

2023

The TikTok Algorithm Is Good, But Is It Too Good? Exploring the Responsibility of Artificial Intelligence Systems Reinforcing Harmful Ideas on Users

Julianne Gabor

The Catholic University of America, Columbus School of Law

Follow this and additional works at: <https://scholarship.law.edu/jlt>



Part of the [Agency Commons](#), [Communications Law Commons](#), [Computer Law Commons](#), [Consumer Protection Law Commons](#), [Internet Law Commons](#), [Litigation Commons](#), [Other Law Commons](#), [Science and Technology Law Commons](#), [Social Welfare Law Commons](#), and the [Torts Commons](#)

Recommended Citation

Julianne Gabor, *The TikTok Algorithm Is Good, But Is It Too Good? Exploring the Responsibility of Artificial Intelligence Systems Reinforcing Harmful Ideas on Users*, 32 Cath. U. J. L. & Tech 109 (2023).

Available at: <https://scholarship.law.edu/jlt/vol32/iss1/6>

This Comments is brought to you for free and open access by Catholic Law Scholarship Repository. It has been accepted for inclusion in Catholic University Journal of Law and Technology by an authorized editor of Catholic Law Scholarship Repository. For more information, please contact edinger@law.edu.

THE TIKTOK ALGORITHM IS GOOD, BUT IS IT TOO GOOD? EXPLORING THE RESPONSIBILITY OF ARTIFICIAL INTELLIGENCE SYSTEMS REINFORCING HARMFUL IDEAS ON USERS

Julianne Gabor *

The rise of technology has caused society to experience results that were unanticipated twenty-five years ago. The rate of the spread of information has drastically increased, and people all over the world have come together to connect over niche interests. People are now able to readily voice their opinions on a greater stage than just within their community, possibly reaching celebrities or world leaders. The internet has led to social responsibility for many corporations because of the public's ability to raise issues that have a direct effect on consumerism.¹ Overall, there are many positives to the internet.

However, there is a dark side hidden from the public eye. Online bullying ran rampant in the 2010s and continues today.² However, the underlying epidemic

* Juris Doctor candidate, Columbus School of Law, 2024; Bachelor of Arts in Economics and Government & Politics, University of Maryland, College Park, 2021; *The Catholic University Journal of Law and Technology*, Lead Articles Editor, 2023–2024, Associate Editor 2022–2023. Thank you to my family and friends for their love and support throughout law school. In addition, thank you to the Volume 32 Executive Board and Associate Editors for their hard work on revising my article and commitment to the success of JLT.

¹ See Charles El-Zeind, *6 Easy Ways to Use Social Media and Corporate Social Responsibility*, SUSTAINABLE BUS. TOOLKIT (July 31, 2021), <https://www.sustainablebusinesstoolkit.com/social-media-and-corporate-social-responsibility/>.

² See Zia Muhammad, *New Research Shows 45% of Gen Z Faces Cyberbullying Online*, DIGIT. INFO. WORLD (Sept. 12, 2022), <https://www.digitalinformationworld.com/2022/09/new-research-shows-45-of-gen-z-faces.html>.

is not caused by individuals, but by the social media platform itself. Many sites have horrendous images of self-harm and posts encouraging others to cause harm to themselves.³ Apps are able to show a user unwanted content and provide little recourse to stop it.⁴ While some find comfort in knowing that others are also struggling, constantly viewing this content can push some into a worse mental state than before.⁵ The public usually puts the burden on the user to simply stop using the app, but this understates the issue.⁶ Users may not realize the harm that is being caused or how it happens, so completely withdrawing from the app when they still get some benefit does not seem like a solution.⁷ In addition, there is public pressure to be part of social media and to be in the know of every trend.⁸ Social media platforms put their head in the sand when there is an issue because they are not burdened with the liability.

TikTok rocketed to popularity in the early 2020s and has been dominant ever since.⁹ It is known for being different than other social media platforms because users put their viewing experience into the hands of the algorithm—which is notorious for “being inside a user’s head”—by allowing the app to generate content for them, as opposed to other platforms where a user must follow other users to see content.¹⁰ It allows algorithms to analyze users most vulnerable thoughts and show them content in this area, with the mission to find what content is addictive enough to keep them on the site and increase their watch time.¹¹ This predatory practice is especially dangerous because users are not aware that this is the main goal of the app.¹² By putting fate in the hands of the algorithm, users may lose control of what they view, which is worsened by the lack of user input, such as the ineffective “Not Interested” button.¹³

³ See Julie Jargon, *She Tried to Block Eating-Disorder Content on TikTok. It Still Pops Up Daily.*, WALL ST. J. (Sept. 24, 2022), <https://www.wsj.com/articles/she-tried-to-block-eating-disorder-content-on-tiktok-it-still-pops-up-daily-11663982648>.

⁴ See *id.*

⁵ See *id.*

⁶ See Alyssa Young, *Pressures and Dangers of Social Media: A Personal Story*, LEE HEALTH, <https://www.leehealth.org/health-and-wellness/healthy-news-blog/mental-health/pressures-and-dangers-of-social-media-a-personal-story> (last visited Dec. 9, 2023).

⁷ See Jargon, *supra* note 3.

⁸ See *id.*

⁹ See Drew Harwell, *How TikTok Ate the Internet*, WASH. POST (Oct. 14, 2022, 5:00 AM), <https://www.washingtonpost.com/technology/interactive/2022/tiktok-popularity/>.

¹⁰ *Investigation: How TikTok’s Algorithm Figures Out Your Deepest Desires*, WALL ST. J. (July 21, 2021), <https://www.wsj.com/video/series/inside-tiktoks-highly-secretive-algorithm/investigation-how-tiktok-algorithm-figures-out-your-deepest-desires/6C0C2040-FF25-4827-8528-2BD6612E3796>.

¹¹ *Id.*

¹² *Id.*

¹³ *Id.*

There have been many suits against social media platforms claiming they are addictive and allow harmful content.¹⁴ However, TikTok poses a new problem. Should social media platforms face liability for pushing users into niche, harmful content areas that are less regulated by design, when they are aware that a user is struggling with mental health and therefore susceptible to adopting the harmful ideas being shown?¹⁵

This situation invites parallels to *Commonwealth v. Carter*, a case in which an individual was charged with involuntary manslaughter for perpetuating ideas of suicide to someone who was mentally vulnerable, leading the victim's will to be overpowered and ultimately resulting in the victim's suicide.¹⁶ Both situations prey on one's vulnerabilities, but TikTok does it on a much larger scale.¹⁷ Considering the social harms it poses, it is hard to see why TikTok is able to hide behind a liability shield under the law, when producers of other products, such as firearms and cigarettes, have been held responsible despite similar non-liability statutes.¹⁸ Due to the large-scale social harms it poses, TikTok, along with other social media platforms, should be held liable for the declining mental health of its users when this decline is caused by the platform's unhealthy algorithm. To accomplish this, Section 230 of the Communications Decency Act, the statute which bars liability for social media platforms, should be amended to include a reasonable person standard, which would force TikTok to reflect on how it screens and processes content on its site.¹⁹

First, I will evaluate TikTok's algorithm, showing how it poses more harm than the traditional social media platform. TikTok has not publicly released its process for recommending videos to users, but a study by the *Wall Street Journal*

¹⁴ See Ashley Gold, *Lawsuit Pushes Addiction Case Against Social Media Firms*, AXIOS (Jan. 11, 2023), <https://www.axios.com/2023/01/11/social-media-addiction-lawsuits-case>.

¹⁵ *Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires*, *supra* note 10 (noting that niche content areas are less regulated than popular topic areas because there are less users reporting harmful content and less oversight from the app and explaining how the users in the content area may be promoting the harmful content themselves).

¹⁶ See generally *Commonwealth v. Carter*, 52 N.E.3d 1054 (Mass. 2016).

¹⁷ See generally *id.*; see generally *Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

¹⁸ See *Gun Industry Immunity*, GIFFORDS L. CTR., <https://giffords.org/lawcenter/gun-laws/policy-areas/other-laws-policies/gun-industry-immunity/> (last visited Nov. 26, 2023); see David Goguen, *Tobacco and Cigarette Litigation: History & Trends*, NOLO, <https://www.nolo.com/legal-encyclopedia/tobacco-litigation-history-and-development-32202.html> (last visited Nov. 26, 2023).

¹⁹ Communications Decency Act, 47 U.S.C. § 230 (2023); Jeffery Johnson, *Reasonable Person Standard: Legal Definition & Examples*, FORBES (Sept. 19, 2022), <https://www.forbes.com/advisor/legal/personal-injury/reasonable-person-standard/> (discussing how the "reasonable person" standard is an objective standard which asks what a reasonable person should have known and what would have been reasonable to do in that situation).

theorizes how the “For You Page” works.²⁰ I will show how this model is more harmful than other social media platforms due to the lack of user control over content when compared with viewing only content that a user chooses to follow. Second, to show the social harm it poses to users, I will draw parallels between the TikTok algorithm and *Commonwealth v. Carter*, the case where an individual was convicted of involuntary manslaughter because the coercive power of her verbal actions overpowered the will to live of another person.²¹ Finally, I will discuss Section 230 of the Communications Decency Act, which protects social media platforms from liability for content posted to their sites, to show how social media platforms have been shielded from litigation despite years of controversy in the public eye,²² despite the fact that we have previously seen other types of commodities being held liable for their wrongs despite similar protective statutes.²³ Considering TikTok’s parallels to *Carter*, and its applicability to current Section 230 case law, TikTok should be held liable through judicial action—just as cigarette and firearm manufacturers have been held liable despite similar provisions.

I. HOW DOES TIKTOK KNOW EXACTLY WHAT I’M THINKING? ANALYZING THE MECHANICS OF TIKTOK’S ALGORITHM

TikTok’s parent company, ByteDance, has not publicly released how it recommends videos to users, but a 2021 study by the *Wall Street Journal* collected data based on bots acting as users and then utilized that data to create theories on how the “For You Page” works.²⁴ A bot is a “software program that performs automated, repetitive, pre-defined tasks” that “imitate or replace” the actions or behaviors of humans at a faster pace.²⁵ The *Wall Street Journal* created over 100 bots to pose as users on the app, and those bots collectively viewed hundreds of thousands of videos.²⁶ TikTok has not said much about its algorithm, but it has stated that what a user shares, likes, and who they follow

²⁰ *Investigation: How TikTok’s Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

²¹ *See Commonwealth v. Carter*, 52 N.E.3d 1054, 1064 (Mass. 2016).

²² 47 U.S.C. § 230(c)(2)(B).

²³ *See Gun Industry Immunity*, *supra* note 18; *see Goguen*, *supra* note 18.

²⁴ *Investigation: How TikTok’s Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

²⁵ *What Are Bots? - Definition and Explanation*, KASPERSKY, <https://www.kaspersky.com/resource-center/definitions/what-are-bots> (last visited Nov. 26, 2023).

