFORMS AND FED. PROSECUTIONS (2021).

MEDICAL MISINFORMATION ON SOCIAL MEDIA: HOW SECTION 230 IMPEDED REGULATION OF ONLINE MISINFORMATION

liable because they actively filtered out offensive language, employed moderators to enforce content guidelines for users' content and, thus, were able to be treated as the publisher for the user-generated content.⁵⁵ Lawmakers saw this decision as a disincentive for computer service companies to regulate offensive content on their platforms because of concern that liability for screening content would cause companies to stop screening instead of improving screening capabilities.⁵⁶ Section 230 was a response to combat screening disincentives proposed by Representatives Christopher Cox and Ron Wyden as an amendment to the CDA. Section 230(c)(2), the "Good Samaritan" clause, gives companies freedom to screen their own platforms without legal repercussions.⁵⁷ The law was created to incentivize self-moderation by social media companies, not excuse it.

But because Section 230 currently only provides companies with the ability to police their platforms, not the incentives, its goal is not fully realized. The proposed new amendment would require social media companies to improve algorithms that moderate content and information spreading structures such as recommendation algorithms so that users can become more aware of sources of information.

If these system changes are not implemented to protect users, cases to sue companies for any content causing bodily harm can be brought into court on the grounds of this new statutory change. Otherwise, misinformation-based cases will be readily dismissed, like sex trafficking cases were before FOSTA.

VIII. IMPLICATIONS OF SECTION 230(C)(1)

It is important to note that companies under this new exemption are only liable for facilitating the spread of misinformation. This means that companies cannot be held liable if they have acted to substantively mitigate the spread of misinformation. One example of this is algorithms on social media platforms that tag content containing medical advice. During the Covid-19 pandemic, major platforms like Facebook and Twitter highlighted content that included information about the coronavirus and warned users to double-check the information with the Center for Disease Control and Prevention (CDC) website for accuracy. By giving users an instantaneous way to

⁵⁵ Stratton Oakmont, Inc. v. Prodigy Services Co., 23 MEDIA L. REP. 1794 (N.Y. Sup. Ct. 1995).

⁵⁶ Danielle K Citron and Benjamin Wittes, *The Problem Isn't Just Backpage: Revising Section 230 Immunity*, 2 GEORGETOWN L. TECH. REV. 453 (2018).

⁵⁷ Daisuke Wakabayashi, *Legal Shield for Social Media is Targeted by Lawmakers*, THE NEW YORK TIMES (Dec. 15, 2020).

verify information, Facebook essentially slows the spread of misinformation without restricting the articles themselves.

This additional information also prompts readers to question the accuracy of what they are viewing on the internet. In a study, researchers found that by prompting survey respondents to think about the accuracy of news headlines, the accuracy of information shared also increased⁵⁸. Thus, social media platforms are able to increase the accuracy of information on their platforms by prompting users to think about the accuracy of their information. Doing so would be progress in mitigating the spread of misinformation.

However, this may also not be enough. Misinformation on social media partly spreads because of a desensitization to spreading inaccurate information on social media platforms. In a study, researchers found that subjects asked to assess a set of information labeled samples as true around fifteen percent of the time, and were willing to share the same set of information more than thirty percent of the time.⁵⁹ This showed that misinformation was shared not so much as a result of an inability to discern the accuracy of information but rather because “[P]eople were apparently willing to share content that they could have identified as being inaccurate.”⁶⁰ The study explained that the social media context distracts people’s attention from providing accurate information. Instead, the “attention economy” of social media platforms diverts users’ attention towards maximizing engagement, resulting in an increase in the share of low-quality news content.⁶¹ Because of the structure of social media platforms, users have begun to forgo the accuracy of content in favor of gaining engagement on the platform, which exacerbates the prevalence of misinformation.

The structural algorithms behind social media platforms’ design results in the degree of the spread of misinformation. Thus, to combat misinformation, policy should target the systems that allow such misinformation to spread. By holding social media companies liable for physical harm caused by misinformation, those companies are incentivized to change their algorithms and provide buffers against the spread of misinformation, like additional informational prompts for users consuming information on social media platforms. User warning and consent can come in the form of joining public groups and interactions like “follow” and “friend,” allowing

⁵⁸ Gordon Pennycook et al., *Shifting Attention to Accuracy Can Reduce Misinformation Online*, 592 NATURE (2021).

⁵⁹ Gordon Pennycook and David G Rand, *The Psychology of Fake News*, 25 TRENDS IN COGNITIVE SCI. ISSUE 5, 388-402 (2021).

⁶⁰ *Id.*

⁶¹ *Id.*

MEDICAL MISINFORMATION ON SOCIAL MEDIA: HOW SECTION 230 IMPEDED REGULATION OF ONLINE MISINFORMATION

users to still publish and spread information, only in a way that mitigates cognitive bias. These changes, in conjunction, will help to prevent medical misinformation incidents and enhance general public safety.

CONCLUSION

Following the explosive development of internet services over the past few decades, the influence that social media platforms now have on people's lives is beyond the scope of many United States' existing laws. Legislators have and continue to look for ways to manage internet services so that society as a whole can reap its benefits without having to sacrifice the rights and well-being of individuals. Crises like the Covid-19 pandemic and the Rohingya genocide have showcased the drastic effect that social media platforms like Facebook and Twitter can have on information, altering people's perception of both problems and solutions.

Currently, we are balancing information and misinformation through a "free-market equilibrium"—by believing that true information will trump over misinformation either through users' support of accurate information or the goodwill of social media platforms to eradicate misinformation. As such, Section 230 of the Communications Decency Act (CDA) grants near-full immunity to platforms against publishing misinformation generated by users. However, this does not provide them with incentives to maintain information accuracy. Computer algorithms that these platforms use to organize what users see are prone to creating echo chambers through cognitive biases, effectively diminishing interaction between true information and misinformation. This phenomenon of "echo chambers" combined with the goal of these platforms to optimize engagement time can explain a large part of the prevalence of misinformation on social media.

The alteration of incentives provides an effective way to limit health-related misinformation, an area that directly affects the lives of all Americans and contains an abundance of scientifically-proven information—much of the existing health-related misinformation on the internet is directly contradicted by existing scientific consensus. By creating standards for medical accuracy online, governmental agencies and health experts would be tasked to set the standards for proven misinformation. That is not to say those standards could not be changed. Scientific research constantly makes breakthroughs on the most basic assumptions of research. That is what gives science the rigorousness that makes consensus in the scientific field meaningful. Yet without limitations on the spread of misinformation on the internet, the rigorousness of true information becomes less useful.

To resolve this, legislation needs to target the structural incentives behind social media platforms. By amending Section 230 to hold interactive computer service companies liable for health misinformation spread on their platforms, these companies would be incentivized to enhance algorithms that detect misinformation, provide users with closer access to reliable sources of information, and set up systems that improve internet literacy by targeting cognitive behaviors.

These algorithms will present themselves in the form of timely warnings, accuracy requirements for content claiming to be medical information, and other direct verification methods for social media users. Users can still choose to publish or receive unverified medical information, but computer algorithms will no longer recommend that information to users without warning or consent. By bringing on these changes, the fundamental structure of social media will be changed. These changes will mitigate cognitive biases that drastically worsen the misinformation pandemic, and overall health safety of the public will rise.