



Children's Advocacy Institute
UNIVERSITY OF SAN DIEGO SCHOOL OF LAW

SB 287 (Skinner)

Social Media Platform Child Suicide & Drug Overdose Prevention Act



“Facebook’s internal research is aware that there are a variety of problems facing children on Instagram; they know that severe harm is happening to children.”

- Frances Haugen,
Facebook Whistle Blower

Q Does the bill present a “yes or no” question for legislators?

It does. SB 287 simply asks whether social media platforms should be able, free from financial accountability, to use technologies, devices, or techniques that the platforms know or, if operating with ordinary care in its affairs, should know will cause children to die by suicide, facilitate the sale of fentanyl to children, cause eating disorders in children, or make children addicted to social media.

If a legislator believes platforms should be free from financial accountability when they knowingly or carelessly cause these harms to children, with the costs of addressing these knowingly inflicted harms imposed entirely on parents, taxpayers, teachers, and children, they should vote no. If they think platforms should be financially accountable when they knowingly harm children or fail to take basic care not to injure them, they should vote yes.

Q What is the public policy aim of the bill?

The aim of the bill is simple: to prevent some of the world’s largest and most profitable social media platform companies from using technologies in ways they know or should know are deadly and harmful to children, especially girls.

77.14%

Of drug deaths among teens in 2021 caused by Fentanyl.

84%

and 38% of tweens use social media

Q How many children use social media?

By the time a child user reaches age thirteen, advertisers have collected an average of 72 million data points about him/her.[3] While thirteen is allegedly the minimum age federal law permits users to create social media accounts, 50% of ten-to-twelve-year-olds, 33% of seven-to-nine-year-olds,[4] and 38% of “tweens” use social media.[5] By the time they are teenagers, about 92% of young people in the U.S. are active on social media.[6]

Q What is it about the business models of social media platforms that will always prioritize “user engagement” over everything else, including child safety?

Social media platforms offer users access to their platforms and services for free. Platforms earn “substantially all” of their revenue through digital advertising. “Advertising isn’t just a way for [Facebook] and its ilk to perhaps earn a little bit of revenue in between hosting family photos and personal musings. It’s the very purpose of the site’s existence, and the same goes for Twitter and LinkedIn.” “There’s a reason why [Facebook’s] 10-K filing with the U.S. Securities and Exchange Commission (SEC) uses the acronym ARPU, as in average revenue per user” and why investors track “monthly median engagement levels” measuring increases or decreases in the average number of likes, comments posted, and ads clicked.[7]

“[Facebook's] leadership knows ways to make Facebook and Instagram safer and won't make the necessary changes because they have put their immense profits before people.”

- Frances Haugen,
Facebook Whistle Blower



Thus, the “product” that social media companies sell to advertisers is the opportunity to display targeted ads to the platform’s billions of users. The time and attention users spend on such platforms (a metric that companies call “user engagement”) are critical ingredients in sustaining social media platforms’ profitability. The more time spent on the platform, the more ads will be seen, and the more valuable the advertising becomes.

Plus, the more time a user spends on the platform, the more data the platform can derive about the user, which, in turn, can be used to keep the user on the platform. “Facebook’s data, algorithms and use of machine learning have continued to improve This means that users are seeing more and more relevant content, and this of course leads to more engagement on the platform.”[8]

It is important to note that “user engagement” does not necessarily track a user’s subjective enjoyment or objective health and well-being. In fact, user enjoyment and well-being are sometimes negatively correlated to user engagement. In other words, many users spend *even more time* on social media when engaging with content they don’t like[9] or objectively makes them unhealthier.[10] For these reasons, at their core, social media companies are financially motivated to use design features that maximize user engagement instead of features that would increase user enjoyment or well-being but decrease engagement.[11] Addicted consumers are particularly profitable because their consumption behavior goes beyond normal engagement levels. The extra profit generated by addicted consumers is known as the “addiction surplus.”[12] Because addicted consumers are generally more profitable than consumers who are not addicted, businesses are motivated to create addiction and capture the “addiction surplus.”[13]

Here is the most alarming facet of this business model when it comes to child safety: there is no natural endpoint to the motivation of social media companies to engage their child users more and more and more. *To achieve revenue and market share growth every quarter — to not peak or decline — platforms must figure out ways to keep children on their platforms even more than they are now, and so on, and so on, without end.*



What are the legislative options for addressing harms caused by social media platforms?