²⁶ *Investigation: How TikTok’s Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

plays a role in what the algorithm recommends.²⁷ However, the *Wall Street Journal* bots showed that how long a user lingers over content, or how long they stay and watch or rewatch a video, is the most important factor in the types of videos shown to users.²⁸ Others have confirmed the theory that retention time is the single most important factor to the algorithm.²⁹

Each bot took on a different alias, with its own location, age, and interests.³⁰ Each was given a list of interests that was not shared with the app and was tasked with expressing their interests through rewatching or pausing videos with similar hashtags.³¹ Much discussion focused on username @kentucky_96, a bot from Kentucky, that was assigned an age of twenty-four years old and the interests of sadness and depression.³² In the study, @kentucky_96 acted like a human using the app.³³ In the beginning of the bot's journey, TikTok showed @kentucky_96 popular videos on the app vetted by moderators to see if the user is religious, likes dance videos, or has been feeling sad.³⁴ The purpose of this process is to find the unexpressed interests of the user in order to create a page especially for them.³⁵ Less than three minutes into @kentucky_96's use of the app, after viewing its fifteenth video, it watched a thirty-five second video twice, thus giving the algorithm a clue that it was feeling down based on the author, audio track, caption, and hashtags of the video.³⁶ Four minutes, or twenty-three videos, later, it was shown a video that had #sad and content about a breakup, while still being shown popular videos on the app and location-specific videos about Kentucky.³⁷ At video fifty-seven, there was a video about hurt feelings with #heartbreak and at video sixty, a video about emotional pain with #pain.³⁸ By this point, the algorithm thought that @kentucky_96 wanted relationship content, but at video eighty, or fifteen minutes of watch time, the user was shown relationship-specific content but did not seem interested.³⁹ It was not until @kentucky_96 paused on a video with #mentalhealthmatters, #depression, and

²⁷ *Id.*

²⁸ *Id.*

²⁹ See Andrew Hutchinson, *TikTok Shares New Insights into Key Engagement and Retention Behaviors in the App [Infographic]*, SOCIAL MEDIA TODAY (July 26, 2022), <https://www.socialmediatoday.com/news/tiktok-shares-new-insights-into-key-engagement-and-retention-behaviors-in-t/628174/> (July 26, 2022).

³⁰ *Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

³¹ *Id.*

³² *Id.*

³³ *Id.*

³⁴ *Id.*

³⁵ *Id.*

³⁶ *Id.*

³⁷ *Id.*

³⁸ *Id.*

³⁹ *Id.*

#anxiety, while swiping past videos about missing an ex, that the algorithm solidified the user's interests.⁴⁰ By thirty-six minutes of watch-time, or 224 videos, videos about depression and mental health outnumbered other interests previously thought to be connected to the user.⁴¹ After this turning point, ninety-three percent of the videos shown to @kentucky_96 were related to depression.⁴² A spokesperson for TikTok said that the remaining videos were to help users to discover different content, but in this case, videos for discovery were rare, with most of the non-depression related content being advertisements.⁴³ TikTok has said that the bots are not representative of real users because humans have diverse interests, but even bots assigned diverse interests eventually wound up in a rabbit hole of content.⁴⁴ The bots only escaped the niche content areas, known as rabbit holes, when their interests were changed; they were told to stop watching videos by programmers controlling the bots.⁴⁵ However, many bots were still pushed into rabbit holes, with general interests becoming narrower and more personalized as watch-time increased.⁴⁶ One bot was assigned a general interest in politics and was eventually shown election conspiracy and QAnon videos.⁴⁷ A TikTok spokesperson said that computer analysis and human moderators catch the most inappropriate content with the help of user reports, but as users scroll deeper into niche interests, they are more likely to see harmful content that violates the terms of service.⁴⁸ Overall, TikTok learned most bots' interests in less than two hours and some in less than forty minutes of watch time.⁴⁹

Guillaume Chaslot, founder of Algotransparency, and a former Google engineer, said that TikTok is different than YouTube, another highly criticized algorithm, which Chaslot worked on, because of its power and ability to learn vulnerability much faster.⁵⁰ Chaslot said that in a case like @kentucky_96,

⁴⁰ *Id.*

⁴¹ *Id.*

⁴² *Id.*

⁴³ *Id.*

⁴⁴ *Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires*, *supra* note 10; see Saumya Kaila, *The Psychology of Why We Fall Down Internet Rabbits Holes*, THE SWADDLE (Mar. 10, 2022), <https://theswaddle.com/the-psychology-of-why-we-fall-down-internet-rabbit-holes/> (explaining that rabbit hole means to get "interest[ed] in something to the point of distraction—usually by accident, and usually to a degree that the subject in question might not seem to merit").

⁴⁵ *Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

⁴⁶ *Id.*

⁴⁷ *Id.*

⁴⁸ *Id.*

⁴⁹ *Id.*

⁵⁰ *Id.*

depression content is seen as helpful to the algorithm to foster engagement on the app; it pushes more and more depression content, which then forces users toward more extreme content in order to increase watch time on the app.⁵¹ Furthermore, he explained that the app finds your vulnerabilities and what makes you click or interact, but that this engagement does not necessarily equate to liking or enjoying the content.⁵² It simply shows users what is most likely to make them stay on the app.⁵³

TikTok emphasized the use of the “Not Interested” function, but Chaslot argues that this is not enough to protect users.⁵⁴ Even if this function is used, content creators on the app have found ways to evade the restriction.⁵⁵ For example, TikTok has expanded its ban on videos related to eating disorders by blocking searches for hashtags related to the content and creating a tool to flag certain hashtags.⁵⁶ Instead, when users search #anorexic, they are shown a message from the National Eating Disorders Association.⁵⁷ However, users have misspelled common eating disorder terms to avoid the blockages.⁵⁸ Viewers can still see videos with #anotexic and #orthoreixa—both misspellings of eating disorder diagnoses—and access the same content as before the blocks.⁵⁹ A teen who had previously struggled with eating disorders blocked accounts promoting eating disorders, reported videos promoting eating disorders, and used the “Not Interested” function, yet these videos continued to show up on her “For You Page”; the videos may not promote extreme ideas anymore but still have a common message.⁶⁰ Dr. Rachel Fortune, a consulting eating-disorder physician for Newport Healthcare, stated that users must “make a conscious effort to run away from it and actively reject the content.”⁶¹

The burden being placed on users to run away from harmful content is an especially big issue when the potentially harmful message is hidden a few seconds into a video. At first glance, a user is unable to comprehend what the video may encompass; it is necessary to watch a few seconds to recognize the content of the video and make the decision to either stay and watch or swipe off. As previously mentioned, the biggest influencing factor of the algorithm is retention time.⁶² Because a user has to watch a video to see its content, it creates

⁵¹ *Id.*

⁵² *Id.*

⁵³ *Id.*

⁵⁴ *Id.*

⁵⁵ See Jargon, *supra* note 3.

⁵⁶ *See id.*

⁵⁷ *See id.*

⁵⁸ *See id.*

⁵⁹ *See id.*

⁶⁰ *See id.*

⁶¹ *See id.*

⁶² *Investigation: How TikTok’s Algorithm Figures Out Your Deepest Desires, supra*

a never-ending cycle of continually seeing harmful content, swiping off, and seeing it pop up again because the algorithm registers that the user watched for a few seconds and then continues to show the user videos with similar content.⁶³ Because the “Not Interested” and blocking functions have been shown to provide little recourse for users, most will be forced to make a cognizant effort to swipe away from these videos as fast as possible or bear the risk of seeing them again.⁶⁴ For some users, this will require great strength to resist the dark content rabbit hole because the only two choices available are either swipe away once they recognize the video content, or stay and watch, both of which inevitably increase that type of content on your “For You Page” due to retention time.⁶⁵

TikTok has revolutionized the way social media works by creating an incredible algorithm that is able to feed users exactly what they are looking for, or expose them to new content that is related to what they already watch.⁶⁶ However, for some users struggling with mental health, this creates an inescapable burden to run away from content that is sensitive for that user.⁶⁷ The purpose of TikTok’s algorithm is to find content that keeps users on the app, usually resulting in users being pushed into niche content areas.⁶⁸ Because TikTok’s content is composed of videos as opposed to photos like other platforms, a user must watch a few seconds to process what the video is to make the decision to swipe away or not. This engagement time means the videos will continue to play, with few ways to proactively prevent the content from coming up.⁶⁹ TikTok creates a harmful environment for users due to the badgering nature of the algorithm and lack of self-help mechanisms available to users; because users can’t engage in self-help, they will have to continue to swipe through harmful content which will only increase their watch time and lead these types of videos to plague their “For You Page.”⁷⁰ This creates parallels to the actions of the defendant in *Carter*, which may lead users down a path of harm to themselves or others.⁷¹

As “influencing” as a job continues to gain popularity, so do influencers’ rights as independent contractors.⁷² Many platforms are starting to pay

note 10.

⁶³ *Id.*

⁶⁴ *Id.*

⁶⁵ *See id.*

⁶⁶ *Id.*

⁶⁷ *Id.*

⁶⁸ *Id.*

⁶⁹ *Id.*

⁷⁰ *See generally* Jargon, *supra* note 3.

⁷¹ *See generally* Commonwealth v. Carter, 52 N.E.3d 1054 (Mass. 2016).

⁷² Brent Barnhart, *TikTok Creator Fund: How It Works and What to Expect*, SPROUT

individuals who excel on the app and create content that is popular.⁷³ TikTok has created the “Creator Fund” which “aims to provide financial support to top-tier content creators.”⁷⁴ The Creator Fund is an “official fund established by the app to compensate creators for their content” and “‘reward’ users for their ‘incredible’ TikTok videos and creativity.”⁷⁵ Initially, TikTok made a \$200 million investment in the Creator Fund, with plans to “grow the Fund to over \$1 billion within three years.”⁷⁶ Compensation is based on views and engagement, and aims to capture different types of people in different content areas.⁷⁷ The Fund is similar to the algorithm because it “encourage[s] even more activity and time spent on the app,” with 56 percent of users claiming they are “inspired to create videos after seeing compelling creator content.”⁷⁸ It also discourages the creators from jumping to other apps to create content by paying them through the Creator Fund, thus deterring “creators from seeking out agencies or influencer contracts independently.”⁷⁹ To qualify for the Creator Fund, a user must produce original content, be in good standing, have at least 10,000 followers, 100,000 views in the previous thirty days, and be at least eighteen years old.⁸⁰ The algorithm gives virtually anyone the potential to gain popularity on TikTok by connecting users with niche content area followers.⁸¹ TikTok also allows individual users to give other users a “virtual gift” as well as a subscriber program for live streamers.⁸² TikTok also recently launched a contextual-advertising program, TikTok Pulse.⁸³ This program allows brands to “buy ads alongside the top 4% of content in different categories,” with “50% of the revenue split with the creator whose videos appeared before the in-feed ad.”⁸⁴ This program is harder to qualify for, requiring 100,000 followers.⁸⁵ Still, any user has the potential to monetize their content on the app.

Getting paid to be popular has its drawbacks, with many “rage-baiting” to

SOCIAL (June 28, 2022), <https://sproutsocial.com/insights/tiktok-creator-fund/>.

⁷³ Kimberlee Meier, *Top 7 Social Media Platforms That Pay Content Creators in 2022*, VIMEO BLOG (Mar. 31, 2022), <https://vimeo.com/blog/post/social-media-pay/>.

⁷⁴ Barnhart, *supra* note 72.

