

---

---

ALGORITHMS, FILTERS, AND ANONYMOUS MESSAGING: THE  
ADDICTIVE DARK SIDE OF SOCIAL MEDIA

Elizabeth M. Jaffe\*

**I. Introduction**

Within the constant, ever changing world of social media, a new harm is developing which directly impacts consumers. This is, in effect, creating an environment where consumers are unaware they are being targeted. Recent lawsuits attempting to impose liability on the social media giants for their actions are beginning to call attention to the darker side of social media. While there is no doubt social media has incredible benefits, this new harm is creating an environment with a need for newly created liabilities.

**II. The Dangerous Side of Social Media and Associated Harms**

Lying to consumers about harmful effects is nothing new. For many years, tobacco companies were well aware of the addictive nature of their products, yet it was not until Jeffrey Wigand blew the whistle on them that this knowledge would become public.<sup>1</sup> In 1989, Wigand was hired as vice president of research and development at

---

\*Associate Professor, Atlanta's John Marshall Law School. The author would like to thank Katie McClelland and Jessica Regalado for their invaluable assistance with this article.

<sup>1</sup> See Cassi Feldman, *60 Minutes' most famous whistleblower*, CBS NEWS (Feb. 4, 2016), *archived at* <https://perma.cc/9J27-5AJ4> (exposing Wigand's company, Brown & Williamson, for knowingly selling an addictive product).

Brown and Williamson Tobacco Corp.<sup>2</sup> Four years later, he was fired.<sup>3</sup> He then became one of the most well-known whistleblowers, disclosing to the world the highly addictive nature of tobacco products and the tobacco companies were well aware of it.<sup>4</sup> In fact, the tobacco companies were actively making the cigarettes stronger and more addictive.<sup>5</sup> They even genetically altered the crops to contain twice as much nicotine as before.<sup>6</sup> The addictiveness of tobacco has been compared to the addictiveness of social media.<sup>7</sup>

Similarly, Joe Rannazzisi, head of the Office of Diversion Control of the DEA, discovered major problems within the opioid industry that led to rampant addiction.<sup>8</sup> Rannazzisi blew the whistle on three companies, Cardinal Health, McKesson, and AmerisourceBergen, finding they were responsible for the distribution of hundreds of illegitimate orders for opioids that involved millions of pills.<sup>9</sup> For example, one pharmacy located in Kermit, West Virginia, ordered 9 million hydrocodone pills over a period of two years despite

---

<sup>2</sup> See Carol M. Bast, *At What Price Silence: Are Confidentiality Agreements Enforceable?*, 25 WM. MITCHELL L. REV. 627, 628 (1999) (stating Wigand served as Vice President from January 1989 to March 1993).

<sup>3</sup> See *id.* (declaring Wigand's employment ended in 1993 and thereafter violated confidentiality agreements).

<sup>4</sup> See *id.* (recognizing the far-reaching consequences Wigand's decision to whistle blow had on the tobacco industry).

<sup>5</sup> See *id.* at 689 (disclosing how nicotine is addictive and can lead to other health problems). See also *How Big Tobacco made cigarettes more addictive*, TRUTH INITIATIVE (Jan. 23, 2018), archived at <https://perma.cc/HA6k-GZKZ> (adding bronchodilators to cigarettes to help smoke reach lungs more easily and adding sugars, flavors, and menthol to improve the taste).

<sup>6</sup> See *How Big Tobacco made cigarettes more addictive*, *supra* note 5 (articulating that "Big Tobacco companies genetically engineered their tobacco crops to contain two times the amount of nicotine and adjusted their cigarette design so that the nicotine delivered to smokers increased by 14.5 percent.").

<sup>7</sup> See Megan McCluskey, *How Addictive Social Media Algorithms Could Finally Face a Reckoning in 2022*, TIME (Jan. 4, 2022), archived at <https://perma.cc/WZC7-CATC> (comparing Big Tobacco to Big Tech in terms of addictiveness and profitability because social media companies are focused on making platforms more engaging to users and finding ways to keep the user scrolling through content to create a stream of revenue through advertisements).

<sup>8</sup> See Bill Whitaker, *Ex-DEA agent: Opioid crisis fueled by drug industry and Congress*, CBS NEWS (Oct. 17, 2017), archived at <https://perma.cc/CPT9-B9BC> (describing Rannazzisi's role in the DEA and major problems associated with overprescribing opioid drugs).

<sup>9</sup> See *id.* (stating "[t]he three largest distributors are Cardinal Health, McKesson, and AmerisourceBergen."). These distributors control approximately 85 or 90 percent of the drugs going downstream. *Id.*

the population of Kermit being only 392 people.<sup>10</sup> Instead of reporting the suspicious orders to the DEA, the orders were simply filled.<sup>11</sup> Moreover, Purdue Pharma, was also fraudulently representing OxyContin as being less addictive than it actually was.<sup>12</sup> The opioid manufacturers were also making massive donations to policymakers in exchange for opioid regulations.<sup>13</sup> Rampant addiction to opioids developed as a result, and to date, there have been over 200,000 deaths over two decades.<sup>14</sup>

Now, just as the whistle was blown on the tobacco and pharmaceutical companies, the whistle is being blown on social media.<sup>15</sup> Frances Haugen was formerly employed at Facebook as a data scientist.<sup>16</sup> She studied how misinformation was spread through Facebook's algorithm and how foreign adversaries exploited it.<sup>17</sup> She worked on Facebook's integrity operation focused on user safety.<sup>18</sup>

---

<sup>10</sup> *See id.* (recognizing the suspicious orders distributors were placing).

<sup>11</sup> *See id.* (observing the millions of suspicious pill orders that were filled). Jim Geldhof, a DEA veteran, describes a common occurrence that a pharmacy had "bought 50 times an amount that a normal pharmacy purchases and they are in a town of 5,000 people." *Id.*

<sup>12</sup> *See* Karen Feldscher, *What led to the opioid crisis – and how to fix it*, HARV. T.H. CHAN SCH. PUB. HEALTH (Feb. 9, 2022), *archived at* <https://perma.cc/F6NC-RP99> (exposing the profit motive of the pharmaceutical industry, specifically Purdue Pharma).

<sup>13</sup> *See id.* (commenting on the influence opioid manufacturers have on policy decisions through their financial contributions).

<sup>14</sup> *See* Whitaker, *supra* note 8 (addressing the hundreds of thousands of lives taken by opioids).

<sup>15</sup> *See* McCluskey, *supra* note 7 (stating that lack of accountability regarding business practices in the tobacco industry is parallel to social media companies). The same way big tobacco was called out for creating an addictive and profitable product, the same kind of action is needed for social media companies because the platforms they host are unhealthy for users. *Id.*

<sup>16</sup> *See id.* (reporting how Frances Haugen's role as Facebook's product manager allowed her to see how the information gathered from user data was harmful among young users across various platforms that Facebook hosts). *See also* Bobby Allyn, *Here are 4 key points from the Facebook whistleblower's testimony on Capitol Hill*, NPR (Oct. 5, 2021), *archived at* <https://perma.cc/BU9W-SPGZ> [hereinafter Allyn, *Key points from the Facebook whistleblower*] (listing Haugen's career in Big Tech).

<sup>17</sup> *See id.* (highlighting the deceptive practices that are powered by Facebook's algorithm, as well as past controversies with Russia during the Cambridge Analytica case where user data was harvested in an attempt to influence the 2016 presidential election).

<sup>18</sup> *See* Billy Perrigo, *Inside Frances Haugen's Decision to Take on Facebook*, TIME (Nov. 22, 2021), *archived at* <https://perma.cc/7LHL-LNX5> (noting Haugen's desire to work for Facebook only if she was tasked with handling misinformation on the site to ensure safety). For example, one specific concern Haugen had was regarding

Within a few months of beginning employment in 2018, she started seeing major holes in the operation that left developing countries especially vulnerable to misinformation spread by the algorithm.<sup>19</sup> In 2020, Facebook dissolved the integrity committee, which led to Haugen's decision to blow the whistle.<sup>20</sup> After making this decision, she spent another year gathering the documents she would later take to Congress and the SEC.<sup>21</sup> In total, she copied thousands of pages of confidential information.<sup>22</sup> These documents showcase that Facebook knew its platforms are addictive, especially for children.<sup>23</sup> Haugen has also filed eight claims with the SEC stating that Facebook knowingly lied to investors.<sup>24</sup>

---

Facebook's safety measures in Ethiopia which was unacceptable because no action was taken until the misinformation turned violent ensued. *Id.*

<sup>19</sup> *See id.* (outlining that of the 3.2 million hours spent by Facebook counteracting misinformation, only 13% was spent on content that originated outside the United States). This is especially dangerous in places that do not have independent fact-checking programs. *Id.* She first became aware of the problem after looking at content from India, and ethnic violence is being fueled by Facebook content in Ethiopia. *Id.*

<sup>20</sup> *See id.* (reporting that Facebook denies dissolving the committee, and it claims that the individuals involved were spread throughout the company to increase its influence).

<sup>21</sup> *See* Perrigo, *supra* note 18 (indicating that in summer of 2021, Haugen "disclosed tens of thousands of pages of internal Facebook documents to Congress and the Securities and Exchange Commission (SEC)."). "The documents ... sparked a reckoning in September over what the company knew about how it contributed to harms ranging from its impact on teens' mental health and the extent of misinformation on its platforms, to human traffickers' open use of its services." *Id.*

<sup>22</sup> *See* Allyn, *Key points from the Facebook whistleblower*, *supra* note 16 (reporting that before Haugen left Facebook, she made numerous copies of confidential documents about Facebook's platforms which "lay bare exactly what Facebook knew about its products.").

<sup>23</sup> *See id.* (suggesting that Facebook intentionally targets children with an addictive product). One Facebook study done in the U.K. shows that 13.5% of teen girls experienced an increase in suicidal thoughts after using Instagram. *Id.* A separate study showed that "17% of teen girls say their eating disorders got worse after using Instagram." *Id.* Instagram made 32% of teen girls feel worse about their bodies. *Id.*

<sup>24</sup> *See id.* (describing the claims Haugen brought to the SEC). One alleged lie was that Facebook said it did not know organizers of the January 6th siege on the Capitol were using its platform to plan the insurrection. *See* Allyn, *Key points from the Facebook whistleblower*, *supra* note 16. Another lie was the overstatement of Facebook's ability to remove hate speech. *Id.* A third claim regarded Facebook and Instagram's ability to cause real-life harm. *Id.*

Facebook, along with other social media platforms, used “engagement-based ranking” as the basis for their algorithms.<sup>25</sup> The longer a user continues to use the platform, the more content is chosen by the algorithm over content from the user’s friends or people they follow.<sup>26</sup> The engagement-based ranking system prioritizes content that receives higher engagement over content with lower engagement.<sup>27</sup> However, content that receives more engagement is often sensationalized, hateful, or misinformation.<sup>28</sup> Even worse, the more we see information repeated, the more likely we are to believe it is true.<sup>29</sup> This all leads to a perfect storm where false or harmful information is shown to millions of users who are likely to believe in its truth.<sup>30</sup>

Facebook and Instagram are not the only platforms responsible for these negative consequences.<sup>31</sup> YouTube, owned by Google, recently started reporting a “violative view rate” (VVR) for its

---

<sup>25</sup> See Perrigo, *supra* note 18 (opining on Facebook’s engagement-based ranking, commonly known as the algorithm).

