8-30-2021

Section 230 Immunity and Being Cancelled: A Cause of Action Against Twitter

Jordan Doll
Cardozo Arts & Entertainment Law Journal

Section 230 Immunity and Being Cancelled: A Cause of Action Against Twitter

BY JORDAN DOLL/ ON AUGUST 30, 2021

Introduction

Section 230 immunity of the Communications Decency Act ("CDA") shields internet service providers from liability for the content posted on their sites by third-party users in almost all instances.[1] One such instance is libelous tweets, meaning Twitter cannot be sued when a user tweets something defamatory.[2] If Section 230 immunity was revoked, the allegedly defamed individual could sue Twitter directly for these tweets, rather than only having a cause of action against the individual user. [3] Removal of Section 230 immunity provides a hypothetical remedy against the platform itself when someone is defamed online.

Under the current statutory regime, a potential plaintiff can sue another Twitter user that posts an allegedly defamatory tweet.[4] Social media has led to an increase in libel lawsuits [5] because Twitter and similar platforms exponentially expand the number of people who are able to speak to a mass audience.[6] However, the target of a defamatory tweet cannot

sue the platform itself, despite Twitter providing the environment that allowed the tweet to reach an audience and cause potential reputational and psychological harm.[7]

Online public shaming, which often involves defamatory comments, is ubiquitous on Twitter and other social media sites where the sites' algorithms help promote the spread of defamatory posts or tweets.[8] Tweets that criticize another person's behavior and language can often snowball into many users targeting a single individual, especially because content that sparks an intense emotional response is more likely to go viral.[9] This snowball effect is sometimes referred to as someone being cancelled.[10] Sometimes this online public shaming involves defamatory or untrue comments that receive widespread dissemination, causing negative psychological effects or other fallout to the target that may be undeserved.[11] In cases of online shaming where the remarks are defamatory, the law offers limited redressability and the harmed individual has no cause of action against Twitter, which has allowed the online shaming to occur on a particularly large scale.[12] Without Section 230 immunity, Twitter could be held liable for these allegedly defamatory Tweets; if Twitter faced liability for the defamation, the threat of a lawsuit from any single tweet may render Twitter more cautious in policing speech and public shaming would occur online with less dramatic effect.[13] Removing Section 230 immunity would alter how we publicly shame online because Twitter and similar social media sites would be incentivized to remove posts used as tools of public shaming out of fear that they contain defamation for which the platform could be held liable.

What is Section 230 Immunity

Section 230 immunity helped create the modern internet, allowing platforms like Twitter, Facebook and Google to flourish unrestrained from potential liability arising from the contents their users share.[14] The broad immunity encouraged websites to monitor their sites for pornography and other potentially obscene or offensive content without fear of liability for other user-generated content.[15] Section 230 immunity was, in part, a response to Stratton Oakmont v. Prodigy, where a New York state court held that a web services company could be sued for allegedly defamatory posts by its users.[16] Because the website moderated or removed some of the posts on its website, the Court reasoned that it acted as a publisher of all user-generated posts, even for those the site did not directly monitor.[17] [18] Congress immediately recognized the ruling's potential for creating massive liability because any website that attempted to monitor some of the content posted on its platform would be liable for all defamatory or otherwise illegal content posted on its platform.[19] This could create a chilling effect on speech because the potential for mass liability would create a disincentive for websites to function.[20] At scale, it would not be feasible for platforms to moderate every single post for possible illegal or defamatory content.[21] To avoid such potential liability, the provider had a huge incentive to restrict content posted, limiting the innovation that can result from an unconstrained marketplace of ideas.[22] Congress acted to avoid this chilling effect by enacting Section 230 immunity.[23]

Courts have construed Section 230 immunity for websites broadly.[24] In *Zeran*, the Fourth Circuit reasoned, "[i]nteractive computer services [or websites] have millions of users. The amount of information communicated via interactive computer services is therefore staggering. The specter of tort liability in an area of such prolific speech would have an obvious chilling effect" on speech. [25]

Often relying on the statute's findings section that reflected the congressional desire to allow the internet to develop unburdened by tort liability, courts almost never hold websites accountable for what their users post.[26] Section 230 immunity allowed for social media sites to grow and thrive, unrestrained by potential liability of the massive amounts of content hosted on these platforms.[27]