⁷⁵ *Id.*

⁷⁶ *Id.*

⁷⁷ *Id.*

⁷⁸ *Id.*

⁷⁹ *Id.*

⁸⁰ Dan Whateley, *How Much Money TikTokers Make, According to Top Creators*, BUS. INSIDER, <https://www.businessinsider.com/how-much-money-tiktokers-make-real-creator-examples-2021-7> (Sept. 14, 2023).

⁸¹ *Id.*

⁸² *Id.*

⁸³ *Id.*

⁸⁴ *Id.*

⁸⁵ *Id.*

grow their views because “it is the most effective way to go viral.”⁸⁶ This is confirmed by social media anthropologists, who say “it’s a tried-and-true technique to grow followers in an attention-based economy where generating engagement, positive or negative, is lucrative digital currency.”⁸⁷ Creators themselves have said that they have noticed an increase in rage-baiting.⁸⁸ TikTok’s algorithm will push this content to users if it will keep them on the app because “[s]ocial media platforms don’t care if the message is uplifting or toxic. As long as people are interacting with it, platforms will spread it even further.”⁸⁹ While some acknowledge rage-baiting as their growth strategy, many deny the claim.⁹⁰ Researchers have found that users “shared disgust and righteous indignation more often than content that elicited emotions like joy and sadness.”⁹¹ Algorithms play into this phenomenon because it “benefits the platforms to give you what you want and what you don’t want.”⁹² This content can include content that is “racist and hateful and misinformed and conspiratorial in nature” because this is what enraptures viewers.⁹³ While the monetization of content creators is a positive for many, there are also drawbacks that can create an uptick of harmful content on the platform that overwhelms the more positive content.

A. Section 230

Section 230 of the Communications Decency Act (“CDA”) states that “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.”⁹⁴ Sections 230(c)(1) and (2), collectively known as the “Good

⁸⁶ Tanya Chen, *On TikTok, Creators Farm Rage to Get Clicks and Make Money. But It Can be a Fast Race to the Bottom.*, INSIDER (Dec. 8, 2022), <https://www.insider.com/why-creators-rage-farm-on-tiktok-and-why-its-on-the-rise-2022-12>.

⁸⁷ *Id.*

⁸⁸ *Id.*

⁸⁹ *Id.*

⁹⁰ *Id.*

⁹¹ *Id.*; see Helen A.S. Popkin, *As Your Social Media Emotions Go Viral, Anger Spreads the Fastest*, NBC NEWS (Sept. 17, 2013), <https://www.nbcnews.com/technology/your-social-media-emotions-go-viral-anger-spreads-fastest-4b11186087>.

⁹² Chen, *supra* note 86.

⁹³ *Id.*

⁹⁴ 47 U.S.C. § 230(c)(1); VALERIE C. BRANNON & ERIC N. HOMES, CONG. RSCH. SERV., R46751, SECTION 230: AN OVERVIEW (2021) (covering TikTok, along with sites such as Facebook, because they fall under the definition of “interactive computer service” which is “any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server.” Also considering TikTok to be an “information content provider” which is “any person or entity that is responsible, in whole

Samaritan” provision, protect platforms from liability for harmful content; however, it does not force them to remove it and shields them if they choose to either keep or remove harmful content.⁹⁵ This has been interpreted as “creat[ing] a federal immunity to any cause of action that would make service providers liable for information originating with a third-party user of the service.”⁹⁶ There are exceptions for “suits brought under federal criminal law, intellectual property law, any state law ‘consistent’ with Section 230, certain privacy laws applicable to electronic communications, or certain federal and state laws relating to sex trafficking,” but these exceptions are rare.⁹⁷

The CDA was intended to “modernize the existing protections against obscene, lewd, indecent or harassing uses of telephone.”⁹⁸ Section 230 expresses that Congress wanted to “promote the free development of the internet, while also ‘remov[ing] disincentives’ to implement ‘blocking and filtering technologies’ that restrict ‘children’s access to . . . inappropriate online material.’”⁹⁹ Case law on Section 230 predates the internet boom, thus creating confusion on how the principles of Section 230 apply to new problems facing social media platforms. In 1991, *Cubby, Inc. v. CompuServe, Inc.* held that a website was not liable for content posted on it because it did not review any of the content before it was posted, and thus lacked knowledge of the content.¹⁰⁰ However, four years later, *Stratton Oakmont, Inc. v. Prodigy Services Co.* held that a website was liable because it moderated message boards and deleted comments for “offensiveness and ‘bad taste,’” so it acted like a publisher of third-party content with responsibility for the postings on the site.¹⁰¹ In response to this ruling, Congress passed Section 230 because it was worried that content would not be monitored, but in reality, this immunity reduced the incentive to remove harmful content even further.¹⁰²

Section 230 was last updated in 1996, so the impact of social media was not anticipated at the time.¹⁰³ Originally, lawmakers thought that platforms would

or in part, for the creation or development of information provided through the internet or any other interactive computer service.”).

⁹⁵ 47 U.S.C. § 230(c)(1)–(2).

⁹⁶ Neil Fried, *Why Section 230 Isn’t Really a Good Samaritan Provision*, DIGIT. FRONTIERS ADVOC., <https://digitalfrontiersadvocacy.com/blogs-and-op-eds/f/why-section-230-isnt-really-a-good-samaritan-provision> (last visited Nov. 25, 2023).

⁹⁷ BRANNON & HOLMES, *supra* note 94.

⁹⁸ S. REP. NO. 104–23, at 59 (1995).

⁹⁹ BRANNON & HOLMES, *supra* note 94, at 2.; *see* 47 U.S.C. § 230(b)(3)–(4).

¹⁰⁰ *Cubby, Inc. v. CompuServe, Inc.*, 776 F. Supp. 135, 141 (S.D.N.Y. 1991).

¹⁰¹ *Stratton Oakmont v. Prodigy Servs. Co.*, No. 31063/94, 1995 N.Y. Misc. LEXIS 229, at *10–11 (N.Y. Sup. Ct. May 24, 1995).

¹⁰² Michael D. Smith & Marshall Van Alstyne, *It’s Time to Update Section 230*, HARV. BUS. REV. (Aug. 12, 2021), <https://hbr.org/2021/08/its-time-to-update-section-230>.

¹⁰³ *Id.*

be self-interested in filtering content for safety concerns, but in the past decade, we have seen that allowing harmful posts has economic value for platforms, with the added benefit of little social backlash.¹⁰⁴ The rising spread of misinformation and conspiracy theories have led sites to advertise no restrictions on content being posted, leading to many extremists sharing ideas with few repercussions.¹⁰⁵

The creators of Section 230 theorized that it would act as a Good Samaritan statute by “remov[ing] the liability risk for someone with no duty of care, thereby eliminating the disincentive to act and creating a net increase in the likelihood of aid.”¹⁰⁶ However, Section 230 does the opposite by removing a platform’s duty of care to its users, making the platform *less* likely to protect its users.¹⁰⁷ The public, lawmakers, and interest groups have called to amend Section 230 to require application of a reasonable duty of care to internet sources.¹⁰⁸ The current era of internet exceptionalism, creates the “notion that the internet is so unique, yet so delicate, that we must shield it from ordinary legal mechanisms that promote accountability,” and is embodied by Section 230 because “courts have applied [it] in a way that largely immunizes online platforms from the obligation that most businesses have to take reasonable steps to prevent the use of their services from causing harm.”¹⁰⁹

Traditionally, under common law principles, businesses have a “duty to take reasonable steps to not cause harm . . . [creating] an affirmative obligation in certain circumstances for a business to prevent one party using the business’s services from harming another party.”¹¹⁰ If these principles were to be applied to the internet, platforms would be liable for fostering an unsafe environment or failing to prevent users from harming each other, thus removing the “current misincentive regarding content moderation,” which is radically different from the current statutes.¹¹¹ Platforms would still have discretion over how they run their site, with the freedom to test out to which degree they allow harmful

¹⁰⁴ *Id.*

¹⁰⁵ See Thomas L. Bourneuf, *Free Speech in the Era of False News: Section 230 Protections in the Modern Era*, N.Y.U. J.L. & BUS. (2021), <https://www.nyuylb.org/single-post/free-speech-in-the-era-of-false-news-section-230-protections-in-the-modern-era>.

¹⁰⁶ Fried, *supra* note 96.

¹⁰⁷ *Id.* (stating that “[o]rdinarily, businesses have a duty of care at common law to avoid unreasonably causing harm to their customers, as well as to take reasonable steps to prevent harms to them from other customers,” which conflicts with Section 230).

¹⁰⁸ Neil Fried, *The Myth of Internet Exceptionalism: Bringing Section 230 into the Real World*, AM. AFF. (May 5, 2021), <https://americanaffairsjournal.org/2021/05/the-myth-of-internet-exceptionalism-bringing-section-230-into-the-real-world/>.

¹⁰⁹ Fried, *supra* note 96.

¹¹⁰ *Id.*

¹¹¹ *Id.*

content.¹¹² However, if they cause “harm in a particular circumstance, [they] could be held accountable on the back end, just like most other unregulated businesses.”¹¹³ Instituting the duty of reasonable care would create a “back-end accountability” that “would also prompt more responsibility from the start, encouraging ‘responsibility by design.’”¹¹⁴ Online platforms will operate similarly to the free market and multiple online sources will be created for different ideas and views, still allowing the free flow of ideas among people interested in the topic.¹¹⁵

There is support for Congress to amend or repeal Section 230.¹¹⁶ In the 116th Congress alone, there were twenty-six bills proposed to amend Section 230 with a variety of suggestions, such as conditional immunity for narrower exceptions and limiting “publisher immunity” to urge sites to monitor their content closer.¹¹⁷ The National Telecommunications and Information Administration and the Department of Justice have submitted proposals highlighting issues of Section 230, such as First Amendment infringements.¹¹⁸ Specifically, they contemplate if Section 230 is completely repealed, whether “the First Amendment may nonetheless prevent private parties or the government from holding providers liable for publishing content . . . prevent[ing] some claims premised on decisions to host or restrict others’ speech, but its protections are likely less extensive than the current scope of Section 230 immunity.”¹¹⁹ Although many are pushing for change, little advances have been made in the legislative area.¹²⁰ However, we have seen judicial solutions in the past for products that cause societal harm, such as cigarettes and firearms.¹²¹ Due to the lack of traction in Congress, the key to creating liability for social media platforms may be in the judicial branch.

B. The Current Application of Section 230

We have seen an increase in claims surrounding Section 230 in the past few years. *Klayman v. Zuckerberg* shows how Section 230 claims apply to the current state of social media platforms.¹²² In *Klayman*, a user came across a page on Facebook with over 360,000 members titled “Third Palestinian Intifada” and

¹¹² *Id.*

¹¹³ *Id.*

¹¹⁴ *Id.*

¹¹⁵ *Id.*

¹¹⁶ *Id.*

¹¹⁷ BRANNON & HOLMES, *supra* note 94.

¹¹⁸ *Id.*

¹¹⁹ *Id.*

¹²⁰ *Id.*

¹²¹ *See Gun Industry Immunity*, *supra* note 18; *see Goguen*, *supra* note 18.