Broadly speaking, there are five legislative options:

(i) *PUC-like oversight and regulation of all operations.* We have rejected this as politically impractical, innovation-stifling, and ineffective, as social media companies are vastly larger than the utilities the amply resourced PUC has difficulty overseeing.

(ii) *Over-arching statutory oversight regulation of all child-affecting operations.* We have rejected this approach for the same reasons as we rejected (i).

(iii) *Outlawing in statute particular features like auto-scroll.* Some experts have urged us to reject such an approach embedded in statute because the public-facing features are only one part of the addiction/ trafficking equation and so such an approach would be ineffective. Experts also warn that isolating one feature in an inflexibly specific statute would be hard to enforce as the border between a lawful feature and an unlawful one can be elusive, and so innovation suppressing. Likewise, features evolve, so outlawing a feature in year one risks outlawing a feature that would soon be outdated anyway.

(iv) *Do nothing.* As detailed horrifically below, our children are in unprecedented anguish and crisis. Children do not grow up in the long run. They are growing up *now*. Waiting a year for a child is like waiting decades for adults. We cannot claim to be pro-child legislators and citizens and do nothing. Moreover, the core business model of social media companies — getting more and more users and getting those users to stay on the platform as long as possible — is directly at odds with elevating child safety, which is why these companies have not and will never themselves effectively prioritize child safety proportional to our moral obligation to children. As whistleblower Haugen said: **“Until the incentives change, Facebook will not change. Left alone, Facebook will continue to make choices that go against the common good, our common good.”**[14]

(v) *Make it unlawful for platforms to knowingly and carelessly cause the worst and most documented harms – leaving the “how” to the platform.* We have settled on and prefer this enforcement and incentive-based approach which firmly establishes an unlawful result but does not decree how each company tailors its operations to achieve that result.

[1] Facebook Whistleblower, *supra* note 1.



What platforms would be subject to this bill?

To avoid unfairly impacting small- and medium-sized enterprises, this bill only applies to social media platforms controlled by companies earning more than \$100 million in annual revenue. In addition, the bill excludes various types of digital platforms, including video games, email providers, search engines, streaming services, file-sharing and collaboration tools, internal business services, and limited-functionality sites like comment sections and product reviews.

The bill will cover many of the most popular social media platforms used by children, including TikTok, Instagram, and Facebook.



Would the bill apply if a company accidentally created an addictive product or if the company did not and should not have known it was harmful to children?

No. Social media platform companies would not be held liable under this bill solely for practices or features of social media platforms that the platform operator did not know, and had no reason to know, would cause or contribute to child suicides, eating disorders, fentanyl overdoses, or addiction. Only knowingly causing these harms or being so careless as to cause children harm creates liability under the bill. Legally speaking, these are high bars.

**During the time
Instagram use rose to 1
billion users, suicide
among girls age 10 to 14
doubled**

Psychiatric
Research & Clinical
Practice Online



What harms have been documented to be caused by social media devices and techniques?

To avoid unfairly impacting small- and medium-sized enterprises, this bill only applies to social media platforms controlled by companies earning more than \$100 million in annual revenue. In addition, the bill excludes various types of digital platforms, including video games, email providers, search engines, streaming services, file-sharing and collaboration tools, internal business services, and limited-functionality sites like comment sections and product reviews.

The bill will cover many of the most popular social media platforms used by children, including TikTok, Instagram, and Facebook.

An unprecedented spike in teen girl suicide.

Two charts tell the story:

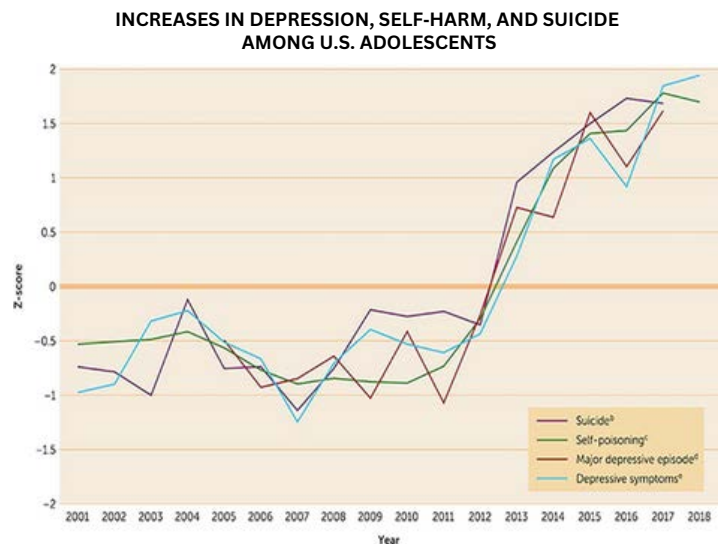
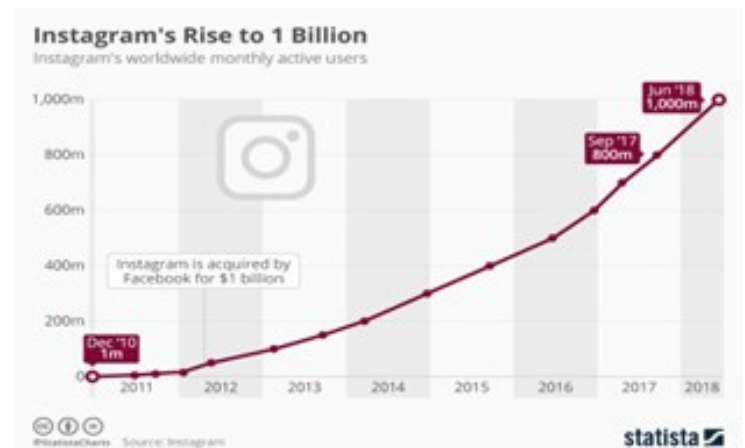


FIGURE 1. Indicators of poor mental health among U.S. girls and young women, 2001–2018 (note, before COVID) [15]

This never-before-seen spike in suicides among teen girls has occurred during this exact same time frame as the following:



To deny the never-before-seen child and teen mental health crisis we are currently engulfed in is at least partly caused by social media bizarrely requires denying the research conducted by the world's largest social media company about its very own operations. From Facebook's own leaked research: "Among teen users [of Instagram] who reported suicidal thoughts ... 6% of American [teen] users traced the desire to kill themselves to Instagram." [16] A March 2020 presentation posted by Facebook researchers to Facebook's internal message board reported that "66% of teen girls on IG experience negative social comparison (compared to 40% of teen boys)" and that "[a]spects of Instagram exacerbate each other to create a perfect storm."

Research outside of Facebook's own affirms the cause-and-effect relationship between these charts. Excessive use of digital and social media has a documented relationship to increases in suicide-related outcomes in teens and children, such as suicidal ideation, plans, and attempts. [17]

Fifty-nine percent of U.S. teens have reported being bullied on social media, and 25% of 9- to 17-year-olds report having had an online sexually explicit interaction with someone they believed to be an adult. [18]

Several studies have found correlations between time spent on social media and rates of suicide and depression among teens, especially girls. [19] In reviewing studies about social media use and girls:

Most lump all screen-based activities together (including those that are harmless, such as watching movies or texting with friends), and most lump boys and girls together. Such studies cannot be used to evaluate the more specific hypothesis that Instagram is harmful to girls. It's like trying to prove that Saturn has rings when all you have is a dozen blurry photos of the entire night sky.

But as the resolution of the pictures increases, the rings appear. The subset of studies that allow researchers to isolate social media, and Instagram in particular, show a much stronger relationship with poor mental health. [20]

Meanwhile, *reducing* social media use has been shown to result in mental health benefits. [21]

Consider these findings from a 2021 U.S. Surgeon General Advisory:

- From 2009 to 2019, the proportion of high school students reporting persistent feelings of sadness or hopelessness increased by 40%.
- The share seriously considering attempting suicide increased by 36%.
- The share creating a suicide plan increased by 44%.
- Between 2011 and 2015, youth psychiatric visits to emergency departments for depression, anxiety, and behavioral challenges increased by 28%.
- Between 2007 and 2018, suicide rates among youth ages 10-24 in the US increased by 57%. [22]

During the same period, the rates of suicide among 12- to 16-year-olds in the United States by use of a firearm increased 146%. [23]

In explaining the crisis' origins, the Surgeon General noted a "growing concern about the impact of digital technologies, particularly social media, on the mental health and wellbeing of children and young people" and called for greater accountability from social media companies. [24] "Business models are often built around maximizing user engagement as opposed to safeguarding users' health and ensuring that users engage with one another in safe and healthy ways. **This translates to technology companies focusing on maximizing time spent, not time well spent.**" [25]

A horrifying spike in teen girl eating disorders.