<sup>26</sup> See *id.* (detailing how the algorithm works). The algorithm “chooses which posts, out of thousands of options, to rank at the top of users’ feeds.” *Id.*

<sup>27</sup> See Allison Slater Tate, *Facebook whistleblower Frances Haugen says parents make 1 big mistake with social media*, TODAY (Feb. 7, 2022), archived at <https://perma.cc/DF42-5XH4> (demonstrating how algorithms retain engagement). Facebook’s algorithms are choosing what the user focuses on. *Id.* “In the first ten minutes on Instagram, people will see content from their friends or pages they follow . . .” *Id.* After users have been on the site for a few hours, the algorithm will show them “new and engaging content in an effort to keep them there.” *Id.*

<sup>28</sup> See *id.* (stressing how algorithms can start off innocent and develop into problematic content). For example, when a user searches “healthy recipes,” the algorithm could lead them to content about eating disorders rather than simply just a healthy recipe because the eating disorder content received higher engagement. *Id.*

<sup>29</sup> See Perrigo, *supra* note 18 (articulating that continuous exposure to content can lead a user to believe it to be true, often distorting reality). For example, when users are exposed to the same idea that it is “it’s O.K. to be violent to Muslims[.]” this can result in a society becoming destabilized. *Id.*

<sup>30</sup> See *id.* (explaining how when misinformation circulates online, users can become susceptible into believing it to be true). During the 2016 U.S. presidential election, one of Haugen’s friends started to believe conspiracy theories online and Haugen soon realized that she lost him to online misinformation and that she was unable to reach him. *Id.*

<sup>31</sup> See Natasha Lomas, *YouTube’s recommender AI still a horror show, finds major crowdsourced study*, TECHCRUNCH (July 7, 2021), archived at <https://perma.cc/V6AK-66L4> (suggesting that YouTube’s algorithm recommends societal ills, such as “hate speech, political extremism and/or conspiracy junk/disinformation . . .”).

videos.<sup>32</sup> VVR is a measure of the number of views whose content violates YouTube's content policies.<sup>33</sup> In 2017, about twenty eight of every ten thousand videos violated YouTube's policies making the VVR around 0.28%; in 2021, that number dropped to about 0.16% - 0.18%.<sup>34</sup> While this seems like a useful metric, YouTube has not released information about how many of these videos its algorithm still recommends to users.<sup>35</sup> However, in an independent study done by Mozilla, it found that the algorithm did in fact suggest videos that violated YouTube's own policies.<sup>36</sup> Of the videos users reported that they regretted watching 71% were recommended by the algorithm.<sup>37</sup> Regretted videos also received 70% more views compared to non-regretted videos that users watched, which further shows that sensationalized content is often recommended by algorithms.<sup>38</sup>

---

<sup>32</sup> See *id.* (defining "violative view rate" as a metric to disclose the percentage of views, which come from content that violates YouTube's policies).

<sup>33</sup> See *id.* (defining what a violative view rate is).

<sup>34</sup> See *id.* (calculating "[t]he most recent VVR stands at 0.16% - 0.18% — which Google says means that out of every 10,000 views on YouTube, 16-18 come from violative content."). The VVR figure is "down by more than 70% when compared to the same quarter of 2017[.]" *Id.*

<sup>35</sup> See *id.* (noting that the VVR metric may misdirect statistics). VVR is of limited use because Google has not released data to "contextualize and quantify how far its AI was involved in accelerating views of content [that] its own rules state shouldn't be viewed on its platform." See Lomas, *supra* note 31.

<sup>36</sup> See *id.* (detailing the study conducted by Mozilla). The study was done through a browser extension which allowed users to "self-report YouTube videos they "regret" watching. *Id.* A "regretted" video is one where the user self-reports having a bad experience watching. *Id.* The extension then creates a report about whether the video had been recommended by the algorithm. *Id.* Some of these videos include "borderline content," such as low-quality videos, that were more difficult for the algorithm to categorize. *Id.* YouTube itself uses borderline content as a category, but it has not provided an exact definition of what falls inside this category. Lomas, *supra* note 31.

<sup>37</sup> See *id.* (reporting that 71% of regretted videos were recommended by YouTube's algorithm).

<sup>38</sup> See *id.* (providing examples of users regretting recommended videos). One person watching videos about the U.S. military was then recommended a video entitled "Man humiliates feminist in viral video." *Id.* Another person was recommended a video about gun rights after watching a video about software rights. *Id.* After watching an "Art Garfunkel music video," a user was recommended a video titled "Trump Debate Moderator EXPOSED as having Deep Democrat Ties, Media Bias Reaches BREAKING Point." *Id.* Overall, the content recommended by the algorithm disproportionately selects sensationalized content that is intended to bring about triggering or misinforming content because content of this type brings views despite lacking in quality. See Lomas, *supra* note 31.

Similar to Facebook, regrettable content was found to be recommended at much higher rates in “non-English speaking countries.”<sup>39</sup> Finally, some have even gone as far as to say that YouTube’s algorithm automizes radicalization, but the truth is that there is not enough data to definitively prove this point.<sup>40</sup>

A document shared with the New York Times by a member of “TikTok’s engineering team in Beijing” offers an unprecedented look at how the platform’s algorithm works, through a document labeled “TikTok Algo 101.”<sup>41</sup> Like other platforms, TikTok takes into account a user’s likes and comments to recommend content.<sup>42</sup> It also examines other video information such as “captions, sounds, and hashtags” to recommend similar videos.<sup>43</sup> However, the main driving force behind the algorithm is the time a user spends watching each video.<sup>44</sup> The released document states that the ultimate goal of the app is to add active users by maximizing retention and time spent on the app.<sup>45</sup> TikTok Algo 101 also displays a simplified mathematical equation of how content is chosen by the algorithm.<sup>46</sup>

People usually choose to put their best foot forward on social media and create highly curated feeds.<sup>47</sup> Other users will see this content through rose-colored glasses that makes them feel worse about their own lives.<sup>48</sup> Negative effects of this repeated comparison include

---

<sup>39</sup> See *id.* (reporting that on YouTube users regretted watching 60% more videos in non-English speaking countries).

<sup>40</sup> See *id.* (identifying that in 2017 European politicians accused YouTube of “automating radicalization” when content from ISIS, a terrorist organization, was being widely spread across the platform).

<sup>41</sup> See Ben Smith, *How TikTok Reads Your Mind*, N.Y. TIMES (Dec. 5, 2021), archived at <https://perma.cc/T2UM-2K7Q> (revealing the document “TikTok Algo 101,” which was confirmed to be authentic by a TikTok spokesperson).

<sup>42</sup> See *id.* (highlighting the ways that TikTok recommends videos to its users).

<sup>43</sup> See *id.* (including additional factors considered in the recommendation of videos).

<sup>44</sup> See *id.* (reporting that TikTok’s algorithm relies on time spent watching each video to curate the feed to encourage user retention).

<sup>45</sup> See *id.* (identifying that TikTok’s goal is to maximize user retention and continuously add users).

<sup>46</sup> See *id.* (showing the simplified version of TikTok’s user retention algorithm is “Plike X Vlike + Pcomment X Vcomment + Eplaytime X Vplaytime + Pplay X Vplay”).

<sup>47</sup> See Tammy Qiu, *A Psychiatrist’s Perspective on Social Media Algorithms and Mental Health*, STAN. UNIV. HUM. CTR. A.I (Sept. 14, 2021), archived at <https://perma.cc/ZQF9-7FM2> (highlighting that users tend to curate feeds specialized to them because it can provide the user with rewarding stimuli of content that is tailored for them).

<sup>48</sup> See *id.* (detailing negative side-effects to users that may reduce the user’s self-esteem, depressed mood, and general decrease in life satisfaction, which may lead to

decreased self-esteem and life satisfaction as well as depression.<sup>49</sup> Children who spend over five hours on social media per day “are nearly twice as likely to have suicidal tendencies” than those who spend under an hour online.<sup>50</sup> One study found that as little as two hours a day was the tipping point where teens began to report adverse effects of social media on their mental health.<sup>51</sup> To manage these negative feelings, many people turn to cyberbullying as a coping mechanism.<sup>52</sup>

When a user receives likes, comments, and other notifications dopamine is released.<sup>53</sup> However, the user never knows when their posts will receive engagement, and this leads to a rush similar to gambling and drug addiction.<sup>54</sup> Then, the user will repeat the behavior

---

other harmful coping mechanisms such as cyberbullying to combat self-esteem issues). *See also* Christine M. Stabler, *The Effects of Social Media on Mental Health*, PENNMEDICINE LANCASTER GEN. HEALTH (Sept. 1, 2021), *archived at* <https://perma.cc/94LH-54LC> (explaining that “picture-perfect” social media posts create false illusions and lead to unfair comparisons).

<sup>49</sup> *See* Qiu, *supra* note 47 (concluding that exposure to “frequent and extreme upward social comparison” can lead to negative side effects such as “erosion of self-esteem, depressed mood, and decreased life satisfaction.”).

<sup>50</sup> *See* Andrew Solomon, *THE MYSTIFYING RISE OF CHILD SUICIDE*, NEW YORKER (Apr. 4, 2022), *archived at* <https://perma.cc/8LVN-YP5D> (according to Jean Twenge, a psychology professor at San Diego State University, teens who spend “five or more hours a day online are nearly twice as likely to have suicidal tendencies as those who spend less than an hour” online).