Public Shaming Online

Twitter and similar social media sites, protected by Section 230 immunity, allow public shaming to happen in a particular way. Internet communication is instantaneous and it often encourages hyperbole and exaggeration.[28] Internet speech has an informality that frees discourse of the polite conventions that often constrain it in the real world.[29] Speech online also has greater access to an audience and can be disseminated very quickly through re-tweeting and other methods of boosting content.[30] When widely shared tweets or other social media posts shame an individual online it is sometimes called cancel culture or canceling an individual.[31]

When the term became popular on Black Twitter,[32] canceling someone or something was not an attempt to incite a boycott or de-platform them, but signaled a personal decision to move on from that person or concept.[33] Propelled by changing power dynamics in the wake of #MeToo, being canceled morphed in the mid-2010s to a catch-all phrase where someone is de-platformed and their reputation is dashed for what the internet deems unacceptable behavior or speech. Especially in conservative media, the term is equated with the joke or real fear of when people on the left are unnecessarily upset about behavior or speech and mete out outsized consequences as a result.[34] Among Americans familiar with the term, there is little consensus about what being canceled actually means, with a particular divide over whether the loaded phrase refers to holding people accountable for their actions or punishing people for their actions.[35]

Though there is debate and disagreement surrounding the meaning of the term being canceled, this article will use the phrase to describe a particular internet phenomenon where a large group of people, often on Twitter, identify behavior or speech that falls short of a community standard; as a result, the online users call out that individual's behavior or speech, and the individual faces consequences such as losing their online platform, receiving online verbal abuse, and experiencing reputational harm.

The effects of canceling someone are likely both positive and negative. Canceling people can be a productive form of public shaming when it encourages deterrence of problematic behavior and discourages members of society from committing similar violations.[36] It can be unproductive public shaming when the accusations are defamatory, harming the targeted individual reputationally and psychologically without enforcing an accepted norm.[37] Especially in the latter form, the target does not have avenues for redress. As Amanda Koontz, a professor of sociology from the University of Central Florida, explained: "[Cancelling someone] puts great responsibility on an individual, and it does not [always] encourage actual societal change."[38]

This article is not casting judgment either way on some of the targeted speech online. Some behavior and speech targeted online that leads to someone being canceled is condemnable. This article considers more attempts at public shaming online that are defamatory or otherwise reach the threshold of a tort offense, and what remedy might be available to allegedly defamed individuals if Section 230 immunity was removed.

Removing Section 230 Immunity Would Allow a Cause of Action Against Twitter

Twitter cannot be held liable when the weight of the internet comes down on a person because of Section 230 immunity. For example, in *Brikman v. Twitter, Inc.*, Birkman sought to hold Twitter accountable for allegedly defamatory posts by an anonymous third-party user. Twitter refused to take down the posts when asked, citing that the tweets did not violate its community standards. The New York trial court found the Section 230 immunity preempted the plaintiff's claims.[39] Citing the universal trend of construing Section 230 immunity broadly, the Court ruled that Twitter was an ISP or website, which qualified for Section 230 immunity, and thus, was not the publisher of the allegedly libelous tweets.[40] Twitter would have had to "directly and materially contributed" to the tweets to be held liable.[41] Even though this is a trial court ruling, it is reflective of the larger trend of courts rarely holding websites liable for third-party content.[42] This inability to hold Twitter liable for an anonymous individual's tweets allows Twitter to avoid liability when someone is being canceled, even if the tweets are defamatory or rise to the threshold of another tort, like intentional infliction of emotional distress.[43]

That being the case, someone who is canceled online has only a cause of redress against the individuals who tweeted what might well be defamatory. Because Twitter and other social media platforms make gathering and sharing information simple and instantaneous, defamatory content can be shared quickly by many and become viral.[44] Defamed individuals would likely prefer a cause of action against Twitter, not only because Twitter has deeper pockets than the vast majority of its individual users,[45] but, because of Twitter's algorithm and control of the content on its site, Twitter alone may be able to suppress a provocative but defamatory message from being widely re-shared. In contrast, it is impractical to sue each user in individual actions for re-sharing the defamatory content.[46] If Twitter was incentivized to remove tweets that were defamatory because of the threat of liability, it would likely be more

discerning of potentially defamatory tweets and would be more likely to be responsive to take-down requests out of the fear of being sued.[47]

One might argue that removing Section 230 immunity would not have a drastic effect on how Twitter operates because defamation actions, both in and outside cyberspace, are rarely successful.[48] Social media libel lawsuits are especially prone to failure because the internet's informality renders the speech more likely to be interpreted by a court as opinion or hyperbole rather than an actionable statement of fact.[49] However, even the threat of lawsuits can compel a website to change its behavior to avoid early stage litigation costs, even if the suit would likely be dismissed early on.[50] Even frivolous lawsuits may make Twitter more responsive to take-down requests.

