

**To what degree should freedom of speech be limited in order to prevent cyberbullying on  
social media?**

Ethan Lasich

School of Cybersecurity, Old Dominion University

IDS 300W: Interdisciplinary Theory and Concepts

Professor Pete Baker

August 7th, 2025

Bullying has always been a rampant issue for many, mainly affecting young children and teenagers. The US government defines bullying as the “unwanted, aggressive behavior among school aged children that involves a real or perceived power imbalance. The behavior is repeated, or has the potential to be repeated, over time” ([stopbullying.gov](http://stopbullying.gov), 2023). With the widespread adoption of social media, bullying has moved from schools and playgrounds to the digital world, and it has brought with it the same harm, as well as new problems to deal with.

Many believe that cyberbullying is actually more harmful than traditional bullying, mostly due to the anonymity of others on the internet, the faster spread of information, and it also follows the victim everywhere, including their home. In a study conducted by Fabio Sticca and Sonja Perren, they had participants rank the severity of hypothetical bullying situations with variables such as traditional or cyber bullying, public and private, and the anonymity of the bully, and the scenarios that involved cyber bullying were “perceived as worse than traditional ones,” (Sticca & Perren, 2013). If cyberbullying is truly so terrible for many people, then what is stopping companies from just banning any cyberbullies from their sites. Ethically it would make sense to remove comments that harm or offend other users, but what does the law say? Many believe that the first amendment of the Constitution actually protects hate speech and cyberbullying under the first amendment, “The First Amendment was designed to protect offensive speech, that makes it difficult for governments to draft anti-bullying laws that are both effective and constitutional.” (Leess, 2011). With that in mind it raises a question, to what degree should freedom of speech be limited in order to prevent cyberbullying on social media? Although the question might sound simple, figuring out the perfect balance to limit cyberbullying and avoid infringing on free expression becomes a very complex issue, requiring

insights from psychology, law, philosophy, and even economics, warranting the use of an interdisciplinary approach to find an answer.

From a psychological standpoint, cyberbullying can have countless negative effects on the mental health of individuals, especially compared to traditional bullying because of its invasive nature. Cyberbullying affects its victims emotionally and is linked to causing mental health issues such as anxiety, depression, low self esteem, and worst of all suicidal thoughts. A three year study done in India drew data from a survey of 16,292 adolescents, and their findings backup this claim, with their conclusion being “The findings suggest that cyberbullying victims are at higher risk of depressive symptoms and suicidal thoughts and these adverse effects persist for longer period.” (Maurya, Muhammad, & Dhillon et al, 2022). Depression and suicide are not minor side effects of cyberbullying, as they pose a severe risk to a person’s well being and life. Long term depression can lead to problems with academic performance, relationships, and negatively impact a person’s development as a person, and in their brain. In the most extreme cases, victims may see suicide as the only escape from the bullying, giving more of a moral obligation to social media companies to hold users accountable for what they post. Looking into the sociological reasons behind why a perpetrator decides to cyber bully, a lot of it stems from power and anonymity. There is a psychological phenomenon written about by John Suler in 2004, which he calls the online disinhibition effect, but it still holds up today. The effect explains how the lack of face to face interaction online leads to a lack of empathy towards whoever a bully is interacting with, and therefore makes it easier to portray aggressive and harmful behavior (Suler, 2004). Another thing that drives cyberbullying is peer dynamics and social media likes. For example if somebody posts a hateful comment online and is rewarded with like and comments, it normalizes the behavior and incentivises others to join in on it, creating a mob

mentality and a feeling of entrapment against the victim. This combination of psychological and sociological effects makes cyberbullying a complex behavior that is difficult to address through a single lens.

Despite how cyberbullying is so psychologically damaging, the opportunity for legal response in the US is quite limited under the first amendment right to free speech. The Constitution is made to protect individuals differing opinions or beliefs, even if those ideas are offensive or harmful to others. In a 2011 Supreme Court case Chief Justice John Roberts said “Speech is powerful. It can stir people to action, move them to tears of both joy and sorrow, and—as it did here—inflict great pain. On the facts before us, we cannot react to that pain by punishing the speaker. As a Nation we have chosen a different course—to protect even hurtful speech on public issues to ensure that we do not stifle public debate.” (“Is Hate Speech Legal?” n.d.). What he is basically saying is that by regulating hate speech, it eventually leads to a path where dissenting opinions are silenced, which are needed for a functional democracy. Due to this case and similar legal precedent, it complicates many attempts to regulate cyberbullying. Typically “hate speech can only be criminalized when it directly incites imminent criminal activity or consists of specific threats of violence” (“Hate Speech and Hate Crime,” n.d.), which means that a lot of the time, insults, criticism, or cruel social media posts are seen as legally permissible, even if they are psychologically damaging. With the legal system’s high threshold for restricting free speech, it leaves many individuals vulnerable to hate speech, and makes it more difficult for social media companies to keep its users safe.