<sup>51</sup> *See* Joe Graminga, *Social media use may play important role in youth suicide, expert says*, HEALIO (Oct. 5, 2020), *archived at* <https://perma.cc/99ZW-4YAD> (explaining that Professor Jamie Zelany conducted a study of patients in “an intensive outpatient program at Western Psychiatric Hospital” in Pittsburgh, Pennsylvania). The study reported that 67% of participants felt worse about their lives because of social media, 73% felt like they had to post content to make themselves look better to others, 60% felt that they had to post content in order to get likes, and 80% felt personally affected by drama on social media. *Id.*

<sup>52</sup> *See* Qiu, *supra* note 47 (describing how eroded self-esteem from social media can lead to cyberbullying).

<sup>53</sup> *See id.* (depicting the science behind why social media is so addictive). *See also* Jeanne Ricci, *The Growing Case for Social Media Addiction*, CSU (June 28, 2018), *archived at* <https://perma.cc/N3HD-QMUC> (detailing how social media provides rewarding experiences that generate dopamine).

<sup>54</sup> *See* Qiu, *supra* note 47 (describing how “it’s the intermittent absence of the like that keeps us engaged” on social media). *See also* Ricci, *supra* note 53 (reporting that the stimuli activation caused by social media is the same as the stimuli activation in the brains of drug addicts).



in hopes that the next time they post, it will receive the reward they are seeking.<sup>55</sup>

Not only are their brains seeking dopamine, users are also looking for a sense of belonging within a community.<sup>56</sup> Accompanying this is a fear of missing out on what others are enjoying.<sup>57</sup> Not understanding the latest joke or meme is an isolating experience when it feels like everyone gets it, except for you.<sup>58</sup> This can also lead to “impostor syndrome” where a person doubts themselves, their success, and their intelligence.<sup>59</sup>

Not only does social media affect the user’s views of themselves and others, it can cause physical symptoms as well.<sup>60</sup> Anxiety and depression caused by social media addiction can manifest in the form of “nausea, headaches, muscle tension, and tremors.”<sup>61</sup> It can also decrease both quality and quantity of sleep.<sup>62</sup>

The foregoing negative effects are not just theoretical, here are real people with real stories of how social media has negatively impacted their lives.<sup>63</sup> Christopher James Dawley (CJ) was found dead by his parents upstairs in the family home from a self-inflicted gunshot wound to his head.<sup>64</sup> Right before he committed suicide, he texted his best friend “God’s speed,” and posted on Facebook “Who turned out

---

<sup>55</sup> See Qiu, *supra* note 47 (emphasizing social media addiction does not just stem from “the intermittent absence of the like” but also the “pleasure rush of the like”).

<sup>56</sup> See *The Social Dilemma: Social Media and Your Mental Health*, MCLEAN HOSP. (Jan. 21, 2022), archived at <https://perma.cc/AWQ2-NDSG> [hereinafter *Social Dilemma*] (noting how social media users are looking for a sense of belonging within a community online, receiving positive feedback from their content).

<sup>57</sup> See *id.* (describing how the “fear of missing out” plays a large role in social media addiction).

<sup>58</sup> See *id.* (describing how missing out on online experiences can lead to anxiety and depression).

<sup>59</sup> See *id.* (defining the experience of impostor syndrome as “feeling chronic self-doubt and a sense of being exposed as a ‘fraud’ in terms of success and intellect.”).

<sup>60</sup> See *id.* (explaining how social media use can lead to physical ailments).

<sup>61</sup> See *id.* (describing the physical ailments as “nausea, headaches, muscle tension, and tremors”).

<sup>62</sup> See *Social Dilemma*, *supra* note 56 (explaining how a 2018 British study linked social media use to disrupted and decreased sleep).

<sup>63</sup> See Samantha Murphy Kelly, *Their teenage children died by suicide. Now these families want to hold social media companies accountable*, CNN BUS. (Apr. 19, 2022), archived at <https://perma.cc/KK5Z-GQN4> (describing multiple stories from parents who lost children to suicide as a result of how social media companies run their platform).

<sup>64</sup> See *id.* (detailing the background of CJ’s suicide).

the light?”<sup>65</sup> With his phone still in his hand, he pulled the trigger.<sup>66</sup> His mom summed up the terrible tragedy when she said, “[h]e was so addicted to it that even his last moments of his life were about posting on social media.”<sup>67</sup> Before the suicide, CJ would often stay up until three in the morning on social media and became both “sleep deprived and obsessed with his body image.”<sup>68</sup>

Another young boy, Ian Mitchell, was also found dead from a self-inflicted gunshot wound.<sup>69</sup> It appears as though he was playing Russian roulette with a loaded gun while recording himself on Snapchat (“Snap”), and his phone was found next to his lifeless body.<sup>70</sup>

Selena Rodriguez, only eleven years old, committed suicide after a two-year-long social media addiction to Instagram and Snapchat.<sup>71</sup> Selena began communicating with older men, who asked her to “send sexually explicit images” to them.<sup>72</sup> She complied, and the images were eventually sent to her classmates with severe social repercussions for Selena.<sup>73</sup>

#### A. *Anonymous Messaging*

Anonymous messaging is far from a new advent, having taken several faces over the last decade, however, all forms met the same issue- the messages inevitably get nasty and lead to harm.<sup>74</sup> An earlier

---

<sup>65</sup> See *id.* (describing CJ’s last text to his friend and last Facebook post before his suicide).

<sup>66</sup> See *id.* (depicting the scene of CJ’s suicide where he had a “22-caliber rifle in one hand, his smartphone in the other” when he shot himself).

<sup>67</sup> See *id.* (quoting what CJ’s mother saw when she found her son after he committed suicide).

<sup>68</sup> See *id.* (recounting CJ’s online social media habits leading up to his suicide).

<sup>69</sup> See Kelly, *supra* note 63 (summarizing the life of 16-year-old Ian Mitchell who died of a self-inflicted gunshot wound while on snapchat).

<sup>70</sup> See *id.* (describing how Ian’s parents found a video of him on Snapchat that looked like he was playing Russian roulette with a gun).

<sup>71</sup> See *id.* (summarizing the death of eleven-year-old Selena Rodriguez and her prior social media addiction).

<sup>72</sup> See *id.* (describing how Selena used social media to share sexually explicit images with older men).

<sup>73</sup> See *id.* (describing how the images Selena shared with the older men were then leaked to her classmates causing severe social repercussions at school).

<sup>74</sup> See Olivia Smith, *Why Yolo, the Anonymous Snapchat App, is a Disaster Waiting to Happen*, GRIT DAILY (Feb. 28, 2020), archived at <https://perma.cc/M42W-VAUR> (detailing how anonymous platforms often lead to bad behavior, like cyberbullying); Alexia Maddox, *Sendit, Yolo, NGL: anonymous social apps are taking over once*

version of anonymous messaging took shape in Formspring, a website turned app which allowed both registered and unregistered persons to anonymously send questions to users.<sup>75</sup> Though the forum eventually shut down, it first was linked to several teen suicides around the world.<sup>76</sup> A 14-year-old New York boy uploaded a video describing his Formspring bullying and regret about using the app just months before committing suicide.<sup>77</sup> Similarly, a 15-year-old English girl stepped in front of a train less than 24 hours after receiving cruel messages on the forum.<sup>78</sup>

Another anonymous messaging platform, YikYak, began innocently as an app geared towards college campuses.<sup>79</sup> It allowed users to make virtual bulletin posts that could be up or down voted and commented on completely anonymously within a five-mile radius of

---

*more, but they aren't without risks*, THE CONVERSATION (July 14, 2022), archived at <https://perma.cc/TS5H-G4V3> (describing how anonymous social media apps have “harmful consequences for young people, including cyberbullying . . .”). See also Jade Wickes, *Why do we keep going back to anonymous messaging apps?*, THE FACE (July 27, 2022), archived at <https://perma.cc/EF65-Q74U> (describing the history of anonymous social media platforms).

<sup>75</sup> See Dan Seifert, *Social question and answer site Formspring to shut down on March 31st*, THE VERGE (Mar. 15, 2013), archived at <https://perma.cc/Q2ZJ-YUF9> (describing that at the inception of Formspring’s creation, the website would allow users to make posts anonymously or with registered accounts and receive answers from other registered or anonymous users).

<sup>76</sup> See Casey Newton, *Killer app: why do anonymous Q&A networks keep leading to suicides?*, THE VERGE (Sept. 17, 2013), archived at <https://perma.cc/7BRZ-GCLJ> (detailing that several teens were bullied through this anonymous forum and the treatment they received led to them committing suicide). Formspring’s predecessor, Ask.fm, received scrutiny due to the posts that were directed towards a young teen that ultimately committed suicide, this was the same story for Formspring once it launched. *Id.*

<sup>77</sup> See Xgothemo99xx, *It Gets Better, I promise!*, YOUTUBE (May 4, 2011), archived at <https://perma.cc/2V4F-ULW6> (displaying the remorse he felt for using a social networking site that attributed to his bullying and suicidal behavior and directing viewers not to use the forum). See also Newton, *supra* note 76 (alluding to the teen’s YouTube video where he displayed regret for using the forum, and his subsequent suicide after posting the video). “Formspring began working . . . to redesign its interface to discourage bullying. . . . But later that year, a 14-year-old boy from New York committed suicide after recording a video for the ‘It Gets Better’ project in which he describes being bullied on Formspring.” *Id.*

<sup>78</sup> See Newton, *supra* note 76 (setting forth another example of the detrimental effects bullying on the forum has on teens).

<sup>79</sup> See John Patrick Pullen, *You Asked: What Is Yik Yak?*, TIME (Feb. 4, 2015), archived at <https://perma.cc/2NTF-GKEF> (addressing how at the app’s inception, it was an innocuous social media platform where college students could post anonymously, similar to that of a bulletin board).

the user's location.<sup>80</sup> The app quickly rose in popularity, but simultaneously became notorious for cyberbullying and threats.<sup>81</sup> While some users only felt the impact in the way of scars on their mental health, others documented physically harming themselves, and even worse being encouraged to “finish the job.”<sup>82</sup> More recently, YOLO and LMK—third party apps that Snapchat users could separately download and then integrate to their stories—allowed complete anonymity for users to submit questions or comments directly to other users that upon replying would become public posting to their story.<sup>83</sup> As all those before it, the anonymous messaging quickly turned toxic for users who were being harassed and bullied but could not figure out by who.<sup>84</sup> An Oregon sophomore was taunted and verbally attacked through the two apps for months before eventually committing suicide.<sup>85</sup>

After being linked to cyberbullying, harassment, and suicides most platforms took action; Springform consulted MIT in an attempt to revise their interface to discourage bullying and further joined an anti-bullying campaign.<sup>86</sup> The day after a suicide lawsuit was filed,

---

<sup>80</sup> See *id.* (demonstrating how the platform works by geofencing to keep grade school campuses off of the app and allowing the college users to interact with anonymous posts by voting up or down); Isabella Simonetti, *Yik Yak, Notorious for Student Bullying and Harassment, Relaunches With Mysterious Ties to a College Marketing CEO*, OBSERVER (Nov. 8, 2021), archived at <https://perma.cc/5HY9-7PYQ> (asserting that “[U]sers within a five-mile radius of one another to post on an anonymous discussion board. Yik Yak also allows users to up or downvote posts.”).