With social media sites incentivized to remove allegedly defamatory content, the internet would see a loss of online mobs of users who rapidly share defamatory posts and call for the cancellation of an individual. Eliminating or revising Section 230 immunity could significantly change how Twitter functions and how we publicly shame people online.

Jordan Doll is a third-year student at Benjamin N. Cardozo School of Law and an Associate Editor of the Arts & Entertainment Law Journal.

1. 47 U.S.C. § 230 (1996).
2. Brikman v. Twitter, Inc., No. 19-cv-5143 (RPK) (CLP), 2020 U.S. Dist. LEXIS 170798, 1-2 (E.D.N.Y. Sep. 17, 2020) (finding a plaintiff does not have a cause of action against Twitter for defamation for the allegedly defamatory tweets of an anonymous user).
3. Sarah Morrison, Section 230, the internet free speech law Trump wants to repeal, explained, VOX (Oct. 6, 2020 1:19 pm), https://www.vox.com/recode/2020/5/28/21273241/section-230-explained-trump-social-media-twitter-facebook
4. Lyrissa Barnett Lidsky & RonNell Anderso Jones, Reasonable Readers and Unreasonable Speakers: Libel Law in a Networked World of Reasonable Readers and Unreasonable Speakers: Libel Law in a Networked World, 23 Va. J. Soc. Pol'y & L. 155, 160 (2016).
5. The tort of libel is defined by Black's Law Dictionary as "[a] defamatory statement expressed in a fixed medium, esp[ecially] writing but also a picture, sign, or electronic broadcast. Libel, Black's Law Dictionary (11th ed. 2019).
6. Lyrissa Barnett Lidsky, Silencing John Doe: Defamation & Discourse in Cyberspace, 49 Duke L.J. 855, 866 (2000).
7. Id. at 872.
8. Michal Lavi, The Good, the Bad, and the Ugly Behavior, 40 Cardozo L. Rev. 2597, 2602-06 (2019).
9. Anhana Susrla, Algorithms are making cancel culture even worse, Fast Company (Feb. 3. 2020), https://www.fastcompany.com/90458174/hate-outrage-and-cancel-culture-are-snowballing-thanks-to-this.

10. Nicole Dudenhoefer, Is Cancel Culture Effective? Pegasus: The Magazine of the University of Central Florida (accessed Apr. 4, 2021), https://www.ucf.edu/pegasus/is-cancel-culture-effective/.
11. Lavi, supra note 8, at 2613-14.
12. Id. at 2636.
13. The PACT Act and Section 230: The Impact of the Law that Helped Create the Internet and an Examination of Proposed Reforms for Today's Online World, Cong. Hearing, 5 (2020) (testimony of Rep. Chris Cox) ("Without Section 230, millions of American websites–facing unlimited legal liability for what their users create–would not be able to host user-generated content at all.").
14. Barbara Ortutay, AP Explains: The rule that made the modern internet, Associated Press (Oct. 28, 2020), https://apnews.com/article/what-is-section-230-tech-giants-77bce70089964c1e6fc87228ccdb0618.
15. Zeran v. Am. Online, Inc., 129 F.3d 327, 330-31 (4th Cir. 1997).
16. Stratton Oakmont v. Prodigy Servs. Co., INDEX No. 31063/94, 1995 N.Y. Misc. LEXIS 229, 13, 3 (Sup. Ct. May 24, 1995).
17. Id. at 10.
18. See Jordan Doll, Note, Section 230(e)(2) Immunity and Unpredictability: How a Broad Reading of the CDA's Immunity Carve-out Runs Contrary to the Unfettered Development of the Internet, 39 Cardozo Arts & Ent. L.J. (2022).
19. Joshua Dubnow, Ensuring Innovation as the Internet Matures: Competing Interpretations of the Intellectual Property Exception to the Communications Decency Act Immunity, 9 Nw. J. Tech. & Intell. Prop. 297, 300 (2010).
20. Melissa A Troiano, Bankruptcy 2.0(05): Chapters, Changes, And Challenges: Comment: The New Journalism? Why Traditional Defamation Laws Should Apply to Internet Blogs, 55 Am. U.L. Rev. 1447, 1455 (2006).
21. Zeran, 129 F.3d. at 333.
22. Jessica Melugin, Preserving Section 230 is Key to Maintaining the Free and Open Internet, Competitive Enterprise Institute (June 23, 2021), https://cei.org/studies/preserving-section-230-is-key-to-maintaining-the-free-and-open-internet/.
23. George Fishback, How the Wolf of Wall Street Shaped the Internet: A Review of Section 230 of the Communications Decency Act, 28 Tex. Intell. Prop. L.J. 275, 285 (2020).
24. Lavi, supra note 8, at 2636.
25. Zeran, 129 F.3d. at 333.
26. 47 U.S.C. § 230(a) ("The Internet and other interactive computer services have flourished, to the benefit of all Americans, with a minimum of government regulation.").
27. Casey Newton, Everything You Need to Know About Section 230, The Verge (Dec. 29, 2020, 4:50 PM), https://www.theverge.com/21273768/section-230-explained-internet-speech-law-definition-guide-free-moderation.
28. Lidsky, supra note 6, at 863.