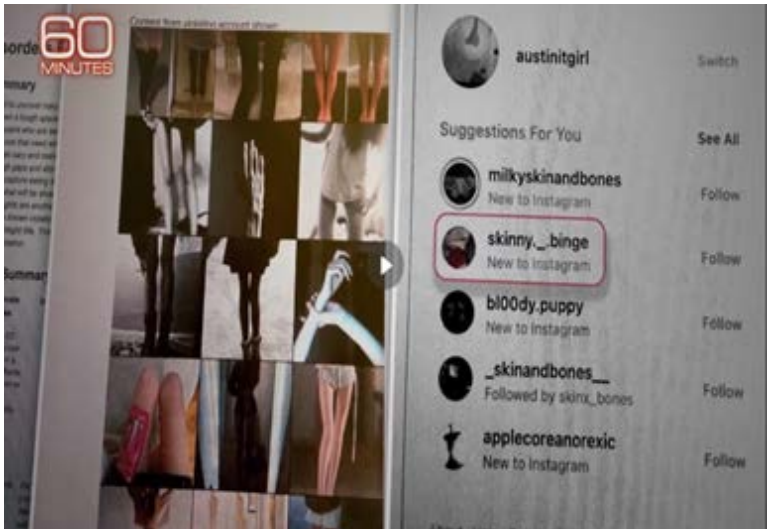
As Facebook's own leaked research admitted about Instagram, "We make body image issues worse for one in three teen girls." The research goes on: Teens blame Instagram for increases in the rate of anxiety and depression ... This reaction was unprompted and consistent across all groups."

Indeed, "Facebook knew Instagram was pushing girls to dangerous content: internal document In 2021, according to the document, an Instagram employee ran an internal investigation on eating disorders by opening a false account as a 13-year-old girl looking for diet tips. She was led to graphic content and recommendations to follow accounts titled 'skinny binge' and 'apple core anorexic.'" [26]

Instagram knowingly worsened existing body image issues for 1 in 3 teenage girls

The New York Times

Just a glance at the content pushed to girls, including those who do not search for it, underscores the urgency of legislative action:



Recently, Center for Countering Digital Hate researchers set up new accounts in the United States, United Kingdom, Canada, and Australia at the minimum age TikTok allows 13 years old. “These accounts paused briefly on videos about body image and mental health, and liked them. What we found was deeply disturbing. Within 2.6 minutes, TikTok recommended suicide content. Within 8 minutes, TikTok served content related to eating disorders. Every 39 seconds, TikTok recommended videos about body image and mental health to teens.”[27]

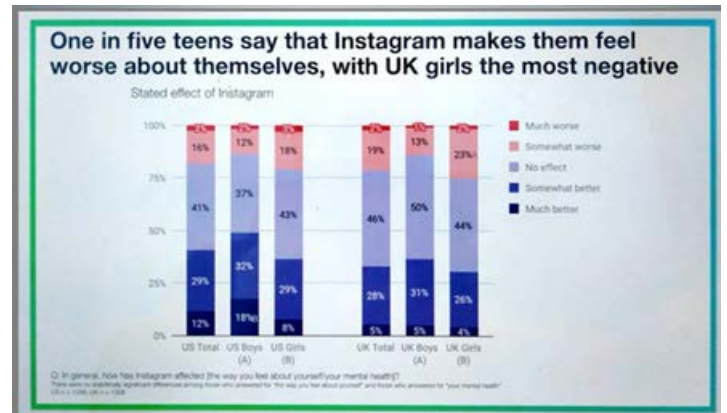
As one expert observed, “Instagram perpetuates the myth that our happiness and ability to be loved are dependent on external things: For girls, it’s appearance[.]”[28] The picture-perfect images on Instagram’s news feeds are so potent that they cement these superficial and harmful values into adolescent brains without them even knowing it.”[29]

Indeed, “observational studies have linked spending more than 2 hours a day on social networking sites and personal electronic devices with high rates of suicidality and depressive symptoms among adolescent girls[.]”[30]

As summed up by Senator Marsha Blackburn (R-TN) at the whistleblower hearing, using Facebook’s own leaked data (emphases supplied):

- Here are some of the numbers that Facebook chose not to share. ... And here you go: 66% of teen girls on Instagram and 40% of teen boys experience negative social comparisons, this is Facebook’s research.
- 52% of teen girls who experience negative social comparison on Instagram said it was caused by images related to beauty.
- Social comparison is worse on Instagram because it is perceived as real life, but based on celebrity standards.
- Social comparison mimics the grief cycle and includes a downward emotional spiral, encompassing a range of emotions from jealousy to self-proclaimed body dysmorphia.[31]

See this slide from internal Facebook Instagram research:[32]



Social media is the source of teen fentanyl deaths.