<sup>81</sup> See Pullen, *supra* note 79 (recognizing that although the app was popular among college students, it soon became a platform for users to exhibit “cyberbullying, racist barbs and hate speech.”).

<sup>82</sup> See Simonetti, *supra* note 80 (discussing the harmful posts and the effect they had on the mental health of the students, in addition to one student being encouraged to take his life).

<sup>83</sup> See Rachel Sharp, *Snap blocks messaging apps YOLO and LMK after Oregon mom sued them over the suicide of her 16-year-old son who suffered months of abuse from anonymous bullies*, DAILY MAIL (May 12, 2021), archived at <https://perma.cc/297S-GUPF> (detailing the way in which the third-party apps function to provide anonymity to engage with the user who was posting the interactive forum across two platforms).

<sup>84</sup> See *id.* (addressing how the anonymous forum quickly became full of humiliating and explicit content and was used to bully other users without letting them know who hid behind the screens).

<sup>85</sup> See *id.* (acknowledging the life of Carson, an Oregon sophomore, who committed suicide after being “subjected to months of cyberbullying via YOLO and LMK[.]”).

<sup>86</sup> See Newton, *supra* note 76 (characterizing Formspring and MIT’s collaboration as a response to the hateful comments posted on their platforms to help combat online bullying and harassment).

Snapchat suspended both YOLO and LMK pending an investigation and started Here For You, a mental health initiative to aid their users cope with anxiety and depression.<sup>87</sup> Nearly a year later, Snapchat announced new policies which included banning anonymous messaging from third party apps as well as stiffer screening.<sup>88</sup> YikYak took measures by geofencing the app, users would be restricted if they tried opening the app near a middle or high school; however, once the teen left school campus the app functioned as normal.<sup>89</sup> The app did however notably keep internal track of anonymous users' location and IP address making sure users weren't left completely unaccountable; in one case they turned over said information to police in connection with a college campus murder.<sup>90</sup>

Despite the clear pattern, and the troubles and backlash its predecessors have faced, in late June of this year 'Sendit for Instagram' launched, with thousands of downloads it sat at number one non-gaming app on the Appstore for nearly a week.<sup>91</sup>

### B. Filters

Filters are literally changing the way that we see ourselves. Edited photos of celebrities in magazines and on billboards have been

---

<sup>87</sup> See *Bride v. Snap Inc.*, No. 21-cv-03473-JD, 2021 U.S. Dist. LEXIS 149150, at \*2 (N.D. Cal. Aug. 9, 2021) (mentioning that the claims for relief arise out of the harassment and cyberbullying that took place on YOLO and LMK). See also Smith, *supra* note 74 (announcing Snapchat and Yolo's initiative to help support users struggling with mental health issues as a result of online bullying).

<sup>88</sup> See *Announcing New Policies for Snap's Developer Platform*, SNAP INC. (Mar. 17, 2022), archived at <https://perma.cc/4USV-S9F4> (acknowledging Snapchat's newly created policies following harmful content caused by their application features). Snap determined in an investigation that the anonymous message platforms that were infiltrating their platform still posed risks for users. *Id.*

<sup>89</sup> See Pullen, *supra* note 79 (discussing the parameters on the usage of YikYak). Brooks Buffington, the co-founder of YikYak stated the app was designed for college-age kids because "it requires a certain level of maturity[.]" *Id.*

<sup>90</sup> See Sean Sposito, *Yik Yak Shares User Information With Police*, GOV'T TECH. (Nov. 16, 2015), archived at <https://perma.cc/85LW-ZY73> (noting that YikYak's privacy policy allows for tracking of user's information). After a student at the University of Missouri made threats on YikYak, the app handed over the student's location and IP address to local authorities. *Id.*

<sup>91</sup> See Sarah Perez, *Anonymous social apps shift their attention to Instagram in the wake of Snapchat's ban*, TECHCRUNCH (June 29, 2022), archived at <https://perma.cc/KZW7-2E88> (commenting that although previous anonymous Q&A apps have led to problematic social experiences, 'Sendit for Instagram' was still able to have a successful launch).

around for a while, but seeing edited versions of ourselves and our friends is a new phenomenon.<sup>92</sup> When filters first appeared on social media they were gimmicky and fun.<sup>93</sup> Today, filters are not just used to make a picture black and white or add dog ears and a nose to people's faces.<sup>94</sup> They are used to whiten teeth, tan skin, hide fat, and fix any other imperfections.<sup>95</sup> This is especially dangerous for vulnerable teens who already have great difficulty in developing a positive self-image.<sup>96</sup>

Filters are not just the norm, sometimes they are the default. In 2019, Apple released an update that would later be named "Eye Contact."<sup>97</sup> This feature makes it appear that the person is looking directly into the device's camera, even when they are looking elsewhere.<sup>98</sup> The concerning part about this filter is not that it exists, but that it is automatically turned on, and users must manually disable it in their device's settings if they do not want it on.<sup>99</sup> Despite the feature existing since 2019, many people did not know about it until

---

<sup>92</sup> See *Social Dilemma*, *supra* note 56 (researching the societal impacts that social media has had on both teen and adult's confidence). "Apps that provide the user with airbrushing, teeth whitening, and more filters are easy to find and easier to use." *Id.*

<sup>93</sup> See Tate Ryan-Mosley, *Beauty filters are changing the way young girls see themselves*, MIT TECH. REV. (Apr. 2, 2021), archived at <https://perma.cc/C48A-RLPF> (demonstrating that prior to recent developments in augmented-reality technology, filters were used for their humor and entertainment value).

<sup>94</sup> See *id.* (reiterating that filters were used to "play a kind of virtual dress-up: change your face to look like an animal, or suddenly grow a mustache . . .").

<sup>95</sup> See *Social Dilemma*, *supra* note 56 (emphasizing that these edits made it hard for "teens to tell what's real and what isn't . . .").

<sup>96</sup> See *id.* (observing social media's influence on mental and physical health). "Social media use can affect users' physical health even more directly. Researchers know the connection between the mind and the gut can turn anxiety and depression into nausea, headaches, muscle tension, and tremors." *Id.*

<sup>97</sup> See Ben Cost, *Chilling 'fake eye contact' Facetime feature exposed in viral tweet*, N.Y. POST (July 11, 2022), archived at <https://perma.cc/CK6U-MCQ8> (reporting Apple's latest software as originally being named "FaceTime Attention Correction," but the name changed after the public's dispute to Apple's new augmentation software which was defaulted to be on and can only be disabled manually).

<sup>98</sup> See *id.* (describing how the filter works). The filter "digitally alters people's image so they seem like they're looking directly at the person they're Facetimeing with even if they're staring elsewhere." *Id.*

<sup>99</sup> See *id.* (stating that many concerns have been raised because the "Eye Contact" filter switches on automatically; the filter needs to be manually disabled to turn it off).

2022 when it was brought to light on social media.<sup>100</sup> Understandably, people reacted negatively knowing that their images were being altered without their consent or knowledge.<sup>101</sup>

Not only are filters changing our perceptions of our faces, they are changing our actual faces as well. An app called FixMe “allowed users to mark up their faces as a cosmetic surgeon might” right before surgery.<sup>102</sup> Recently, plastic surgeons have seen an increase in requests from patients to look like the filtered versions of themselves.<sup>103</sup>

Another filter is partially responsible for several deaths and numerous injuries.<sup>104</sup> In 2013, Snapchat released a filter that showed the speed at which the user was traveling at that time.<sup>105</sup> Two years later, the driver of a car, Christal McGee told the passengers in her car that she was trying to reach one hundred miles per hour so that she could post a photo displaying the speed.<sup>106</sup> The speed limit was fifty five miles per hour, moments later, while traveling one hundred and seven miles per hour she crashed into Wentworth Maynard’s car

---

<sup>100</sup> See *id.* (highlighting that although the filter launched in 2019, it did not become well known to the general public until 2022).

<sup>101</sup> See *id.* (indicating that users were surprised that their images were being altered without their knowledge with most users exclaiming that the feature was “creepy”).

<sup>102</sup> See Ryan-Mosley, *supra* note 93 (explaining that the app was later redesigned, and filters that promoted plastic surgery were banned because of the public debate about negative impact on individuals’ mental health and body dysmorphia).

<sup>103</sup> See *Social Dilemma*, *supra* note 56 (asserting that many users have sought out plastic surgeons to make their appearance resemble the filters seen on social media platforms).

<sup>104</sup> See Bobby Allyn, *Snapchat Ends ‘Speed Filter’ That Critics Say Encouraged Reckless Driving*, NPR (June 17, 2021), archived at <https://perma.cc/55V9-T49N> (declaring that Snapchat also released a problematic filter that resulted in numerous deaths and injuries). The Snapchat filter took a real time speed reading the user was traveling at which prompted others to compete for who had the highest number. *Id.*

<sup>105</sup> See *id.* (pointing to the recent lawsuits Snapchat faced from families who had their loved-ones injured or killed while using the filter to demonstrate they were driving at excessive speeds).

<sup>106</sup> See Tony Marrero, *Snapchat video appears to show driver going 115 mph on night of crash that killed 5*, TAMPA BAY TIMES (Oct. 28, 2016), archived at <https://perma.cc/G95B-BW2S> (reporting that in the lawsuit against Snapchat an allegation that Christal McGee was using the Snapchat speed filter on her phone at the time of the crash). See also Alison Frankel, *Crash victim gets new chance to prove Snap ‘speed filter’ caused accident*, REUTERS (Mar. 17, 2022), archived at <https://perma.cc/Q4EA-XDGW> (demonstrating that McGee, among others, used the filter to post a photo on Snapchat showing how fast she was driving).

causing him permanent brain damage.<sup>107</sup> She then posted a selfie on her Snapchat stating that she was “lucky to be alive.”<sup>108</sup> Maynard and his wife sued Snap alleging negligent design of the speed filter.<sup>109</sup> Snap claimed that McGee’s criminal conduct broke the causal chain, and that Snapchat’s terms of use forbids using the app for illegal purposes.<sup>110</sup> Further, the filter is accompanied by a warning prohibiting use of the filter while driving, and Snap claims that it owes no duty to prevent users from utilizing the filter while driving.<sup>111</sup> The Georgia Court of Appeals found Snap not liable.<sup>112</sup> The Maynard’s have appealed to the Georgia Supreme Court.<sup>113</sup>

That same year, Amonie Barton, Gia Scavo Abgarian, and Candice Walker died in a fiery car accident after they crashed into the back of a tractor trailer after they had been sending Snapchats showing their speed throughout the night.<sup>114</sup>

In 2016, Jolie Bartolome posted a Snapchat of her boyfriend, Pablo Cortes III, driving one hundred and fifteen miles per hour less than ten minutes before the crash.<sup>115</sup> The speed limit where the crash

---

<sup>107</sup> See Frankel, *supra* note 106 (noting that McGee was going approximately fifty-seven miles per hour over the speed limit when she crashed, causing permanent brain damage to Maynard).