¹²² *See Klayman v. Zuckerberg*, 753 F.3d 1354, 1356 (D.C. Cir. 2014).

called for Muslims to rise up and kill Jewish people, with many similar pages calling for the same.¹²³ Facebook eventually removed the page.¹²⁴ The court held that the claim was entitled to dismissal because “the Communications Decency Act mandates dismissal if (i) Facebook is a ‘provider or user of an interactive computer service,’ (ii) the information for which Klayman seeks to hold Facebook liable was ‘information provided by another information content provider,’ and (iii) the complaint seeks to hold Facebook liable as the ‘publisher or speaker’ of that information,” and Facebook satisfied all three.¹²⁵ First, Facebook was held to be an interactive computer service because it provided users access to a computer server.¹²⁶ In this case, this was the “server that host[s] its social networking website,” because when users “browse the site and review the pages of other users, they do so by gaining access to information stored on Facebook’s servers.”¹²⁷ Klayman failed in his argument that Facebook should not qualify due to its ability to control posted content because this additional requirement was not written into the CDA.¹²⁸ Additionally, the court felt this argument was circular because Section 230 prohibits liability for actions taken to restrict harmful content, and it does not make sense to “say that interactive computer services must lack the capacity to police content when the Act expressly provides them with immunity for doing just that.”¹²⁹ Second, the claim that Facebook furthered the message of the content by not removing it failed because the argument rested on “information provided by another information content provider,” and not Facebook itself.¹³⁰ It is established that a “website does not create or develop content when it merely provides a neutral means by which third parties can post information of their own independent choosing online.”¹³¹ *Klayman* is an important interpretation of how Section 230 currently applies to social media platforms, something that was not anticipated at the time of the last amendment to the statute.

The Supreme Court recently vacated and remanded *Gonzalez v. Google LLC* and its application of Section 230 toward social media platforms facilitating crime to be reevaluated in light of *Twitter, Inc. v. Taamneh*, a companion case to *Gonzalez*.¹³² In *Twitter, Inc.*, the Supreme Court reversed the ruling of the

¹²³ *Id.* at 1355.

¹²⁴ *Id.*

¹²⁵ *Id.* at 1357.

¹²⁶ *Id.*

¹²⁷ *Id.*

¹²⁸ *Id.* at 1358.

¹²⁹ *Id.*

¹³⁰ *Id.*

¹³¹ *Id.*

¹³² *Gonzalez v. Google LLC*, 143 S. Ct. 1191, 1192 (2023).

Ninth Circuit and held that the petitioners did not correctly state a claim under the Anti-Terrorism Act.¹³³ The Supreme Court declined to address the application of Section 230, but stated that because the claims in *Gonzalez* were “materially identical” to *Twitter, Inc.*, the complaint “independent of § 230—states little if any claim for relief.”¹³⁴ This sidestep by the Supreme Court further complicates the modern application of Section 230. Some have suggested that the courts will “avoid making massive changes to the law given the complexities of platform content moderation and the highly disruptive ramifications that would accompany such a decision.”¹³⁵ However, this is precisely why such action is needed. As technology continues to advance, a clear stance is needed to protect consumers from the dangers of social media before it is too late to create a workable and effective doctrine.¹³⁶ Although *Gonzalez* has been remanded, it is still important to discuss the split between the Ninth Circuit and the Supreme Court’s refusal to address the issue.

The plaintiffs in *Gonzalez* were family members of victims of a terrorist attack and claimed that the attackers conspired with, and were aided and abetted by Google, Twitter, and Facebook.¹³⁷ They claimed that these platforms allowed ISIS to have a space to post recruitment content to gain new members and further its mission.¹³⁸ Additionally, they claimed that Google “placed paid advertisements in proximity to ISIS-created content and shared the resulting ad revenue with ISIS.”¹³⁹

The plaintiffs argued that the immunities of Section 230 did not apply to claims arising under the Anti-Terrorist Act (ATA) because the Act “permits private civil enforcement of counter-terrorism provisions that otherwise give rise to criminal liability, and § 230(e)(1) includes an exception providing that ‘[n]othing in this section shall be construed to impair the enforcement of . . . any . . . Federal criminal statute.’”¹⁴⁰ The court rejected this contention, agreeing with Google that the exception “extends only to criminal prosecutions . . . not to civil actions based on criminal statutes.”¹⁴¹ However, the court stated that this exception does not preclude the application of § 230(c)(1), known as the Good

¹³³ *Twitter, Inc. v. Taamneh*, 143 S. Ct. 1206, 1231 (2023).

¹³⁴ *Gonzalez*, 143 S. Ct. at 1192.

¹³⁵ F. Paul Pittman et al., *Supreme Court Declines to Reconsider Foundational Principles of Internet Platform Liability*, WHITE & CASE (June 15, 2023), <https://www.whitecase.com/insight-alert/supreme-court-declines-reconsider-foundational-principles-internet-platform-liability>.

¹³⁶ *See id.*

¹³⁷ *Gonzalez v. Google LLC*, 2 F.4th 871, 880 (9th Cir. 2021).

¹³⁸ *Id.*

¹³⁹ *Id.*

¹⁴⁰ *Id.* at 890.

¹⁴¹ *Id.*

Samaritan provision.¹⁴²

The plaintiffs' claim that Google did not do enough to block or remove content promoting ISIS was dismissed because it required Google to be treated as a publisher under the statute.¹⁴³ The plaintiffs also argued that Google did more than republish the content, but rather, at least in part, created and developed the content posted by ISIS to YouTube.¹⁴⁴ If Google was found to have created or developed the information, it would have lost immunity under Section 230.¹⁴⁵ To constitute creation or development, Google must have "materially contribut[ed] to its alleged unlawfulness."¹⁴⁶ The Ninth Circuit looked to *Kimzey v. Yelp! Inc.* for guidance, stating that merely proving a "neutral means by which third parties can post information of their own independent choosing online" is not enough to be considered to have created or developed content.¹⁴⁷ In *Kimzey*, the provider was deemed not to have created or developed the content when it did "nothing to enhance the defamatory sting of the message beyond the words offered by the [third-party] user."¹⁴⁸ Plaintiffs alleged Google made a material contribution when they created the "mosaic" that made a bigger picture when they paired the advertisements with videos because it enhanced engagement with the content.¹⁴⁹

The Ninth Circuit then highlighted Section 230 case law to demonstrate the concept of creation or development. No immunity was found in *Fair Housing Council v. Roommates.com, LLC*.¹⁵⁰ In that case, a website matched renters and tenants based on a questionnaire they filled out prior to signing up for the website.¹⁵¹ The questions were alleged to have violated federal and state anti-discrimination laws in housing.¹⁵² By asking questions about sex, sexual orientation, children, and preferred traits in a roommate, Roommates.com

¹⁴² *Id.*; 47 U.S.C. § 230(c)(1) ("No provider or user of an interactive computer service shall be treated as a publisher or speaker of any information provided by another information content provider.").

¹⁴³ *Gonzalez*, 2 F.4th at 891.

¹⁴⁴ *Id.* at 892.

¹⁴⁵ *Id.*

¹⁴⁶ *Id.* (quoting *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1167–68 (9th Cir. 2008)).

¹⁴⁷ *Id.* at 893 (quoting *Kimzey v. Yelp! Inc.*, 836 F.3d 1263, 1270 (9th Cir. 2016)).

¹⁴⁸ *Id.*

¹⁴⁹ *Id.*

¹⁵⁰ *Id.* at 894 (quoting *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1161 (9th Cir. 2008)).

¹⁵¹ *Id.* at 893–94 (quoting *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1161 (9th Cir. 2008)).

¹⁵² *Id.* at 894 (quoting *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1161 (9th Cir. 2008)).

encouraged users to post content to the site that violated fair housing laws.¹⁵³ The court ruled that the website was “the developer, at least in part, of that information” because it required subscribers to choose between a “limited set of pre-populated answers.”¹⁵⁴ The website lost its immunity due to its lack of neutral tools because it prompted discriminatory content.¹⁵⁵ However, the website in *Carafano v. Metroplash.com, Inc.* did supply neutral tools when it merely provided “a means for users to publish the profiles they created” for their online dating website.¹⁵⁶ Even though it had a matchmaking function that “allowed that content to be more effectively disseminated” it was still considered neutral because the website “did not encourage the posting of defamatory content by merely providing a means for users to publish the profiles they created.”¹⁵⁷ Finally, in *Dyroff v. Ultimate Software Group, Inc.*, a user bought heroin from another user on an online messaging board, and later died due to the drug being laced with fentanyl.¹⁵⁸ The website “used features and functions, including algorithms, to analyze user posts . . . and recommend other user groups.”¹⁵⁹ These were deemed as neutral tools because “[t]hese functions—recommendations and notifications—[were] tools meant to facilitate the communication and content of others,” and “not content in and of themselves.”¹⁶⁰

The downfall of the plaintiffs’ claim was the failure to allege that Google treated ISIS content differently than other types of content or encouraged videos that furthered the mission of ISIS.¹⁶¹ Therefore, the Ninth Circuit found that the platform was neutral and “did not specify or prompt the type of content to be submitted, nor determine particular types of content its algorithms would promote.”¹⁶² The court then went on to state that “machine-learning algorithms can never produce content within the meaning of Section 230. We only reiterate that a website’s use of content-neutral algorithms, without more, does not expose it to liability for content posted by a third-party.”¹⁶³

¹⁵³ *Id.*

¹⁵⁴ *Id.*

¹⁵⁵ *Id.*

¹⁵⁶ *Id.* at 893 (quoting *Fair Hous. Council of San Fernando Valley v. Roommates.Com, LLC*, 521 F.3d 1157, 1171 (9th Cir. 2008)).

¹⁵⁷ *Id.*

¹⁵⁸ *Id.* at 984 (quoting *Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093, 1094–96 (9th Cir. 2019)).

¹⁵⁹ *Id.* (quoting *Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093, 1098 (9th Cir. 2019)).

¹⁶⁰ *Id.*

¹⁶¹ *See id.* at 898.

¹⁶² *Id.* at 895 (quoting *Force v. Facebook, Inc.*, 924 F.3d 53, 70 (2d Cir. 2019) (holding that algorithms that “may have made content more visible or available . . . did not amount to developing the underlying information”).

¹⁶³ *Id.* at 896 (Gould, J., dissenting).

Although the Ninth Circuit dismissed most claims, the revenue-sharing claim survived because it did not rely on the content itself.¹⁶⁴ On YouTube, Google uses AdSense to sell advertising space to viewers.¹⁶⁵ If a user opts into the program, Google will share a “portion of the revenue generated by the advertisement on the user’s videos.”¹⁶⁶ Plaintiffs claimed that “because [Google] shared advertising revenue with ISIS, Google should be held directly liable for providing material support to ISIS”¹⁶⁷ They stated that because each “video must be reviewed and approved by Google before Google will permit advertisements to be placed with that video,” it had reviewed and approved ISIS videos and shared a “percentage of revenues generated from those advertisements with ISIS.”¹⁶⁸ This claim survived because Google did not act like a publisher or speaker of third-party content, but instead gave the third party money for posting the content, thus, not depending on the actual content of the video.¹⁶⁹ This is distinctly different from the other claims because it may be remedied without changing the content posted.¹⁷⁰ Additionally, they did not want accountability for the content itself, but for paying them for making the content.¹⁷¹ Due to the fact that this was unrelated to the content itself, the revenue claim was not immunized by Section 230.¹⁷² Although the Ninth Circuit’s holding has been vacated and remanded, it can still provide some strategic guidance on how to present a claim that is able to survive Section 230 as it becomes increasingly difficult to overcome.¹⁷³

C. *Commonwealth v. Carter*: The Case That Moved the Nation—and Hulu

Commonwealth v. Carter is a 2016 case out of the Supreme Court of Massachusetts that captured the nation and even prompted a Hulu docuseries.¹⁷⁴ Michelle Carter and Conrad Roy were dating at the time of Roy’s death but did not live in the same area, so most of their communication was over the phone, either by text or call.¹⁷⁵ Roy often expressed his troubles with mental health,

¹⁶⁴ *Id.* at 898.