29. Id.
30. Id. at 863-64.
31. Aja Romano, Why We Can't Stop Fighting About Cancel Culture, Vox (Aug, 25, 2020 12:03 PM), https://www.vox.com/culture/2019/12/30/20879720/what-is-cancel-culture-explained-history-debate.
32. The term canceled originated in Black culture and was appropriated by White mainstream culture, which distorted the term's use to something irrevocably associated with political correctness in the minds of many Americans. Before its appropriation, young Black people used the term canceled as sincere calls to consciousness and action before a distorted version of the word caught the attention of conservative media. See Clyde McGrady, The strange journey of 'cancel,' from a Black-culture punchline to a White-grievance watchword, The Washington Post (Apr. 2, 2021, 6:00 AM), https://www.washingtonpost.com/lifestyle/cancel-culture-background-black-culture-white-grievance/2021/04/01/2e42e4fe-8b24-11eb-aff6-4f720ca2d479_story.html.
33. Id.
34. Id.
35. Emily A. Vogels, Monica Anderson, Margaret Porteus, Chris Baronavski, Sara Atske, Colleen McClain, Brooke Auxier, Andrew Perrin, & Meera Ramshankar, Americans and 'Cancel Culture': Where Some See Calls for Accountability, Others See Censorship, Punishment, Pew Research Center (May 19, 2021), https://www.pewresearch.org/internet/2021/05/19/americans-and-cancel-culture-where-some-see-calls-for-accountability-others-see-censorship-punishment/.
36. Lavi, supra note 8, at 2616.
37. Id. at 2619-20.
38. Nicole Dudenhoefer, Is Cancel Culture Effective? Pegasus: The Magazine of the University of Central Florida (accessed Apr. 4, 2021), https://www.ucf.edu/pegasus/is-cancel-culture-effective/.
39. Brikman, No. 19-cv-5143 (RPK) (CLP), 2020 U.S. Dist. LEXIS 170798, at 1-2.
40. Id. at 5-7.
41. Id. at 8, quoting Force v. Facebook, Inc., 934 F.3d 53, 70 (2nd Cir. 2019).
42. Lavi supra, note 8, at 2636-37.
43. Id. at 2621.
44. Lidskey and Jones, supra note 4, at 157.
45. Lidskey, supra note 6, at 859.
46. Id. at 864.
47. Zeran, 129 F.3d at 333 ("Because service providers would be subject to liability only for the publication of information, and not for its removal, they would have a natural incentive simply to remove messages upon notification.").
48. See Lidskey and Jones supra, note 4 at 185; Ganske v. Mensch, 480 F Supp 3d 542 (SDNY 2020) ("[Where] the context is Twitter, an internet forum, 'New York courts have consistently protected statements made in online forums as statements of opinion

rather than fact'") (citing Bellavia Blatt & Crossett, P.C. v. Kel & Partners, LLC, 151 F. Supp. 3d 287, 295 (E.D.N.Y. 2015)).

49. See Redmond v. Gawker Media, No. CGC-11-508414, 2012 WL 3243507, at 6 (Cal. Ct. App. Aug. 10, 2012); Lidskey and Jones, supra note 4 at 166-67.
50. Lidskey and Jones supra, note 4 at 904-05.