<sup>108</sup> See *id.* (highlighting the seriousness surrounding the crash and the injuries that resulted).

<sup>109</sup> See *id.* (discussing how Maynard sued Snapchat to hold them liable for promoting this dangerous filter).

<sup>110</sup> See *id.* (detailing the Maynards’ lawsuit against Snapchat). The Maynards’ allege that Snapchat should have known that “the filter would prompt users like McGee to drive dangerously yet did not build in protections to reduce that risk.” *Id.*

<sup>111</sup> See *id.* (outlining Snapchat’s objection to the Maynard’s claims). Snapchat argued that users “received a pop-up alert warning them not to use snap while driving.” See Frankel, *supra* note 106.

<sup>112</sup> See *Maynard v. Snapchat, Inc.*, 870 S.E.2d 739, 743 (Ga. 2022) (detailing that the Court of Appeals sided with the trial court’s decision and conclusion that Snap was not liable because the misuse of the manufacturer’s product does not extend in this situation). The court concluded that Snapchat owed “no duty to Wentworth . . .” *Id.* at 755.

<sup>113</sup> See *id.* at 743 (describing the procedural posture of the case).

<sup>114</sup> See *Did Snapchat play a role in deaths of 3 young women?*, ABC11 (Feb. 12, 2016), archived at <https://perma.cc/874C-8DKL> (summarizing the details of the deaths of Amonie Barton, Gia Scavo Abgarian, and Candice Walker).

<sup>115</sup> See Michael Paluska, *115.6 MPH Snap posted by girlfriend of driver that caused fatal crash that killed 5 people*, ABC ACTION NEWS (Oct. 28, 2016), archived at <https://perma.cc/KNU3-P7GM> (detailing the circumstances surrounding the deaths of Jolie Bartolome and Pablo Cortes III). See also Marrero, *supra* note 106



occurred was fifty miles per hour.<sup>116</sup> Cortes lost control of his vehicle when he crossed the median and collided head-on with another car driven by Marianela Murillo.<sup>117</sup> Murillo and her two children were killed in the accident along with Bartolome and Cortes.<sup>118</sup> Two others were left in critical condition.<sup>119</sup>

Finally, in 2017, Jason Davis was driving a car also containing Hunter Morby and Landen Brown when the car crashed into a tree, killing all three boys.<sup>120</sup> A Snapchat sent by one of the boys showed that the car was traveling one hundred and twenty three miles per hour.<sup>121</sup> Minutes later, the car crashed into the tree that killed them, and police estimate that the car was traveling one hundred and thirteen miles per hour at the time of impact.<sup>122</sup> The boys' parents brought suit against Snap for the wrongful death of their sons.<sup>123</sup>

Some teenagers believed that they would receive a digital prize for posting speeds exceeding one hundred miles per hour.<sup>124</sup> Since these tragedies, Snap has downgraded “the speed feature from a ‘filter’ to a ‘sticker’” which was done to decrease its prominence.<sup>125</sup> It also

---

(discussing the video just taken prior to the accident where the driver was exceeding the 50-mph speed limit).

<sup>116</sup> *See id.* (noting the speed limit where the accident occurred was 50 mph).

<sup>117</sup> *See Snapchat video clocks car going 115 mph before deadly Hillsborough crash*, WFLA (Oct. 28, 2016), *archived at* <https://perma.cc/8V4V-A8FW> (highlighting the details of the crash and how the victims were struck head on).

<sup>118</sup> *See id.* (reporting on the deaths resulting from the accident).

<sup>119</sup> *See id.* (listing the conditions of the two survivors in the crash).

<sup>120</sup> *See Lemmon v. Snap, Inc.*, 440 F. Supp. 3d 1103, 1105 (C.D. Cal. 2020) (*rev'd and remanded by*, 995 F.3d 1085, 1088 (9th Cir. 2021)) (giving the details surrounding the car accident that killed Jason Davis, Hunter Morby, and Landen Brown). The Ninth Circuit reversed and remanded the District Court's ruling on various grounds, but mainly on the ground that there is no settled law on the issue of causation for accidents alleged to be caused by a social media company. *Id.* at 1113.

<sup>121</sup> *See id.* at 1105 (detailing that on the night of the accident, a snapchat captured the boys traveling at a speed of 123 mph). “‘Within minutes’ of the 123-MPH-Snap, the car ran off the road and crashed into a tree.” *Id.*

<sup>122</sup> *See id.* (stating that investigators estimated the speed of the vehicle to be 113 mph at the time of the crash).

<sup>123</sup> *See id.* (asserting that Snapchat was a critical cause of the accident). The plaintiffs assert claims for the wrongful death of their son. *See Lemmon v. Snap*, 440 F. Supp. 3d at 1105.

<sup>124</sup> *See Allyn*, *supra* note 104 (describing a California lawyer's lawsuit which alleged that some teenage Snapchat users believed they would be awarded digital prizes for recording a speed in excess of 100 mph).

<sup>125</sup> *See id.* (detailing the transition of the speed feature from a “filter” to a “sticker”).

placed a limit on speeds that could be posted at thirty-five miles per hour.<sup>126</sup>

### C. Trending Challenges

Challenges are a more recent social media trend; they begin as a single user filming themselves doing *something* and challenging others to recreate it.<sup>127</sup> More popular challenges go viral and see hundreds of thousands of people following suit by uploading their clip.<sup>128</sup> Most challenges serve no purpose other than gaining views and the subject matter ranges from completely harmless in nature- such as randomly standing still or makeup transformations to extremely dangerous.<sup>129</sup>

One such challenge is Car surfing, the activity of which is not entirely new, having been around since the mid-eighties.<sup>130</sup> However, it has recently regained popularity through the viral challenge which consists of recording a person riding unrestrained on the roof of a moving vehicle, generally standing straight up, “surfing.”<sup>131</sup> The trend

---

<sup>126</sup> See *id.* (summarizing Snapchat’s decision to cap the top speed for which a post could be shared).

<sup>127</sup> See Liam Curtis, *Why TikTok’s Pass Out Challenge is so dangerous – don’t bother with it!*, HITC (Feb. 10, 2020), archived at <https://perma.cc/3GBN-BU7S> (describing different social media challenges and the dangers that some raise).

<sup>128</sup> See *id.* (discussing the viral “Pass Out” and “Choking Challenges” going viral and the extreme dangers associated with them).

<sup>129</sup> See Daniel Victor, *Mannequin Challenge Is the New Viral Video Sensation You Probably Can’t Avoid*, N. Y. TIMES (Nov. 7, 2017), archived at <https://perma.cc/S8VG-2FYU> (discussing how the mannequin challenge asked groups of people to freeze in place while a cameraman walked around the scene to record the details). See also Payal Bajoria, *21 Most Popular Social Media Challenges*, IIM SKILLS (Nov. 20, 2022), archived at <https://perma.cc/YD45-EUdT> (describing the don’t rush challenge which asked women to start the video bare faced in pajamas, put an object such as a hair brush up to the lens taking up the whole frame, and upon pulling back to the item being completely done up including hair and makeup).

<sup>130</sup> See Mariana Zapata, *20 Most Dangerous Social Media Challenges to Avoid*, FAMILYMINDED (Aug. 12, 2022), archived at <https://perma.cc/D6A8-88MK> (recognizing that car surfing has been around since the mid 1980’s). Social media has helped the activity reached more teens and young adults, resulting in at least seven deaths since 2016. *Id.*

<sup>131</sup> See *Inside the Dangerous Trend of Car Surfing and How a Single Joyride Can Kill*, INSIDE EDITION (Oct. 11, 2018), archived at <https://perma.cc/U22V-FXMV> (highlighting the increasing popularity of the car surfing trend). The dangerous challenge draws the attention of young adrenaline junkies. *Id.*

is dangerous for obvious reasons but sees most of its injuries when the moving vehicle takes a turn or hits something in the road launching the surfer off the roof.<sup>132</sup> Depictions in pop culture detailing every step of the activity have correlated with a steady rise in car-surfing deaths and serious injury since 2000.<sup>133</sup> One particularly gruesome challenge in Florida involved four teens, one car surfing, one driving, and the other two riding front and back seat passenger.<sup>134</sup> The vehicle was travelling at 70 mph before it flipped and crashed sending the eighteen-year-old car surfer into a ditch.<sup>135</sup> He sustained fatal injuries and was later pronounced dead, additionally, the backseat passenger was trapped inside the car and died after the car burst into flames.<sup>136</sup> The front seat passenger and driver were able to escape before the crash, but footage of the challenge was shown at the latter's trial for vehicular homicide in connection to the crash.<sup>137</sup> Similarly, a sixteen-year-old Georgia boy

---

<sup>132</sup> See *id.* (stressing that “[n]ot everyone walks away unscathed”). “Texas student Dakota Revell, then 14, was almost killed when she slipped off the back of a moving car back in 2016 . . . [s]he was rushed to the hospital, where she underwent life-saving brain surgery . . . [and underwent a] long recovery process.” *Id.* See also Zapata, *supra* note 130 (identifying the car surfing challenge as one of the more dangerous viral challenges). “Most people die when curves, bumps or stops throw them off the car, causing them to hit the pavement or other things around them.” *Id.*

<sup>133</sup> See Arthur Wang et al., *Neurological injuries from car surfing*, 4 J. NEUROSURGERY PEDIATRICS 408, 413 (July 17, 2009) (arguing that media aimed at impressionable children and adolescents should be made responsibly and include appropriate warnings due to the steady rise in car-surfing fatalities). National statistics from the three states with the highest fatality rates from car surfing which demonstrated a steady rise in car-surfing deaths since 2000. *Id.* “An interesting finding is the chronological overlap between the introduction of *Grand Theft Auto* games, *Jackass*, and YouTube, and the rise and peaks in car surfing fatalities.” *Id.* at 412. The largest peak in car-surfing fatalities coincided with the release of YouTube in 2005. *Id.* “Previous authors have reported average ages between 17 years and 15.7 years.” *Id.*

<sup>134</sup> See Gary Taylor, *DeBary man charged in car-surfing deaths get probation*, ORLANDO SENTINEL (Feb. 3, 2012), archived at <https://perma.cc/7TLN-AQXZ> (explaining the general details of the car crash caused by “car surfing”).