¹⁶⁵ *Id.*

¹⁶⁶ *Id.*

¹⁶⁷ *Id.*

¹⁶⁸ *Id.*

¹⁶⁹ *Id.*

¹⁷⁰ *Id.*

¹⁷¹ *Id.* at 898–99.

¹⁷² *Id.*

¹⁷³ *Gonzalez v. Google LLC*, 143 S. Ct. 1191, 1192 (2023) (per curiam).

¹⁷⁴ *See e.g.*, *Commonwealth v. Carter*, 52 N.E.3d 1054 (Mass. 2016).

¹⁷⁵ *Id.* at 1057.

including his thoughts of suicide.¹⁷⁶ At first, Carter urged him to seek help but eventually got frustrated, chastised him, and then helped plan his suicide.¹⁷⁷ She knew of his previous suicide attempt, scolded him for putting it off, and even made him “promise” to go through with the plan.¹⁷⁸ In July 2014, Roy drove to a parking lot to commit suicide by carbon monoxide poisoning via a gasoline powered water pump in his truck while on the phone with Carter.¹⁷⁹ Roy got out of the truck to abandon his attempt, but Carter instructed him to get back in the truck.¹⁸⁰ Roy had made previous attempts at suicide, but he would normally abandon them, similar to how he got out of the truck on that day.¹⁸¹ After Roy got back in the truck, he died of carbon monoxide poisoning due to inhalation from the gasoline powered water pump.¹⁸² After Roy’s death, Carter texted a friend,

I failed [the victim] I wasn’t supposed to let that happen and now I’m realizing I failed him. [H]is death is my fault like honestly I could have stopped him I was on the phone with him and he got out of the car because it was working and he got scared and I fucking told him to get back in . . . because I knew he would do it all over again the next day and I couldn’t have him live the way he was living anymore I couldn’t do it I wouldn’t let him.¹⁸³

The court found that Carter was guilty of involuntary manslaughter under a theory of wanton or reckless conduct, which is “intentional conduct involv[ing] a high degree of likelihood that substantial harm will result to another.”¹⁸⁴ This is shown by:

specific knowledge or . . . what a reasonable person should have known in the circumstances . . . constitute wanton or reckless conduct . . . if an ordinary normal [person] under the same circumstances would have realized the gravity of the danger If based on the subjective measure, i.e., the defendant’s own knowledge, grave danger to others must have been apparent and the defendant must have chosen to run the risk rather than alter [his or her] conduct so as to avoid the act or omission which caused the harm.¹⁸⁵

Carter argued that, because she was not physically present and did not provide the means by which Roy killed himself, her verbal encouragement to continue

¹⁷⁶ *Id.* at 1057.

¹⁷⁷ *Id.* at 1057, 1064.

¹⁷⁸ *Id.* at 1058 n.6.

¹⁷⁹ *Id.* at 1058 n.5, 1059 n.6.

¹⁸⁰ *Id.* at 1059.

¹⁸¹ *Id.* at 1058 n.5, 1059.

¹⁸² *Id.* at 1056.

¹⁸³ *Id.* at 1059 n.8.

¹⁸⁴ *Id.* at 1060 (quoting *Commonwealth v. Pugh*, 969 N.E.2d 672, 685 (Mass. 2012)).

¹⁸⁵ *Id.*

the attempt did not cause his death and that her actions would therefore not constitute wanton or reckless conduct that overcame Roy's willpower to live.¹⁸⁶ The court was not persuaded and affirmed the juvenile court's grand jury indictment of Carter.¹⁸⁷ Although the court had never considered a manslaughter indictment based on words, it had seen cases of "involuntary manslaughter against a defendant where the death of the victim [was] self-inflicted."¹⁸⁸ In *Commonwealth v. Atencio*, the court stated that it had an "interest that the deceased should not be killed by the wanton or reckless conduct of himself and others . . . found in the concerted action and cooperation of the defendant in helping to bring about the deceased's foolish act."¹⁸⁹ Additionally, in *Persampieri v. Commonwealth*, the court ruled that, when a wife threatened to commit suicide, the husband, knowing of her previous attempts, was wanton and reckless by "instead of trying to bring [the victim] to her senses, taunted her, told her where the gun was, loaded it for her, saw that the safety was off, and told her the means by which she could pull the trigger."¹⁹⁰

Carter's words were overly coercive and had "more weight than mere words overcoming any independent will to live" because of her virtual presence, constant pressure on the victim, and the victim's already fragile mental state.¹⁹¹ Others in that situation would have realized the danger of their words towards a vulnerable individual who had previously attempted suicide, especially, as in the case here, in the position as the victim's girlfriend and because of her constant communication with the victim.¹⁹² The court found that Carter "knew that she had some control over" the victim, and her words "overbore the victim's willpower" to abandon his attempt.¹⁹³

In *Carter II*, the sufficiency of the evidence was considered, and the court revisited the argument that Carter was protected by the First Amendment under the principles of free speech.¹⁹⁴ They affirmed that Carter could not "escape liability just because she happened to use words to carry out [her] illegal [act]" because many crimes can be verbally committed and "they are intuitively and correctly understood not to raise First Amendment concerns."¹⁹⁵ The Supreme Court has affirmed that "speech or writing used as an integral part of conduct in

¹⁸⁶ *Id.* at 1061.

¹⁸⁷ *Id.*

¹⁸⁸ *Id.* at 1061–62.

¹⁸⁹ *Commonwealth v. Atencio*, 189 N.E.2d 223, 224–25 (Mass. 1963).

¹⁹⁰ *Persampieri v. Commonwealth*, 175 N.E.2d 387, 390 (Mass. 1961).

¹⁹¹ *Carter*, 52 N.E.3d at 1063.

¹⁹² *Id.*

¹⁹³ *Id.*

¹⁹⁴ *Commonwealth v. Carter*, 115 N.E.3d 559, 562 (Mass. 2019).

¹⁹⁵ *Id.* at 570 (quoting *United States v. Barnett*, 667 F.2d 835, 842 (9th Cir. 1982)).

violation of a valid criminal statute” is not a category of speech protected by the First Amendment.¹⁹⁶ The Massachusetts Supreme Court further stated that Carter’s words fell into the category of criminal conduct because they were “not punishing words alone, as the defendant claims, but reckless or wanton words causing death. The speech at issue is thus integral to a course of criminal conduct and thus does not raise any constitutional problem.”¹⁹⁷

Carter demonstrates that exposure to messages encouraging suicide may lead to a tragic result for those who already struggle with mental health. Although there has been no formal action against TikTok, due to Section 230, TikTok should take this as a cautionary tale because of the strong parallels its algorithm draws to Carter’s actions. Specifically, the reckless action of knowingly showing users who struggle with mental health videos that either encourage self-harm or harm to others, while providing no self-help mechanisms, may lead to the peril of many users.

With the rise of Generation Alpha on social media, we have started to see litigation involving the mental health crisis allegedly caused using such websites and applications. Recently, a fourteen-year-old girl in England who committed suicide has captured the attention of the nation.¹⁹⁸ Molly Russel used Instagram and Pinterest to view posts about mental health, such as self-harm, depression, and suicide, until her death in November 2017.¹⁹⁹ She used “her Instagram account up to 120 times a day, liking more than 11,000 pieces of content . . . [and] Pinterest more than 15,000 times” in the six months prior to her death, with the coroner reviewing her case stating that the content was “‘pretty dreadful’ and difficult even for adults to look at for extended periods of time.”²⁰⁰

A whistleblower at Meta, the company that owns Instagram, released company data that acknowledged Meta had knowledge that their platform was extremely harmful for users.²⁰¹ The head of community operations at Pinterest, which is not owned by Meta, conceded that Pinterest was “not safe” and that he “deeply regrets” the content that Russel viewed.²⁰²

On October 13, 2022, the coroner issued a *Regulation Report to Prevent*

¹⁹⁶ *Id.* at 571 (quoting *Giboney v. Empire Storage & Ice Co.*, 336 U.S. 490, 498 (1949)).

¹⁹⁷ *Id.* at 572.

¹⁹⁸ See Tama Leaver, *Coroner Finds Social Media Contributed to 14-Year-Old Molly Russell’s Death. How Should Parents and Platforms React?*, THE CONVERSATION (Oct. 3, 2022), <https://theconversation.com/coroner-finds-social-media-contributed-to-14-year-old-molly-russells-death-how-should-parents-and-platforms-react-191757>.

¹⁹⁹ *Id.*

²⁰⁰ Angus Crawford, *Molly Russell’s Inquest to Put Focus on Big Tech*, BBC NEWS (Sept. 20, 2022), <https://www.bbc.com/news/uk-62958851>.

²⁰¹ *Id.*

²⁰² Angus Crawford & Bethan Bell, *Molly Russell Inquest: Father Makes Social Media Plea*, BBC NEWS (Sept. 30, 2022), <https://www.bbc.com/news/uk-england-london-63073489>.

Future Deaths, concluding that Pinterest “operated in such a way using algorithms as to result, in some circumstances, of binge periods of images, video clips and text some of which were selected and provided without Molly requesting them.”²⁰³ The content romanticized self-harm and discouraged discussion with those who can provide help, even going as far as to “portray self-harm and suicide as an inevitable consequence of a condition that could not be recovered from.”²⁰⁴ The normalization of her depressive illness combined with a “limited irrational view without any counterbalance of normality . . . affected her mental health in a negative way and contributed to her death in a more than minimal way.”²⁰⁵ The coroner stated that the social media platforms should take action to prevent future deaths and that they have the power to take such action, resulting in a duty to respond with “details of action taken or proposed to be taken, setting out the timetable for action.”²⁰⁶ The report provided many recommendations, such as independent regulatory bodies to monitor online platforms and a review of internet provisions relating to children; the report stated that although regulations are a governmental role, there is “no reason why the platforms themselves would not wish to give considerations to self-regulation taking into account the matters raised above.”²⁰⁷

A social media platform being deemed as part of the cause of death in a suicide is a monumental step toward effective regulation of social media considering how the platforms have been shielded from liability for so many years.²⁰⁸ The United Kingdom is ahead of the United States in online regulation, specifically in protecting children from harmful content and criminalizing social media actions, even presenting the possibility of prosecution or jail time for social media executives rather than implementing a statutory equivalent to Section 230.²⁰⁹ There have been social movements that promote increased liability for social media platforms, but they have seen little success against the existing incentives to promote free and full discretion of users’ engagement with the platform.²¹⁰ However, this is not necessarily a barrier to achieving liability

²⁰³ ANDREW WALKER, NORTH LONDON CORONER’S SERV., REGULATION 28 REPORT TO PREVENT FUTURE DEATHS (2022) (U.K.).