Fentanyl was the cause of 77.14% of drug deaths among teenagers last year.[33]



The unprecedented spike of children dying from overdosing on fentanyl has been documented to be the fault of social media. According to, for example, *The New York Times* article titled “Fentanyl Tainted Pills Bought on Social Media Cause Youth Drug Deaths to Soar — Teenagers and young adults are turning to Snapchat, TikTok and other social media apps to find Percocet, Xanax and other pills. The vast majority are laced with deadly doses of fentanyl, police say:”

the content pushed to girls, including those who do not search for it, underscores the urgency of legislative action:

- “Law enforcement authorities say an alarming portion of [fentanyl overdoses] unfolded ... from counterfeit pills tainted with fentanyl that teenagers and young adults bought over social media.”
- ““Social media is almost exclusively the way they get the pills,” said Morgan Gire, district attorney for Placer County, Calif., where 40 people died from fentanyl poisoning last year.
- “The phenomenon has led to disturbing new statistics ...:
- “Overdoses are now the leading cause of preventable death among people ages 18 to 45, ahead of suicide, traffic accidents and gun violence, according to federal data.
- Rates of illicit prescription pill use are now highest among people ages 18 to 25, according to federal data.

“There are drug sellers on every major social media platform” one expert is quoted as saying...: As long as your child is on one of those platforms, they’re going to have the potential to be exposed to drug sellers.”

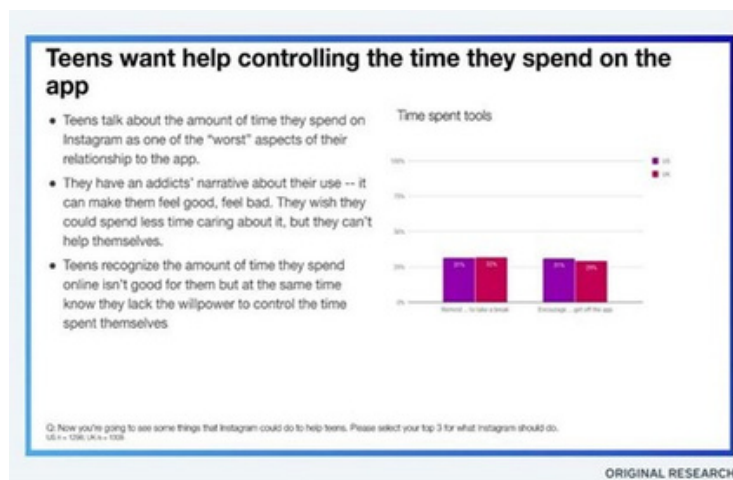
Making addicts of children.

Imagine being medically addicted to the platform that is making body image worse for one-third of our girls. “Adolescence is ...associated with an increased risk for... addictive disorders.”[34] Even so, platforms sometimes copy techniques used in gambling.[35] They recommend video after video without stopping cues, which are especially problematic for youth who may not yet have the self-discipline and maturity to take a break. (“The dopamine rush of endless short videos makes it hard for young viewers to switch their focus to slower-moving activities. ‘We’ve made kids live in a candy store.’”).[36] “TikTok is a dopamine machine,” said John Hutton, a pediatrician, and director of the Reading & Literacy Discovery Center at Cincinnati Children’s Hospital.[37]

These techniques cause addiction, in the medical sense of the word, among untold numbers of children. Facebook itself documented this in leaked research, as seen in the Appendix, which concludes that children they surveyed had “[a]n addict’s narrative about their use.” As Facebook whistleblower Frances Haugen explained:

Facebook has studied a pattern that they call problematic use, what we might more commonly call addiction. It has a very high bar for what it believes [“problematic use”] is. It [means] you self-identify that you don’t have control over your usage and that it is materially harming your health, your schoolwork or your physical health.