<sup>135</sup> See Paul Thompson, *If I die, remember this was his idea’: Shocking video captures final moments of US teenagers killed while ‘car surfing’ at 70mph*, DAILYMAIL (Feb. 7, 2012), archived at <https://perma.cc/YP7N-W8ER> (reporting that two teens, Carlos Velazco and Hunter Perez, were seen clinging to the side of their friend’s car as it traveled 70 mph in a 35-mph zone). In the video, one of the teens says, “[i]f I died, remember that this is Carlos’s idea.” *Id.*

<sup>136</sup> See *id.* (describing the cause of death of the two teenagers). Hunter Perez’s mother stated that “[she] get[s] no pleasure from your pain, but you must be held accountable for two lives lost” when addressing the driver of the car. *Id.*

<sup>137</sup> See Taylor, *supra* note 134 (reporting that the deputies were able to recover the recording from the passenger’s cellphone which showed the ‘car surfing’ until the

fell off the jeep he was surfing on, was run over and succumbed to his injuries days later at a local hospital.<sup>138</sup>

Even the accidents associated with this challenge that do not end in death still have tragic outcomes.<sup>139</sup> A ten-year-old boy sustained two broken shoulders, a broken pelvis, fractured ribs and had internal bleeding after falling and being run over by the car he was surfing on.<sup>140</sup> More common is brain trauma, like the seventeen-year-old Minnesota boy who had to be placed in a medically induced coma for a month following his car surfing accident.<sup>141</sup> The severity of his brain injury required a piece of his skull to be removed and he now suffers from chronic headaches.<sup>142</sup>

Like car surfing, the concept underlying the blackout challenge has been around for some time.<sup>143</sup> However, it peaked in popularity on TikTok in 2021.<sup>144</sup> The challenge dares participants to asphyxiate themselves- either by holding their breath or choking themselves out- in an attempt to get a euphoric high just prior to passing out.<sup>145</sup> While

---

crash). The video was played at trial, however, the driver escaped prison time since he was a youthful offender. *Id.*

<sup>138</sup> See WXIA Staff, 'Car surfing' teen dies after falling, getting run over, ALIVE (Nov. 23, 2015), archived at <https://perma.cc/AR3Q-P9B7> (describing a comparable 'car surfing' tragedy). In this instance, a 16-year-old boy was sitting on the hood of a Jeep 'car surfing', when he fell off and subsequently, was run over. *Id.*

<sup>139</sup> See Chelsea Robinson, *Police: 10-year-old 'surfing' on top of parents' car falls, is run over*, WBALTV11 (Apr. 8, 2019), archived at <https://perma.cc/L5SC-EETN> (describing the injuries sustained when a young boy fell off the roof of his parents car); Anjali Subbaswamy, *Summer's here; so is car surfing*, ALBUQUERQUE J. (June 5, 2018), archived at <https://perma.cc/GK5T-3ZNL> (explaining how a teen was in a coma and had part of his skull removed after a 'car surfing' accident).

<sup>140</sup> See Robinson, *supra* note 139 (highlighting the injuries the boy sustained after 'car surfing'). The young boy was 'car surfing' on the roof of his parent's car, when he slipped and was run over. *Id.*

<sup>141</sup> See Subbaswamy, *supra* note 139 (describing a similar 'car surfing' accident). 'Car surfing' "caused 58 deaths between 1990 and 2008 . . ." *Id.*

<sup>142</sup> See *id.* (noting how 'car surfing' is an "uncommon but significant cause of brain injury and death" but is also completely preventable).

<sup>143</sup> See Selina Maycock, *What is the Blackout Challenge on TikTok? Parents warned over fatal online trend*, GOODTOKNOW (Aug. 24, 2022), archived at <https://perma.cc/X5LA-7TE8> (reporting that the dangerous blackout challenge has been around since 2008).

<sup>144</sup> See Sarah Felbin et al., *The 'Blackout Challenge' Has Resurfaced On TikTok, And It's Still Just As Dangerous As It Was 16 Years Ago*, WOMEN'SHEALTH (Oct. 27, 2022), archived at <https://perma.cc/P7SX-RXN6> (discussing how the 'blackout challenge' is one of many dangerous viral TikTok trends).

<sup>145</sup> See Maycock, *supra* note 143 (detailing how the Blackout challenge encourages TikTok users of all ages to pass out "to obtain a brief euphoric state[.]").

the number of related deaths in the last decade is unknown lawsuits filed in the last year suggest at least seven children- all under the age of fifteen- died attempting the challenge.<sup>146</sup> Some parents were faced with the scaring experience of finding their child dead or clinging to life with objects such as shoestrings, belts, and other ligatures around their children's neck.<sup>147</sup> But what's more is the graphic videos of their child's last moments that were left behind, intended to be posted to social media with the challenge hashtag.<sup>148</sup> While TikTok has taken steps to ban the challenge hashtag, videos remain accessible and deaths continue to pile up worldwide- including in Italy, the UK, Australia and the U.S.<sup>149</sup>

---

<sup>146</sup> See Andrew Paul, *DON'T TRY THIS AT HOME TikTok's dangerous 'Blackout Challenge' allegedly already killed seven kids*, INPUT (July 11, 2022), archived at <https://perma.cc/SJ22-8DKL> (explicating how families are filing lawsuits against TikTok and its parent company for “introducing the stunt to their children via algorithmic promotions intended to keep users on the platform as long as possible.”). These lawsuits want TikTok to be held accountable for improperly monitoring and adjusting the algorithms which perpetuate dangerous challenges, such as the Blackout Challenge. *Id.* But see CDC Study Warns of Deaths Due to the “Choking Game”, CDC (Feb. 14, 2008), archived at <https://perma.cc/L5NY-8WSN> (describing a 2008 study which showed that at least 82 children between the age of 11 and 16 attempting the game). The average age of deaths resulting from the “Choking Game” is 13 with 87% of these deaths being males. *Id.*

<sup>147</sup> See Melissa Chan, *Kids Are Playing the 'Choking Game' to Get High. Instead, They're Dying*, TIME (Mar. 12, 2018), archived at <https://perma.cc/8BKT-3H39> (pointing out that due to “millions of how-to videos on asphyxiation only a finger’s tap away, kids are more likely to play” the choking game with common household objects).

<sup>148</sup> See Maycock, *supra* note 143 (recounting the numerous stories of parents whose children have suffered tragedies associated with the blackout challenge). PEOPLE magazine states “four children under the age of 12 have already died after attempting to copy the challenge ... [and] has led to over 80 deaths in the past, per the study released by the Centers for Disease Control and Prevention.” *Id.*

<sup>149</sup> See *id.* (explaining that TikTok has attempted to block the challenge and when users search for the Blackout Challenge they are notified of its ban). But see Ben Lovejoy, *TikTok blackout challenge said to be responsible for the deaths of seven children*, 9TO5MAC (July 8, 2022), archived at <https://perma.cc/W6N9-ZBXC> (alleging that TikTok has not removed the Blackout Challenge videos). See e.g., Crispian Balmer, *Italy tells TikTok to block users after death of young girl*, REUTERS (Jan. 22, 2021), archived at <https://perma.cc/7EK6-3N8U> (addressing that since TikTok’s supposed ban of the Blackout Challenge, an Italian girl died of asphyxiation after attempting to participate in the challenge); Ellie Hutchings, *What happened to Archie Battersbee? His injury, coma and battle to stay on life support explained*, GOODTOKNOW (Aug. 8, 2022), archived at <https://perma.cc/8TV9-6SWM> (pointing to another life lost to the Blackout Challenge, Battersbee, a 12-year-old boy from Southend-on-Sea in Essex); Leesa Smith, *He was found lifeless on the verandah after watching a TikTok challenge*, KIDSPOT (June 17, 2021),

### III. Where Are we Now?

#### A. Current Cases

Suits against social media platforms have come in an array of claims such as negligence, products liability, failure to warn, and wrongful death.<sup>150</sup> Yet for nearly 25 years internet platforms have enjoyed near full protection from such cases thanks to Section 230 of the Communications Decency Act.<sup>151</sup> More recently though, the Section 230 shield has begun to wear in its reign of absolute immunity

---

*archived at* <https://perma.cc/H8ZS-7GHT> (affirming the deathly consequences of the Blackout Challenge). Boyd-Gergely, a 14-year-old boy from New South Wales, Australia, died from the challenge while he was home alone and his mother recounts: “[i]t’s almost like it was a stunt gone wrong. There is no way in my mind that he would do that intentionally.” *Id.* See also Charmaine Patterson, *Parents Sue TikTok After Daughters Die Doing 'Blackout Challenge': 'We Want People to Be Aware'*, PEOPLE (July 21, 2022), *archived at* <https://perma.cc/L9U9-ZBLF> (reiterating the worldwide spread of the Blackout Challenge). Two young girls, Walton and Arroyo, passed away after attempting the challenge in July 2021 and February 2021, respectively. *Id.*

<sup>150</sup> See *Bride v. Snap Inc.*, No. 21-cv-03473-JD, 2021 U.S. Dist. LEXIS 149150, at \*2 (N.D. Cal. Aug. 9, 2021) (alleging Snap and other social media apps negligence in failing to implement anti-bullying and harassment policies). See e.g., *Rodriguez v. Meta Platforms, Inc. et al.*, No. 3:22-cv-00401, 2022 WL 190807, at ¶¶ 47-51 (N.D. Cal., Jan. 20, 2022) (seeking to prove, in their pleadings, that the court should hold the defendants, formally Facebook Inc., “liable for their own speech and their own silence in failing to warn of foreseeable dangers arising from anticipate [sic] use of their products.”); *Tanton v. Meta Platforms, Inc. et al.*, No. 3:22-cv-00411, 2022 WL 2057366, at ¶¶ 12–13 (M.D. Tenn., June 6, 2022) (alleging in their complaints that the Defendant, Meta Platforms Inc., failed to “provide adequate warnings to minor users and their parents of the danger of mental, physical, and emotional harms arising from the foreseeable use of their social media products.”); *Lemmon v. Snap, Inc.*, CV-19-4504-MWF 2022 U.S. Dist. LEXIS 83399, at \*3 (alleging Snapchat’s speed filters were being used in dangerous ways, so Plaintiffs sued Snapchat on the grounds “that the design of the Speed Filter itself encouraged Plaintiffs to engage in reckless driving.”).