²⁰⁴ *Id.*

²⁰⁵ *Id.*

²⁰⁶ *Id.*

²⁰⁷ *Id.*

²⁰⁸ Crawford & Bell, *supra* note 202.

²⁰⁹ See *World-First Online Safety Laws Introduced in Parliament*, GOV.UK (Mar. 17, 2022), <https://www.gov.uk/government/news/world-first-online-safety-laws-introduced-in-parliament>.

²¹⁰ See David McCabe, *Supreme Court Poised to Reconsider Key Tenets of Online Speech*, N.Y. TIMES (Jan. 19, 2023), <https://www.nytimes.com/2023/01/19/technology/supreme-court-online-free-speech-social-media.html>; see generally #BeBraveZuck, BRAVE

for social media platforms; we have previously seen examples where private companies that cause social harm have been held responsible.²¹¹ This may be the next step for social media.

D. Cigarette and Firearms Facing Judicial Action

Although Section 230 shields TikTok from liability, private actors have been held liable for social harms despite having similar statutes protecting them from liability.²¹² Firearm and cigarette manufactures have started to face consequences for their actions through suits based on claims of misleading and predatory advertisements.²¹³ These actions will be compared to the predatory practices of the TikTok algorithm and the company's misleading claims about a user's ability to control the content they consume on the app.²¹⁴

In response to the Sandy Hook shooting, parents filed a suit against a gun manufacturer for violating consumer protection laws by targeting troubled teenagers in their advertisements.²¹⁵ Firearms companies have broad immunity from civil litigation from the criminal or otherwise unlawful use of their products under the Protection of Lawful Commerce in Arms Act (PLCAA).²¹⁶ The PLCAA has been criticized for encouraging “‘head in the sand’ behavior that shirks any role for those manufacturing, selling, marketing, or importing firearms and related products in identifying dangerous patterns, proactively change behaviors, and lead toward lifesaving innovations and solutions,” and for leaving victims without compensation or justice while providing business advantages to irresponsible manufacturers over other manufacturers that take precautions.²¹⁷ However, the PLCAA has an exception for when the manufacturer knowingly violated a statute relating to the sale or marketing of guns, if that is the proximate cause of the alleged harm.²¹⁸ This exception is known as the “predicate exception” and has been the main source of dispute due

MOVEMENT, <https://web.archive.org/web/20230326192110/https://www.bravemovement.org/campaign/bebravezuck> (last visited Dec. 5, 2023); *see generally* *Internet Child Safety Foundation*, HALLEY MOVEMENT, <https://halleymovement.org/our-programs/internet-safety/> (last visited Dec. 5, 2023); *see generally* PACT, <https://www.ecpatusa.org/mission> (last visited Dec. 5, 2023).

²¹¹ *See Gun Industry Immunity*, *supra* note 18; *see* Goguen, *supra* note 18.

²¹² *See Gun Industry Immunity*, *supra* note 18.

²¹³ *See id.*

²¹⁴ *See id.*

²¹⁵ *Soto v. Bushmaster Firearms Int'l, LLC*, 202 A.3d 262, 275 (Conn. 2019).

²¹⁶ Protection of Lawful Commerce in Arms Act, Pub. L. No. 109-92, 119 Stat. 2095 (2005).

²¹⁷ *Gun Industry Immunity*, *supra* note 18.

²¹⁸ 15 U.S.C. § 7903(5)(A)(iii); *see* Karena Phan, *Despite Biden's Claim, Gun Makers Can Indeed Be Sued*, AP NEWS (Feb. 9, 2022, 8:37 PM), <https://apnews.com/article/fact-checking-590518743186>.

to the determination of what a proper predicate statute is under the PLCAA.²¹⁹

In *Soto v. Bushmaster Firearms International*, plaintiffs claimed that Bushmaster targeted their firearm advertisements toward at-risk teens by placing their advertisements in violent video games and using phrases such as “Consider Your Man Card Reissued” in an attempt to appeal to that demographic.²²⁰ A state statute prohibited “advertisements that promote or encourage violent, criminal behavior.”²²¹ The court ruled that “Congress has not clearly manifested an intent to extinguish the traditional authority of our legislature and our courts to protect the people of Connecticut from the pernicious practices alleged in the present case” and there was “simply no reason to think that the present action represents the sort of novel civil action Congress sought to bar” by passing statutes that shielded gun manufactures from other types of liability.²²² Overall, the court’s interpretation of the PLCAA reflected the idea that targeting troubled teens in advertisements was predatory and deserved to be exempted from immunity protections for firearms manufactures.

Since this case, four states—New York, Delaware, New Jersey, and California—have enacted legislation reducing the liability shield and expanding victims’ ability to bring civil suits for illegal conduct.²²³ These states have required “firearm industry members to implement reasonable controls, such as standard procedures, safeguards, screening or security standards, and safe business practices, in order to prevent specified harms.”²²⁴ California specifically intensified their regulations by implementing requirements such as taking “reasonable precautions to ensure they do not sell, distribute, or provide a firearm-related product . . . [to those] who fail to comply with the firearm industry standard of conduct” and prohibiting the manufacturing, sale, or importation of products that are abnormally dangerous and likely to cause harm to the public health and safety.²²⁵ Meanwhile, New York, Delaware, and California have also “create[d] a new cause of action authorizing any person harmed by a firearm industry member’s violation of the law to file suit and seek compensation or other fair remedies in court.”²²⁶ Although the PLCAA has prevented much of the liability from reaching manufacturers, the courts have

²¹⁹ *Gun Industry Immunity*, *supra* note 18.

²²⁰ Dave Collins, *Sandy Hook Families Settle for \$73M With Gun Maker Remington*, AP NEWS (Feb. 15, 2022), <https://apnews.com/article/sandy-hook-school-shooting-remington-settlement-e53b95d398ee9b838afc06275a4df403>.

²²¹ *Soto v. Bushmaster Firearms Int’l, LLC*, 202 A.3d 262, 272 (Conn. 2019).

²²² *Id.* at 272, 310.

²²³ *Gun Industry Immunity*, *supra* note 18.

²²⁴ *Id.*

²²⁵ *Gun Industry Immunity*, *supra* note 18; *see* CAL. CIV. CODE § 3273.51(b)(2) (2023); *see* CAL. CIV. CODE § 3273.51(c) (2023).

²²⁶ *Gun Industry Immunity*, *supra* note 18.

found a way to circumvent this, with state legislatures following and creating their causes of action for gun manufacturers under the predicate exception.

Cigarette manufactures faced a similar course when they encountered issues of liability surrounding their advertising methods in the mid-twentieth century. Much of the accountability for their social wrongdoing is based on a theory of personal injury from a proven addictive product.²²⁷ In the 1950s, the first wave of lawsuits against cigarette manufacturers emerged, and manufacturers continually prevailed in cases alleging negligent manufacturing, products liability, negligent advertising, fraud, and violation of consumer protection statutes.²²⁸ However, in the second wave of litigation in the 1980s, courts started to shift their rulings.²²⁹ In *Cipollone v. Liggett Group*, the Supreme Court ruled that a suit could not be brought against cigarette producers for failure to inform users of harmful effects of cigarettes, but users could bring suit arising from inaccurate statements in advertisements or conspiring with others to mislead consumers about the dangers of smoking.²³⁰ This result was likely influenced by the Federal Cigarette Labeling and Advertising Act of 1965, which preempted many claims against manufacturers based on liability for their conduct.²³¹ The holding established that even when someone knowingly uses an addictive product, companies can still be liable for harm if the company conceals the danger. Due to the increasingly demonstrated knowledge of many cigarette manufacturers, courts started to rule for the user, ultimately leading to the 1998 Master Settlement Agreement, which held large tobacco manufactures liable to forty-six states for \$206 billion in compensation for public costs related to smoking.²³²

In 2008, a third wave of cigarette manufacturing litigation emerged when the Supreme Court denied preemption arguments for “light” cigarettes, stating that “federal legislation does not preempt plaintiffs from suing under certain state unfair business practice laws,” and thus open[ing] the door for similar lawsuits against tobacco companies.”²³³ More recently, we have seen the continuation of this trend with Juul, a popular e-cigarette manufacturer, settling suits alleging it “violated the Consumer Protection Act tens of thousands of times by creating and marketing products catered toward underage customers . . . [and] deceiv[ed]

²²⁷ *See id.*

²²⁸ Goguen, *supra* note 18.

²²⁹ *Id.*

²³⁰ *Cipollone v. Liggett Grp.*, 505 U.S. 504, 530–31 (1992).

²³¹ *Tobacco Lawsuits Alleging Product Defects*, JUSTIA, <https://www.justia.com/products-liability/types-of-defective-product-cases/tobacco/> (Oct. 2023).

²³² *The Master Settlement Agreement*, NAT’L ASS’N. OF ATT’YS GEN., <https://www.naag.org/our-work/naag-center-for-tobacco-and-public-health/the-master-settlement-agreement/> (last visited Dec. 5, 2022).

²³³ Goguen, *supra* note 18; *see generally* *Altria Grp., Inc. v. Good*, 555 U.S. 70 (2008).

customers about how addictive its products are.”²³⁴ Juul was also ordered to remove certain products from the market because its marketing targeted young people, and many speculate that these types of suits will be the next wave of cigarette litigation.²³⁵

The litigation history surrounding gun and cigarette manufacturers displays that, although there may be statutes that prevent full liability for products that cause social harms, courts may still be able to hold egregious wrongdoers liable.²³⁶ This should be the case with social media platforms. Guns, cigarettes, and social media are all products that society tends to blame the user for harm suffered due to use of the product.²³⁷ Over time, we have seen a narrative switch about how the company attributes the harm of the product. Instead of focusing on why an individual was smoking or chose to kill with a gun, courts shifted toward analyzing how the companies targeted their products at sensitive groups, despite knowing that advertisements could cause guns to fall into the wrong hands or could misinform users about the dangers of smoking.²³⁸

Despite Section 230 deeming TikTok not liable for their wrongs, there are parallels to *Soto* and *Cipollone*, such as the targeting of troubled teens with content that specially appeals to their fragile state and failure to inform users of how the algorithm produces the results on a “For You Page” or to warn that blocked topics may still appear.²³⁹ While social media regulation has been a source of contention for years, the best approach to regulating social media is to follow the precedents set by the judicial regulation of other harmful products, such as firearms and cigarettes, which ultimately held major actors liable for wrongdoing. Additionally, the legislature should follow judicial action by either actively amending Section 230 or implicitly adopting the court’s ruling to strike Section 230’s liability provisions as applied to social media.

Although the courts have historically sided with the social media platforms, we are starting to see a new wave of social media litigation. *Gonzalez v. Google* is an important case for the state of Section 230, despite its recent remand, because it shows that it is possible for a claim to overcome the Section 230 liability shield when paired with causes of actions that arise out of other statutory

²³⁴ Mychael Schnell, *Juul Settles Washington State Deceptive Advertising Suit for \$22.5M*, THE HILL (Apr. 13, 2022, 3:51 PM), <https://thehill.com/policy/healthcare/3266658-juul-settles-washington-state-deceptive-advertising-suit-for-22-5m/>; see Lateshia Beachum & Laurie McGinley, *Juul to Pay \$439 Million in Settlement Over Marketing to Teens*, WASH. POST (Sept. 6, 2022), <https://www.washingtonpost.com/nation/2022/09/06/juul-settlement-vaping-advertising-teens/>.