Even though the legal system works to defend a broad freedom of expression, ethical beliefs offer a more balanced position on ways to deal with hate speech. An article by Melina Bell suggests regulating hate speech using two popular philosophical ideas called John Stuart

Mill's Harm Principle, as well as his Greatest Happiness Principle. The Harm Principle is an idea that believes individual freedom of speech is allowed to be limited in order to prevent harm to others, and the greatest happiness principle states "actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness" (Mill, 1863). Under these ethical considerations, Bell concludes that harmless speech, no matter how offensive, is not justified to be restricted, however speech that is "both harmful and fails to express a genuine opinion, and does not deserve free speech protection." (Bell, 2021). Some more extreme philosophical takes, such as one Joel Feinberg writes about in his book "Offense to Others", argues to associate the concept of offense with, and that any statement "preventing shock, disgust, or revulsion is always a morally relevant reason for legal prohibitions." (Feinberg, 1984). His take would mean even if hate speech did not meet legal thresholds for restrictions, there is still a moral obligation to protect individuals for speech that harms, isolates, or humiliates others. Even if the law cannot create a safe environment on social media, we still have an ethical responsibility as a society to hold cyberbullies accountable for what they say online.

Due to the fact that U.S. law is not always able to hold cyberbullies accountable, most of the responsibility for controlling hate speech ends up falling on social media platforms themselves. Since platforms like Instagram, YouTube, and X/Twitter aren't bound by the first amendment, it means they can create their own terms of service that allow the deletion of any content that they find hateful or abusive, however they are never legally obligated to do so. Even with these rules though, problems arise when it comes to enforcing them or deleting posts quick enough before they're able to cause psychological harm. However, even though the law and the mental wellbeing of its users allows companies to delete hate speech, sometimes economics is also a factor in these decisions. An important thing to note is that controversial posts gather a lot

of user engagement, which is what social media platforms profit from, so even the morally right thing to do is take down the post, it might not always be in their best interest to delete offensive content if it's making them money. According to Section 230 of the Communication Act of 1934, which was updated in 1996, "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider," and this broad statement is interpreted to mean that social media companies face no liability for offensive content that is posted to their platform by users, which means companies have zero legal obligation to take down hate speech, essentially leaving the decision up to them, but is section 230 really what's stopping companies from having to remove cyberbullying and hate speech. Well in 2018, Congress enacted FOSTA, which stated that sex trafficking content is no longer protected under Section 230, and Alan Rozenshtein, a professor of law claims "some platforms disappeared altogether and the ones that remained removed not just sex-trafficking content but also anything having to do with sex work." (Rozenshtein, 2024). He also believes because of that, if Section 230 was revised to have a more narrow interpretation "Platforms would likely respond to such a decision by massively increasing moderation, removing not just illegal content but also large amounts of legal content, just to be on the safe side." (Rozenshtein, 2024), referencing companies' response to FOSTA as the reasoning behind this. Because social media platforms aren't legally obligated to remove anything that doesn't violate criminal law, content regulation comes down entirely to what the social media platforms are motivated by, whether that's money or moral standards.

Another regulatory body that can control hate speech on social media is schools and universities, which mainly focuses on cyberbullying issues between its students, and their decisions are more motivated by ethics and sociology rather than by economics and law. Schools

have the duty to maintain a safe learning environment and protect their students well being, but does that translate to the online world? Justin Patchin, the co-director for the Cyberbullying Research Center mentions that under Tinker V. Des Moines, “educators can restrict students’ speech only if it causes “a substantial disruption” of school activities.” (Patchin, 2018) Furthermore, the supreme court case Mahanoy area school district vs BL challenged the interpretations of Tinker V. Des Moines and concluded that universities can hold students accountable for off campus speech if it is considered severely disruptive. With this in mind, school regulations are made to balance both the preservation of students freedom of expression and also the prevention of speech that incites harassment or poses a psychological threat to other students' well being. From a sociological perspective, most universities' policies aim at promoting inclusivity and deterring conflict between groups of students. Also according to Justin Patchin, “research shows that students who believe schools will punish them for cyberbullying are less likely to torment their classmates than those who don’t fear punishment.” (Patchin, 2018)