<sup>151</sup> See 47 U.S.C.A. § 230(c) (stating the protection offensive material has within the Communications Decency Act). “[A]ny action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected[.]” *Id.* See also Isaiah Portiz, *Tech’s Online Content Shield Dented by Product Liability Claims*, BLOOMBERG LAW (July 22, 2022), *archived at* <https://perma.cc/89J2-3YE8> (indicating that Section 230 enables a broad protection shield for some of the world’s largest internet companies).

for internet platforms.<sup>152</sup> The district court in *Lemmon v. Snap, Inc.*, 2022 U.S. Dist. LEXIS 83399 initially dismissed in favor of Snap citing Section 230 as a bar.<sup>153</sup> The Ninth Circuit however reversed holding that Section 230 did not bar the parent's claim.<sup>154</sup>

While there have been few wins in the column, in a small step there, a British court found that Meta contributed to a 14-year-old's suicide.<sup>155</sup> However, the social media platform faced no recourse because the case was to determine cause of death rather than any civil or criminal liability against Meta.<sup>156</sup> Similarly, The Georgia Supreme Court reversed and remanded the dismissal of a complaint against Snap for the death resulting from its speed filter, allowing the case to live on.<sup>157</sup>

In arguably the largest step yet towards holding social media platforms liable, The US Supreme Court granted certiorari in a case that includes challenges to the immunity held under section 230.<sup>158</sup>

---

<sup>152</sup> See Portiz, *supra* note 151 (highlighting a shift in the Section 230 liability shield in 2021).

<sup>153</sup> See Wyatt Larkin, *Lemmon v. Snap, Inc.: Ninth Circuit Chips Away at Tech Companies' Section 230 Immunity*, HARV. J.L. & TECH. (Oct. 20, 2021), archived at <https://perma.cc/H2A7-8RGN> (reporting that the case was dismissed by the district court pursuant to Section 230 of the Communications Decency Act).

<sup>154</sup> See *id.* (summarizing the Ninth Circuit court's ruling). "The Ninth Circuit answered 'no' and found for the plaintiffs, holding that Section 230 did not bar their claim and reversed the lower court's granting of Snap's motion to dismiss." *Id.*

<sup>155</sup> See Rachyl Jones, *Social Media Contributed to the Death of a 14-Year-Old Girl, a U.K. Court Finds*, OBSERVER (Oct. 3, 2022), archived at <https://perma.cc/L24Z-KE9Q> (attributing the cause of the teen's suicide to Meta).

Content on social media platforms contributed to the suicide of a 14-year-old girl, a British court found Sept. 30. Molly Russell, who died in 2017, interacted with 2,100 posts related to suicide, self-harm and depression in the six months leading up to her death—or 12 posts per day—according to Meta's data . . .

*Id.*

<sup>156</sup> See *id.* (noting that although Meta was found to be responsible, there will be no civil, criminal, or monetary punishments imputed upon Meta because the scope of the proceeding was limited to determining the teen's cause of death).

<sup>157</sup> See *Maynard v. Snapchat, Inc.*, 870 S.E.2d 739, 756 (Ga. 2022) (concluding that the appellate court did not address whether the trial court erred in not addressing the plaintiff's challenge to the trial court's proximate cause analysis). See also Ryan Bennett, *Oh Snap! Georgia Supreme Court Revives Suit Against Snapchat for Alleged Faulty Speed Filter*, JDSUPRA (Apr. 29, 2022), archived at <https://perma.cc/D6T7-H74M> (highlighting the holding of the Supreme Court of Georgia in *Maynard*'s case allowing the suit to move forward).

<sup>158</sup> See Amy Howe, *Court agrees to hear nine new cases, including challenge to tech companies' immunity under Section 230*, SCOTUSBLOG (Oct. 3, 2022), archived at <https://perma.cc/E7VV-8QEM> (announcing the Supreme Court agreeing to review

The result of which could mean reducing or even eliminating the current protections afforded to internet platforms through the Communications Decency Act.<sup>159</sup>

Several new cases have recently been filed across the U.S., with at least 30 since June of this year.<sup>160</sup> Being that the cases are so new to the dockets, little substantive movement has occurred, however, along with the upcoming *Gonzalez v. Google LLC* decision, could result in a huge change in the tides of holding social media accountable.<sup>161</sup>

### B. Pending Legislation

Several bills have been introduced in Congress related to social media companies' algorithms. The Filter Bubble Transparency Act would require that certain internet platforms allow users to interact with the platform without being manipulated by an algorithm.<sup>162</sup> Those platforms must notify the user that it utilizes an algorithm that

---

*Gonzalez v. Google*, in which one of the central issues will be determining the scope of protection that Google will be afforded under Section 230).

<sup>159</sup> See *id.* (summarizing the division in the U.S. Court of Appeals for the 9<sup>th</sup> Circuit regarding the scope of protection Section 230 affords tech companies). Section 230 currently offers protections to social media companies that are host-sites for content created by others, however, the challenge uprooting the protections stems on whether algorithms that target or “recommend” content to users is within the statute’s scope of protection. *Id.*

<sup>160</sup> See Hugo Guzman, *Social Media Platforms Hit With Surge in Mental Health Lawsuits*, ALM (Aug. 17, 2022), archived at <https://perma.cc/F77D-8LNA> (stating that over 30 cases have been filed since the start of June of 2022).

<sup>161</sup> See *Gonzalez v. Google LLC*, 2 F.4th 871, 913 (9th Cir. 2021) (concluding that Section 230 protects social media companies more than Congress initially envisioned). “Whether social media companies should continue to enjoy immunity for the third-party content they publish, and whether their use of algorithms ought to be regulated, are pressing questions that Congress should address.” *Id.*

<sup>162</sup> See Filter Bubble Transparency Act, 117 S. 2024, 117th Cong. § 2(4)(A) (2021) (indicating that the Act would apply to “any public-facing website, internet application, or mobile application, including a social network site, video sharing service, search engine, or content aggregation service.”). Exclusions apply when the owner of the platform did not employ more than 500 employees in the previous six months, “averaged less than \$50,000,000 in annual gross receipts . . .” per year in the previous three years, and “collects or processes on an annual basis the personal data on less than 1,000,000 individuals . . .” *Id.* at § 2(4)(B). A separate exclusion applies to platforms created for the sole purpose of research and is “not made for profit either directly or indirectly.” *Id.*



provides content based on data gathered from that user.<sup>163</sup> It must be a clear and conspicuous notice that the user dismisses in order to attract the attention of the user.<sup>164</sup> The Bill would also require a “prominently placed icon” which would allow the user to easily switch between the opaque algorithm and the input-transparent algorithm.<sup>165</sup>

One example of a feature that would comply with the Bill is Twitter’s “sparkle icon.”<sup>166</sup> This icon, appearing in the upper right corner of the app, allows users to switch between “Home” and “latest Tweets.”<sup>167</sup> “Home” is an opaque algorithm that recommends tweets based on data, and “latest Tweets” is an input-transparent algorithm which shows only content posted by accounts the user follows in chronological order.<sup>168</sup>

A major problem with holding social media companies liable can be found in Section 230 of the Communications Decency Act.<sup>169</sup>

---

<sup>163</sup> See *id.* at § 3(b)(1)(A) (mandating that platforms must provide adequate notice to users that their information is being utilized by an algorithm).

<sup>164</sup> See *id.* (clarifying that adequate notice to a user must be in a, “clear, conspicuous manner” to ensure the user is sufficiently put on notice that their information is being utilized by the platform).

<sup>165</sup> See *id.* at § 2(6)(A) (2021) (stating that an opaque algorithm is defined as one that presents content to users based on data not explicitly provided by the user for that purpose). Information relating to a user’s age is used only to censor content based on that individual not being old enough. See Filter Bubble Transparency Act at § 2(6)(B)(ii). Input-transparent algorithms do not base content on user-specific data. *Id.* at § 2(5)(A).

<sup>166</sup> See Press Release, Thune, Colleagues Reintroduce Bipartisan Bill to Increase Internet Platform Transparency, (June 10, 2021) (outlining that the Filter Bubble Transparency Act would allow users to switch between a personalized timeline and a chronological timeline, similar to the “sparkle icon” Twitter currently uses).

<sup>167</sup> See Sarah Perez, *Twitter’s latest update makes it easier to escape its recommendation algorithm (Update: Twitter removed it!)*, TECHCRUNCH (Mar. 10, 2022), archived at <https://perma.cc/WG7T-NYKY> (stating that Twitter’s “sparkle icon” feature that allows users to switch between latest tweets and home by swiping back and forth).

<sup>168</sup> See *id.* (explaining that before adding the sparkle icon, Twitter allowed users to switch between chronological and algorithm-generated content). However, the icon makes the switch more accessible and well-known to users. *Id.* While the tab feature for switching back-and-forth has been removed, users can still utilize the sparkle icon to switch their feeds. *Id.*

<sup>169</sup> See Michael D. Smith & Marshall Van Alstyne, *It’s Time to Update Section 230*, HARV. BUS. REV. (Aug. 12, 2021), archived at <https://perma.cc/5S88-6JV3> (acknowledging that Section 230(c)(1) was put in place to keep social media platforms in business because they “generate social benefits”). However, Section 230(c)(1) is in conflict with Section 230(c)(2) because “[w]hen you grant platforms complete legal immunity for the content that their users post, you also reduce their incentives to proactively remove content causing social harm.” *Id.* See also 47

The Justice Against Malicious Algorithms Act would modify Section 230 and would hold platforms liable when the provider knew or should have known that a recommendation of content was made based on personal data, the recommendation was made recklessly, and the recommendation “materially contributed to a physical or severe emotional injury.”<sup>170</sup> Smaller platforms would be exempt.<sup>171</sup>

Another major problem is that current law used to protect children is outdated.<sup>172</sup> The Kids Online Safety Act (KOSA) creates a duty of social media companies to act in the best interest of minor users of their platform.<sup>173</sup> This duty requires the company to “prevent and mitigate the heightened risks of physical, emotional, developmental, or material harms” potentially created by the platform.<sup>174</sup> KOSA would require social media companies to create a public report at least once a year identifying foreseeable risks of harm to minors.<sup>175</sup> These companies would also be required to allow minors and their parents to modify recommended algorithms and place limitations on certain

---

U.S.C.A. § 230(c)(2) (summarizing how section 230 protects internet platforms from liability related to content posted by third parties while also allowing them to remove potentially harmful content without facing liability).