²³⁵ Goguen, *supra* note 18.

²³⁶ *Gun Industry Immunity*, *supra* note 18; see Goguen, *supra* note 18.

²³⁷ *Gun Industry Immunity*, *supra* note 18; Goguen, *supra* note 18.

²³⁸ See *Gun Industry Immunity*, *supra* note 18; see Goguen, *supra* note 18.

²³⁹ See Jargon, *supra* note 3.

claims.²⁴⁰ Due to the shift in how the courts are interpreting Section 230, it is important to move forward and support further judicial action since the legislature has been slow to address the issue at hand. The legislature may then either move forward with amending Section 230 or adopt the interpretation of the courts. Either way will lead to a successful increase in user protection on social media platforms. Many other harmful products have followed this pattern, which has led to increased protection for consumers.²⁴¹

II. COMPARING *CARTER* TO CODERS

TikTok causes harm to the mental wellbeing of its users, showing an unprecedented level of destruction that towers over the typical social media addiction claim. Mental health in the United States is rapidly declining because the algorithms of social media platforms perpetuate harmful ideas and reinforce them to users.²⁴² This wanton and reckless conduct is comparable to that of Michelle Carter, and social media platforms should be held responsible in a way that is comparable to how the Ninth Circuit held Google responsible for paying revenue to ISIS.²⁴³ TikTok's business model is similar to these situations because they aim harmful videos toward users who may be struggling and simultaneously encourage harmful posts by incentivizing rage-baiting and other types of damaging content through its Creator Fund.²⁴⁴

TikTok has an immense amount of knowledge about each individual user and gives them either what they are interested in or something directly opposed to what they are interested in that gets them to continue watching, which is extremely harmful for those going through a mental health crisis.²⁴⁵ Arguably, TikTok knows more than Carter did in respect to Roy because Carter only knew what Roy was telling her about his struggles with mental health.²⁴⁶ TikTok's mission to increase watch time combined with its lack of content control amounts to the same level of wanton and reckless conduct as in *Carter* due to its high level of knowledge about every user and the likelihood that harm will result if it continues to push harmful content to users who are likely in a mental health

²⁴⁰ See generally *Gonzalez v. Google LLC*, 143 S. Ct. 1191, 1192 (2023) (“declin[ing] to address the application of § 230 to a complaint that appears to state little, if any, plausible claim for relief”).

²⁴¹ See *Gun Industry Immunity*, *supra* note 18; see *Goguen*, *supra* note 18.

²⁴² See Dylan Walsh, *Study: Social Media Use Linked to Decline in Mental Health*, MASS. INST. TECH. SLOAN SCH. OF MGMT. (Sept. 14, 2022), <https://mitsloan.mit.edu/ideas-made-to-matter/study-social-media-use-linked-to-decline-mental-health>.

²⁴³ See generally *Commonwealth v. Carter*, 52 N.E.3d 1054 (Mass. 2016); see generally *Gonzalez v. Google LLC*, 2 F.4th 871, 880 (9th Cir. 2021).

²⁴⁴ See generally *Barnhart*, *supra* note 72.

²⁴⁵ See *Walsh*, *supra* note 242.

²⁴⁶ See generally *Carter*, 52 N.E.3d 1054.

crisis.²⁴⁷

A reasonable person in the shoes of TikTok algorithm coders would realize the gravity of their actions and increase internal restrictions on harmful content or pare back on the level of specialized content to prevent users from falling deep into subject matter rabbit holes.²⁴⁸ Michelle Carter only knew what the victim told her about how he was feeling.²⁴⁹ He may have had many more struggles than he was willing to tell her, thus limiting the knowledge that was available to her.²⁵⁰ However, TikTok is able to tap into a user's thoughts more accurately than friends or family would be able to. This removes many of the limitations of human relationships because, although users may not know it, when they use TikTok, they are expressing their feelings in a vulnerable way that goes far beyond how they talk to their friends about their troubles.²⁵¹ Individuals are much less likely to sugarcoat negative thoughts and emotions or to hide the extent to which they are struggling when interacting with content on the app.²⁵² The extent of TikTok's knowledge far surpasses anything that Carter may have known about Roy.²⁵³ When cumulated, the harm spread across TikTok's one billion users is much more detrimental to the public than the harm perpetuated by Carter; this creates a public health issue comparable to that of cigarettes or guns in the way that it harms not only individuals, but also the greater public as a whole.²⁵⁴

The fact that the harm is caused by intentionally swiping on the app is not a barrier to holding TikTok accountable; TikTok itself does not need to perpetuate a physical act toward a user to be liable for harm caused.²⁵⁵ States possess an interest in people not being killed by the reckless conduct of themselves and others, such as bringing about a victim's act of self-harm.²⁵⁶ *Carter* cited

²⁴⁷ See *id.* at 1060 (quoting *Commonwealth v. Pugh*, 969 N.E.2d 672, 672 (Mass. 2012)) (stating "specific knowledge or on what a reasonable person should have known in the circumstances . . . constitute wanton or reckless conduct . . . if an ordinary normal [person] under the same circumstances would have realized the gravity of the danger . . . [i]f based on the subjective measure, i.e., the defendant's own knowledge, grave danger to others must have been apparent and the defendant must have chosen to run the risk rather than alter [his or her] conduct so as to avoid the act or omission which caused the harm") (citation omitted).

²⁴⁸ See generally *Johnson*, *supra* note 19.

²⁴⁹ See *Carter*, 52 N.E.3d at 1057–58.

²⁵⁰ See *id.* at 1057.

²⁵¹ See *id.* at 1057, 1059.

²⁵² See *id.* at 1057 n.5.

²⁵³ See *id.* at 1063–64.

²⁵⁴ Rohit Shewale, *46 TikTok Statistics 2023 (Users, Creators & Revenue)*, DEMAND SAGE (Nov. 22, 2023), <https://www.demandpage.com/tiktok-user-statistics/>.

²⁵⁵ See *Carter*, 52 N.E.3d at 1061, 1063.

²⁵⁶ *Id.* at 1062.

examples in which this was true, showing that the reckless disregard of another's safety and the consequences of one's own actions are enough to satisfy this standard.²⁵⁷ Although Carter's direct words to the victim were the way in which she overpowered his will to live, TikTok showing users harmful content is sufficient to satisfy the fact-specific standard because of its coercive quality of showing the user whatever makes them spend the most time possible on the app, even though it might be harmful to them.²⁵⁸

The *Carter* court reasoned that, but for the communication between the two teens, Roy would have saved himself.²⁵⁹ This reasoning is also applicable to social media platforms because, but for the pushing of harmful ideas at users and the difficulty of escaping them, many users may have avoided diminishing mental health.²⁶⁰ Many users will likely engage in self-help mechanisms, such as blocking specific hashtags or users or utilizing the "Not Interested" function, but this will not prevent the content from showing up again.²⁶¹ We have seen users employ loopholes to get around the ban on eating disorder hashtags by simply misspelling common phrases to circumvent the ban and land their content on the "For You Page" of many who attempted to stop similar content from continuing to be recommended.²⁶² Additionally, in the slim chance that users have connected that watch time and hovering over videos is the key to what is suggested, this is an extremely difficult way to engage in self-help.²⁶³ Many times, users will have to watch a few seconds of a video to comprehend the content and process that it is content they do not wish to see. This may decrease the volume of videos they see, but there is no guarantee that they will stop for good because the video is still getting watch time.

A user's only options are to either watch the video or scroll away, with few other ways to communicate to the algorithm that they do or do not want the content on their page.²⁶⁴ Scrolling past this addictive content is a huge struggle for someone already in a bad mental health state.²⁶⁵ This is sufficient to show that, but for the actions of TikTok, some of the content may never have reached the user and if it did, they would have been able to engage in effective self-help to prevent themselves from seeing harmful ideations. As seen in *Carter*, the

²⁵⁷ *Id.* at 1062; *see generally* *Persampieri v. Commonwealth*, 175 N.E.2d 387 (Mass. 1961).

²⁵⁸ *See generally* *Carter*, 52 N.E.3d 1054; *see Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

²⁵⁹ *See Carter*, 52 N.E.3d at 1063–64.

²⁶⁰ *See generally* *Jargon*, *supra* note 3.

²⁶¹ *See generally id.*

²⁶² *See generally id.*

²⁶³ *See generally id.*

²⁶⁴ *See Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

²⁶⁵ *See Carter*, 52 N.E.3d at 1063.

criminality of this conduct arises from the “manipulative” wanton and reckless speech that overpowers the will to live.²⁶⁶ So, although TikTok is not responsible for what people say vis-à-vis Section 230, they could still be responsible for knowingly promoting this conduct to vulnerable users and providing insufficient self-help mechanisms to escape. This is relevant to showing that users are unable to escape harmful content despite the protective mechanism put in place, and an individual’s best efforts to avoid content are not enough; like in *Carter*, although a user may engage in self-help, this is insufficient to overcome the larger harm occurring from an outside source.²⁶⁷

When distinguished from current Section 230 case law, it is apparent that TikTok “materially contribut[ed] to [the] alleged unlawfulness” of the creation or development of the harmful ideas perpetuated by the algorithm.²⁶⁸ TikTok, in effect, “prompt[s] the type of content to be submitted” because creators know that rage-baiting goes viral and causes creators to gain a following faster; by compensating creators to make videos that go viral, TikTok encourages the posting of these types of videos.²⁶⁹ This serves as a basis to revoke TikTok’s Section 230 immunity because the algorithm is assisting in the development of harmful content via compensation.²⁷⁰ In *Gonzalez*, the claim failed because the plaintiffs did not adequately show that the YouTube algorithm was prompting the creation of content,²⁷¹ so TikTok’s retention based algorithm is the key to defeating Section 230 immunity.

TikTok is unlike the “neutral means by which third parties can post information of their own independent choosing online.”²⁷² TikTok is enhancing the “‘sting of the message’ beyond the words offered by the [third-party] user” by encouraging video creation and furthering their message by continuously showing them to users with no way to prevent them from coming up.²⁷³ Unlike *Carafano*, TikTok does more than use a matching function to match users with content they may like, because it encourages the content to be posted.²⁷⁴

²⁶⁶ *See id.*

²⁶⁷ *See id.*; *see Investigation: How TikTok’s Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

²⁶⁸ *Gonzalez v. Google LLC*, 2 F.4th 871, 892 (9th Cir. 2021) (emphasis omitted) (quoting *Fair Hous. Council of San Fernando Valley v. Roommates.Com, LLC*, 521 F. 3d 1157, 1167–68 (9th Cir. 2008)).

²⁶⁹ *Id.* at 895.

²⁷⁰ *See id.* at 893–94.

²⁷¹ *See id.* at 894–95.

²⁷² *Klayman v. Zuckerberg*, 753 F. 3d 1354, 1358 (D.C. Cir. 2014).

²⁷³ *Gonzalez*, 2 F. 4th at 893 (quoting *Kimzey v. Yelp!*, 836 F. 3d 1263, 1270 (9th Cir. 2016) (quoting *Fair Hous. Council of San Fernando Valley v. Roommates.Com, L.L.C.*, 521 F. 3d 1157, 1172 (9th Cir. 2008))).