After looking at how psychology, law, sociology, and economics affect cyberbullying from different angles there is a noticeable amount of conflict in how each discipline approaches limiting free speech to prevent cyberbullying. Sociology helps emphasize the cultural norms and societal conflict that drive cyberbullying, and psychological reasoning shows that cyberbullying clearly causes mental health dilemmas, supporting an ethical reasoning to step in and limit hateful speech. The law on the other hand is restrained by protections in the first amendment and mostly protects offensive speech, as long as it isn’t threatening or on an extreme level of harassment, so if the government doesn’t have a legal basis to restrict speech then why should anyone else? Economics, which explains why social media companies might hesitate to remove controversial content due to its user engagement level, also plays a role in why free speech might

not be able to be limited, even if it's to protect an individual from being targeted with hate. Using an interdisciplinary approach, I believe the best solutions to both respect free expression and keep social media users free from harm involve creating laws with clearer definitions for speech that is not protected by the first amendment, putting more pressure on universities and social media companies to block hateful speech, using philosophical frameworks like John Stuart Mill's Harm Principle to define hate speech, and using psychological principles to inform people on what hate speech can cause to a person's brain. Even though each discipline only provides partial insights into the causes and effects of cyberbullying, common ground can be found by integrating them and finding that both protecting free speech and protecting people online is not only possible, but better for people and society overall.

In conclusion, by looking at all the conflicting perspectives that different disciplines bring about how to prevent cyberbullying, it becomes clear that a single discipline does not hold all the answers to this problem. Instead, integrating the ideas all together for a single interdisciplinary approach provides the best way to find how to balance free speech and protect individuals from psychological harm online. By combining insights from legal sources, sociological understandings, ethical standards, and economic incentives, we can work towards a solution that defends both free expression and creates safe environments online. Only through a collaboration across disciplines does it become possible to find an answer that can satisfy everybody.

## References

- Bell, M. C. (2021). *John Stuart Mill's harm principle and free speech: Expanding the notion of harm. Utilitas*, 33(2), 162–179. <https://doi.org/10.1017/S0953820820000229>
- Feinberg, J. (1984). *\*Offense to others\**. Oxford University Press.
- Hate speech and hate crime. (n.d.). American Library Association. Retrieved August 9, 2025, from <https://www.ala.org/advocacy/intfreedom/hate>
- Is Hate Speech Legal? (n.d.). Foundation for Individual Rights and Expression. Retrieved August 9, 2025, from <https://www.thefire.org/research-learn/hate-speech-legal>
- Leess, M. (2011, May 1). Cyber-Bullying A Growing Threat & a Free-Speech Dilemma. Democracywise. <https://democracywise.syr.edu/?p=2098>
- Maurya, C., Muhammad, T., Dhillon, P. et al. The effects of cyberbullying victimization on depression and suicidal ideation among adolescents and young adults: a three year cohort study from India. *BMC Psychiatry* 22, 599 (2022). <https://doi.org/10.1186/s12888-022-04238-x>
- Mill, J. S. (1863). *Utilitarianism*. Indian Hills Community College. [https://webcontent.indianhills.edu/\\_myhills/courses/PHI105/documents/lu07\\_utilitarianism.pdf](https://webcontent.indianhills.edu/_myhills/courses/PHI105/documents/lu07_utilitarianism.pdf)
- Patchin, J. W. (2018, February). Should schools punish off-campus cyberbullying? *Senator Scott Martin*. Retrieved from <https://www.senatorscottmartinpa.com/wp-content/uploads/sites/87/2020/02/sb05-articles-1.pdf>

Rozenshtein, A. Z. (2024, April 17). *Interpreting the ambiguities of Section 230*. *Yale Journal on Regulation Bulletin*. Retrieved from

<https://www.yalejreg.com/bulletin/interpreting-the-ambiguities-of-section-230/>

Sticca, F., & Perren, S. (2013). Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying. *Journal of youth and adolescence*, 42(5), 739–750.

<https://doi.org/10.1007/s10964-012-9867-3>

StopBullying.gov. (2023, December 1). *What is bullying*. U.S. Department of Health & Human Services. <https://www.stopbullying.gov/bullying/what-is-bullying>

Suler, J. (2004). *The online disinhibition effect*. *CyberPsychology & Behavior*, 7(3), 321–326.

[https://johnsuler.com/article\\_pdfs/online\\_dis\\_effect.pdf](https://johnsuler.com/article_pdfs/online_dis_effect.pdf)

Wikipedia contributors. (2023, December 1). *Mahanoy Area School District v. B.L.* Wikipedia.

[https://en.wikipedia.org/wiki/Mahanoy\\_Area\\_School\\_District\\_v.\\_B.L.](https://en.wikipedia.org/wiki/Mahanoy_Area_School_District_v._B.L.)