<sup>170</sup> See Justice Against Malicious Algorithms Act of 2021, 117 H.R. 5596, 117th Cong. § (2)(a)(2)(1) (2021) (noting that a provider is liable when they make a personalized recommendation with knowledge or recklessness). Subsection (c)(1) does not apply to information garnered through an interactive computer service when “such recommendation materially contributed to a physical severe emotional injury to any person.” *Id.*

<sup>171</sup> See *id.* at § (2)(a)(2)(A) (explaining that the act would exempt providers of platforms that have had less than five million site visitors per month in at least three of the last twelve months). Another exemption applies to content provided to a user based on a content-specific search that the user ran. *Id.* at § (2)(a)(2)(B).

<sup>172</sup> See Sarah L. Bruno et al., *The Kids Online Safety Act*, REEDSMITH (Feb. 24, 2022), archived at <https://perma.cc/59XL-2YJR> (addressing that the current law can be found in the Children’s Online Privacy Protection Act). This law was enacted in the 1990’s, prior to the creation of all the major social media platforms children use today. *Id.* Furthermore, the law only applies to children under the age of 13. *Id.*

<sup>173</sup> See Kids Online Safety Act, 117 S. 3663, 117th Cong. § 3(a) (2022) (describing that KOSA enforces a duty on companies to act in the best interest of the minor users).

<sup>174</sup> See *id.* at § 3(b) (mentioning harms that are protected by the act include “self-harm, suicide, eating disorders, substance abuse,” physical harm, bullying, harassment, sexual exploitation, and promotion of illegal substances).

<sup>175</sup> See *id.* at § 6(a)(1) (mentioning a yearly report would be required to identify the foreseeable risk of harm to minors). This report would be composed by an independent third-party through a reasonable inspection and describing the safety measures addressing the harms to minors. *Id.*

content.<sup>176</sup> Other features of KOSA include tools for parents to track how much time their child spends on the platform and the ability to opt out of features likely to keep children online longer such as auto play.<sup>177</sup> A large part of KOSA is designed to protect children's data from being used against them.<sup>178</sup> All of these features would be automatically applied for minors and must be opted out of, which is opposite of most current safety features.<sup>179</sup>

In addition to federal legislation, several states have proposed legislation to protect children and hold social media companies liable.<sup>180</sup> The Social Media Platform Duty to Children Act was introduced in the California State Assembly.<sup>181</sup> The Bill creates a duty for platforms to not addict children under the age of eighteen to the platform through the use of the child's data like their engagement in the platform, or using that child's data for platform development, design, or implementation.<sup>182</sup> In the statute, addiction is defined as a use of social media that indicates a preoccupation, obsession,

---

<sup>176</sup> See Cat Zakrzewski, *Senators unveil children's online safety bill after months of pressure on Silicon Valley*, WASH. POST (Feb. 16, 2022), archived at <https://perma.cc/NY63-QFWP> ([explaining the safeguards that come along with KOSA](#)). “The bill would require companies to provide parents and minors with controls and create new obligations for platforms to address self-harm, eating disorders and other content that might harm children and teens.” *Id.*

<sup>177</sup> See Bruno et al., *supra* note 172 (mentioning KOSA would require companies to have “easy-to-use” tools to protect children from harmful content online). See Zakrzewski, *supra* note 176 (describing that companies could create easy to use tools that allow parents to track the amount of time that their children spend on social media and opt out of auto play features which extend screen time).

<sup>178</sup> See Zakrzewski, *supra* note 176 (suggesting that the bill would allow research to be conducted regarding how technology could potential harm children and teens). Technology has resulted in “a sense of powerlessness, a loss of control from kids themselves, ... this bill is empowering those children and their parents to take back control and the power over their lives online.” *Id.*

<sup>179</sup> See *id.* (clarifying how KOSA will require companies to create these features but parents will have the option to limit or ban certain types of content).

<sup>180</sup> See *id.* (explaining how states such as California have been developing their legislatures to augment child safety online). A bill introduced by California Assemblywoman Buffy Wicks and Assemblyman Jordan Cunningham called the “California Age-Appropriate Design Code Act,” which is intended to increase safety measures in protecting children from the dangers posed to online content, has impacted children more since the beginning of the COVID-19 pandemic. *Id.*

<sup>181</sup> See Cal. St. Assemb., A.B. 2408, 2021-22 St. Assemb., Reg. Sess. (Cal. 2022) (highlighting how the Act was first introduced on February 17, 2022, and has been amended four times since).

<sup>182</sup> See *id.* at § 3(j)(3) (defining “child” for the purpose of this act). See also *id.* at § 2(i) (listing the ways companies use their platforms to cause harm to children through addiction).

withdrawal from life, or a difficulty in the cessation of using the platform despite the desire to do so, and that causes “physical, mental, emotional, developmental, or material harms to the user.”<sup>183</sup> The Bill only affects platforms owned by a business that made \$100 million in gross revenue the previous year.<sup>184</sup>

Businesses have a few options to continue avoiding liability.<sup>185</sup> First, in the initial version of the bill prior to the amendments that modified the scope of liability, businesses could remove features that were known or should have been known to cause addiction in children before January 1, 2023.<sup>186</sup> Second, they can perform quarterly audits of features with the potential to cause child users to become addicted and correct those features within thirty days of the audit.<sup>187</sup> Third, they can refuse to allow children from using their platform, which business groups, such as TechNet, believe is a real possibility.<sup>188</sup>

The initial bill provided that if a child becomes addicted, their parents would be allowed to recover a civil penalty up to \$25,000 per violation per year.<sup>189</sup> In a class action, each member of the class would

---

<sup>183</sup> See *id.* at § 3(j)(2)(A–B) (explaining the repercussions of addictions from social media platforms).

<sup>184</sup> See *id.* at § 3(k)(1) (outlining exemptions from the bill). Video game platforms are also excluded from the bill. *Id.* at § 3(k)(2).

<sup>185</sup> See Cal. St. Assemb., A.B. 2408, at § 3(d) (listing ways platforms can evade civil penalty).

<sup>186</sup> See *id.* at § 3(e)(6)(A) (explaining that before the most recent amendment, the bill limited the way social media platforms can be created and conducted). The bill attempted to have companies be held liable for damages to children “including, but not limited to, suicide, mental illness, eating disorders, emotional distress, and costs for medical care, including care provided by licensed mental health professionals.” *Id.*

<sup>187</sup> See *id.* § 3(d) (explaining a way to avoid civil liability). The companies should audit their platform to detect the features that can cause addiction. *Id.* at § 3(d)(1-2).

<sup>188</sup> See Adam Beam, *California parents could soon sue for social media addiction*, AP NEWS (May 23, 2022), archived at <https://perma.cc/6QSP-2EVA> (inferring that to avoid liability, social media companies may have to cease operations for children). Many businesses, including TechNet, believe that if the bill becomes law “social media companies and online web services would have no choice but to cease operations for kids under 18 and would implement stringent age-verification in order to ensure that adolescents did not use their sites.” *Id.*

<sup>189</sup> See Cal. St. Assemb., A.B. 2408 § 3(c) (noting how actions for relief may be pursued by the Attorney General, district attorney, county counsel or city attorney and may seek a civil penalty of no more than \$250,000 per a knowing and willing violation).

be guaranteed at least \$1,000.<sup>190</sup> Additionally, they may recover actual and punitive damages.<sup>191</sup> While this seemed like a step in the right direction, legislators have removed the part of the bill that allowed the child's parents to sue.<sup>192</sup> This Bill would have been the first to create a cause of action for parents regarding algorithms.<sup>193</sup> However, the amended version of this bill exclusively provides the authority to obtain relief if an action is brought by the Attorney General, district attorney, and other designated attorneys representing the state or county.<sup>194</sup>

The Minnesota House introduced a Bill that would hold social media companies liable for using an algorithm to present content to anyone they knew or should have known was under the age of eighteen.<sup>195</sup> Only platforms with more than one million accounts could be held liable under the Bill.<sup>196</sup> A few exemptions exist including parental controls and any content blocked from the user by the platform because it is inappropriate or harmful.<sup>197</sup> Another exemption is for any content created by federal, state, and local government or any school, college, or university.<sup>198</sup>

---

<sup>190</sup> See *id.* at § 3(d)(4) (recognizing that the legislature considered, but decided not to include a minimum for class actions)

<sup>191</sup> See *id.* at § 3(e)(6)(A) (referring to a previous version of the bill, which discussed potential damages that parents could obtain on behalf of their children).

<sup>192</sup> See *id.* (striking out most of the provisions that allowed for recovery of \$25,000 per violation per year, the class action guarantee, and a parent's ability to sue on behalf of their child). See also Beam, *supra* note 188 (predicting correctly that lawmakers would remove the ability of parents to sue social media companies).

<sup>193</sup> See *id.* (pointing out how the bill could have been the first in California to allow parents to sue social media platforms if children became addicted to their products). See also Zakrzewski, *supra* note 176 (reporting that a bill of this kind would be the first in the United States which could cause ripple effects throughout the tech industry across the United States).

<sup>194</sup> See Cal. St. Assemb., A.B. 2408 § 3(b) (authorizing select individuals who may initiate relief under this statute).

<sup>195</sup> See H.F. 3724, 2022 Leg., 92nd Sess. § (1)(2)(a–b) (Minn. 2022) (prohibiting owners of social media platforms with over 1 million users to use algorithms to target “user-created content at an account holder under the age of 18,” if the owner of the social media platform knew or had reason to know that the account holder was younger than 18).

<sup>196</sup> See *id.* at § (1)(2)(a) (providing when a company must comply with the bill).

<sup>197</sup> See *id.* at § (1)(3)(a) (listing the exemptions which are available due to parental or internal controls). These controls are “designed to control access of the account of a minor to filter content for age-appropriate material, that suggest, promote, or rank otherwise accessible content . . .” *Id.*

<sup>198</sup> See *id.* at § (1)(3)(b) (explaining that public and private educational institutions will not be required to comply with the bill and are considered expressly exempt).

### *C. Conclusion*

In light of the foregoing, it is evident the dark side of social media must be addressed. There is a plethora of lawsuits that likely never came to fruition due to the lack of remedies available. With consumers being directly targeted on social media and the widespread use of social media, the time has come to consider new remedies and liability. In the ever-changing social media landscape, it is crucial for the law to continue to transform.