²⁷⁴ *See id.*; *see Carafano v. Metrosplash.com, Inc.*, 339 F.3d 1119, 1124–25 (9th Cir. 2003).

TikTok’s algorithm is more like *Roommates*, because users are directed through their view time, likes, follows, etc., to show their thoughts and opinions about any topic possible.²⁷⁵ Because view time is the main factor in populating content, this is comparable to the “limited set of prepopulated answers” that users must choose between, thus making TikTok the developer of the information.²⁷⁶ TikTok does not ask for users’ responses to what they would like to see, but instead automates the process and curates its own results; users’ only options are to stay and watch or swipe away as quickly as possible, both of which contribute to increased engagement time with the video, which then increases its frequency on the “For You Page.” As seen above, the self-help mechanisms are not effective to rid a user’s “For You Page” of content they do not wish to see. By not allowing users an opportunity to affirmatively show the algorithm what they do not want to see, the algorithm prompts users to rage-watch harmful content, without providing recourse to correct it. Although *Dyroff* stated that the functions of that algorithm were “‘tools meant to facilitate the communications and content of others,’ and ‘not content in and of themselves,’” the precedent does not bar liability for TikTok algorithm, because it did not hold that algorithms can never be held responsible.²⁷⁷

III. TICK TOCK, TIKTOK ... YOUR LIABILITY SHIELD MAY HAVE LIMITED TIME LEFT

In the past decade, litigation against social media platforms on the basis of addiction, product liability, and free speech has increased.²⁷⁸ However, as discussed, very few claims have been successful because Section 230 protects platforms from liability for content posted to their sites.²⁷⁹ Users have started to demand more accountability from social media platforms, but there has been little movement towards this goal.²⁸⁰ The social harms flowing from social media platforms are similar to the social harms caused by the manufacturing and marketing of guns, nicotine, and other potentially dangerous products.²⁸¹ These have all required judicial solutions to pierce liability shields, which is the next step for the internet.²⁸² Although Section 230 currently protects TikTok from liability, we should impose liability on social media platforms through the

²⁷⁵ See *Roommates.Com, LLC*, 521 F.3d at 1161.

²⁷⁶ See generally *id.*

²⁷⁷ *Id.* at 894; *Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093, 1098 (9th Cir. 2019).

²⁷⁸ See generally McCabe, *supra* note 210.

²⁷⁹ *Id.*

²⁸⁰ *Id.*

²⁸¹ See e.g., *Gun Industry Immunity*, *supra* note 18; see also Goguen, *supra* note 18.

²⁸² See *Gun Industry Immunity*, *supra* note 18; see Goguen, *supra* note 18.

judicial process, similar to how gun and cigarette manufacturers were initially immune, but nevertheless eventually held liable.

Gun manufacturers have been a popular target due to the social harms caused by their products.²⁸³ Similar to *Soto*, TikTok targets youth by marketing itself as a fun app where you can dance with your friends, but it is not transparent about how its algorithm uses unconscious thoughts to place content on a user's "For You Page."²⁸⁴ This tactic is similar to the way that Bushmaster placed advertisements in violent video games, because they knew that those who played those types of games would be more susceptible to their ads.²⁸⁵ By showing users harmful ideas, both TikTok and Bushmaster knowingly disregarded the safety of the public.²⁸⁶ This predatory tactic produces large social harms by reinforcing harmful messages to troubled teens.

Similar to *Cipollone*, social media platforms have misled users about the harmful effects of their services.²⁸⁷ Society tends to blame cigarette and nicotine users for their addiction to the product, but this undercuts the tactics manufacturers use to appeal to potential buyers.²⁸⁸ By depicting smoking cigarettes as a fun, occasional activity to do in a social or party setting, tobacco manufacturers fail to warn that it has highly addictive qualities that may hook users within a few uses.²⁸⁹ For years, the harmful effects of social media have been a topic of discussion, and while some platforms are starting to change their community guidelines,²⁹⁰ this is far from enough.²⁹¹ The current narrative focuses on what users can do to avoid this content, not on what the platform can do.²⁹² Yes, users can block specific users and hashtags, or mark content as "not interested," but there are loopholes that content can jump through to still present itself on "For You Pages."²⁹³ Social media platforms are not doing enough to inform users of the harmful effects of using their products or to give them ways to mitigate the harmful content they see, similar to how cigarette manufacturers showed the fun and exciting side of tobacco and nicotine, but not the harmful

²⁸³ See *Gun Industry Immunity*, *supra* note 18; see Goguen, *supra* note 18.

²⁸⁴ See generally *Soto v. Bushmaster Firearms Int'l, LLC*, 202 A.3d 262 (Conn. 2019).

²⁸⁵ See generally *id.*

²⁸⁶ See generally *id.*

²⁸⁷ See generally *Cipollone v. Liggett Grp.*, 505 U.S. 504 (1992).

²⁸⁸ Goguen, *supra* note 18.

²⁸⁹ *Id.*

²⁹⁰ See Jon Bateman et al., *How Social Media Platforms' Community Standards Address Influence Operations*, CARNEGIE ENDOWMENT FOR INT'L PEACE (Apr. 1, 2021), <https://carnegieendowment.org/2021/04/01/how-social-media-platforms-community-standards-address-influence-operations-pub-84201>.

²⁹¹ See *id.*

²⁹² See *id.*

²⁹³ Jargon, *supra* note 3.

side effects of use.²⁹⁴ TikTok is engaging in deceitful conduct by not telling users that their self-help mechanisms are not an effective way to prevent the content from showing up, leaving open the opportunity for users to get sucked into the very harmful content rabbit holes they were trying to avoid.²⁹⁵

By comparing social media to cigarette and gun manufacturers, it is clear that they have very similar social consequences. Both guns and cigarettes have statutory liability shields but have nevertheless been held liable for wrongdoing.²⁹⁶ Social media platforms should be treated the same way. Although Section 230 protects TikTok from claims such as creating addictive practices or hosting harmful content, there should be liability imposed for targeting groups that may be struggling with their mental health while also misleading them to believe they have control over their content.

Congress has started to show an intent to increase the liability of social media platforms.²⁹⁷ *Gonzalez* has shown that as the internet continues to change, there are issues that Section 230 did not comprehend,²⁹⁸ fueling a movement to amend Section 230.²⁹⁹ The legislature should incorporate the reasonable person standard into Section 230, as a starting point, because of its traditional role in businesses and clearly defined case law already showing a path for platforms to take.³⁰⁰ Some platforms have taken it upon themselves to satisfy this standard by implementing user controls, but an official amendment to the standard would force platforms to reflect on their practices due to a new self-interest in preventing litigation.³⁰¹ Although the reasonable person standard is far from a solution to the issue, this is the most rational first move, and will not hurt the current benefits of social media.³⁰² As social media progresses, the reasonable person standard will mold into a well-defined standard tailored to social media platforms that sets out what it means for platforms to act in a socially responsible way.

CONCLUSION

In sum, TikTok should be held liable for showing users videos with harmful

²⁹⁴ See *Gun Industry Immunity*, *supra* note 18; see Goguen, *supra* note 18.

²⁹⁵ See Jargon, *supra* note 3 (stating a user can block “#anorexia” but other types of eating disorder content may still come up; while TikTok has specifically blocked the hashtag “#anorexia” eating disorder influencers are still be promoted on TikTok by misspelling hashtags, such as “#anotexia” as opposed to “#anorexia”).

²⁹⁶ See *Gun Industry Immunity*, *supra* note 18; see Goguen, *supra* note 18.

²⁹⁷ BRANNON & HOLMES, *supra* note 94.

²⁹⁸ See *Gonzalez v. Google LLC*, 2 F.4th 871, 912–13 (9th Cir. 2021).

²⁹⁹ BRANNON & HOLMES, *supra* note 94.

³⁰⁰ See Johnson, *supra* note 19.

³⁰¹ See Bateman, *supra* note 290.

³⁰² See generally Johnson, *supra* note 19.

ideas because they show those harmful videos while providing few effective ways to stop the content, which is against the reasonable person standard, thus constituting wanton and reckless conduct. User-retention time is the main factor that determines which videos show up on a “For You Page,” and users can either swipe away or stay and watch; there are few ways for a user to stop the content because blocking does not necessarily prevent it.³⁰³ By considering the harm that TikTok perpetuates by paying creators to make potentially harmful content, one can draw parallels between Google sharing revenue with ISIS to state a case withstanding Section 230 in the Ninth Circuit.³⁰⁴ Although Section 230 shields social media platforms from liability, the TikTok algorithm is a special case that should be treated similarly to how guns and tobacco have been judicially regulated in the past.³⁰⁵

TikTok has strong parallels to the way in which Carter carried out harmful acts that ultimately led to her manslaughter conviction.³⁰⁶ TikTok purposely suggests videos based on what increases watch time on the app, which includes intentional rage-baiting and other content that is detrimental to a user’s mental health.³⁰⁷ This may put ideas into a user’s head that the user did not originally have, thus taking over his or her ability to think independently.³⁰⁸ Watch time on these types of videos will increase and TikTok will register this as the user wanting more of that content, pushing the user further and further into a content rabbit hole.³⁰⁹ This is sufficient to satisfy causation because, but for the way the algorithm suggests videos and lacks self-help mechanisms, some users would not have adopted the thoughts of the rabbit holes they are being pushed into instead of consuming content they chose to seek out.³¹⁰ Finally, it is irrelevant that TikTok carries out its harmful actions via the operation of its algorithm because an injury does not have to be the product of physical harm, as it has been shown that such manipulative conduct is sufficient.³¹¹

TikTok has repeatedly demonstrated that its algorithm preys on the intimate thoughts of users and uses those thoughts to their advantage by reinforcing those vulnerable ideations to increase watch time to get users hooked on the app, all

³⁰³ See Jargon, *supra* note 3.

³⁰⁴ See *Gonzalez v. Google LLC*, 2 F.4th 871, 898–99 (9th Cir. 2021).

³⁰⁵ See *Gun Industry Immunity*, *supra* note 18; see Goguen, *supra* note 18.

³⁰⁶ See generally *Commonwealth v. Carter*, 52 N.E.3d 1054 (Mass. 2016).

³⁰⁷ *Investigation: How TikTok’s Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

³⁰⁸ See generally Jargon, *supra* note 3.

³⁰⁹ See *id.*

³¹⁰ See *Investigation: How TikTok’s Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

³¹¹ See *Carter*, 52 N.E.3d at 1063.

for TikTok's monetary gain.³¹² Despite Section 230, TikTok should be held liable for their social wrongs because of the potential for harm on users.³¹³ The marketing tactics of TikTok are similar to firearms manufacturers because those tactics target vulnerable teens for financial gain, yet the platform does not face repercussions when users engage in harmful activities as a result of content seen on the app.³¹⁴ Therefore, courts need to reinterpret the liability provision within Section 230 to impose permanent liability for social media platforms. Such judicial action will prompt Congress to either adopt its interpretation implicitly, or to amend Section 230.

³¹² *Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires*, *supra* note 10.

³¹³ *See generally id.*

³¹⁴ *See Gun Industry Immunity*, *supra* note 18.