Indeed, Facebook’s internal definition of “problematic use” is identical to the medical definition of behavioral “addiction”: I know it is bad for me, it is objectively hurting me, I want to stop, but I can’t.[38] Here is Facebook’s chart documenting “an addict’s narrative”:



Others in addition to those employed by Facebook who have studied social media come to the same conclusion about child addiction with, for example, the United States Senate Republican Policy Center publishing a white paper titled “Social Media and Mental Health” where the Center described social media as “An Addiction Machine[.]” and stated that “[o]ne former Facebook executive, who quit the company and doesn’t allow his children to use social media, has said, ‘the short-term, dopamine-driven feedback loops that we have created are destroying how society works.’”[39]

Research using widely accepted measures of social media addiction suggests a prevalence rate of approximately five percent across the general population.[40] The rates of addiction for children may be higher: research shows that rates of behavioral addictions are elevated during adolescence.[41] With 22 million American teenagers using Instagram alone,[42] it is likely that more than a million young Americans are addicted to social media ... at least.

Just as tobacco companies employed scientists and doctors to maximize the addictive properties of the tobacco in their cigarettes, Facebook and other social media companies employ neuroscientists and user-interface designers to maximize the addictive properties of their technologies. Until recently, some design leaders boasted openly about such tactics in such industry bestsellers like *Hooked: How to Build Habit-Forming Products*[43] and *Evil by Design*.[44] One design consulting firm, called Dopamine Labs, brags that: “Dopamine makes your app addictive. Lift your engagement and revenue up to 167% by giving your users our perfect [hits] of dopamine. . . . High ROI. PhD tested, CEO approved.”[45]

Examples of addictive design patterns can be loosely grouped into three buckets.[46] First, using “reward schemes,” such as “likes,” takes advantage of our desire for validation. Second, the timing of “intermittent variable rewards,” such as loading icons or notifications, makes us feel rewarded for staying on the platform longer.[47] Third, the use of design patterns such as bottomless-scroll and auto-play, erode our natural stopping cues. The longer users engage with social media, the more data is collected about how they respond to these designs, enabling the enhancement and personalization of addictive design features.[48]

As bluntly explained in the BBC’s *Science Focus*:

And to do this, [the social media companies] build features into their apps that manipulate our brain chemistry. These tricks are borrowed straight from casinos and slot machines, which are widely considered to be some of the most addictive machines ever invented.

Consider the ‘pull-to-refresh’ feature common to social media apps, where dragging the screen downwards prompts the screen to refresh. Not only is the action itself similar to pulling the lever on a slot machine, but it takes advantage of our attraction to unpredictability. Psychologists call this ‘intermittent reinforcements’ (and I call it ‘the reason we date jerks’). Sometimes, when we check social media there’s something exciting waiting for us (a ‘reward’), sometimes there’s not. It’s the unpredictability that keeps us coming back.

Once you’re plopped in front of a slot machine, the casino wants to keep you there – that’s why most casinos are windowless

with no clocks. Similarly, once you’re trapped in a spiral of infinite scrolling, social media apps don’t want you to look up. That’s why their feeds are deliberately designed to be endless.[49]

Several former Facebook executives have acknowledged that design features they invented exploit weaknesses in human psychology. Leah Pearlman, co-inventor of Facebook’s “like” button, admitted that she herself had become hooked on Facebook because she had begun basing her sense of self-worth on the number of “likes” she had.[50] Sean Parker, Facebook’s first president, admitted that user interface designers “exploit[ed] a vulnerability in human psychology” by rewarding users with “a little dopamine hit” to ensure Facebook would “consume as much of [the users’] time and conscious attention as possible.”[51] The inventor of the infinite scroll at Mozilla has said that behind each phone home screen, there are teams of engineers working on increasing addictiveness.[52]



Are businesses that monitor and reform their behavior to prevent child harm treated differently by the bill?

Yes. This bill includes a safe harbor from all liability for social media platform companies that a) institute quarterly audits of their practices and platform features to detect any practices or features that have the potential to cause the listed harms to children, and b) remove or correct such practices or features within 30 days of completing the audit. Companies that can show in court that they complied with that procedure prior to engaging in the alleged violation of this bill would not be subject to any liability at all.



Won’t platforms fearing liability “over-correct” and deny welcome and appealing content to children?

Highly unlikely, for four reasons.

First, the \$100 million per year platforms would still be powerfully motivated to deliver appealing content to children. That would and still be their profit motivation. Their profit motivation will ensure that most questions will be resolved in favor of what is profitable, meaning “ties” will favor appealing and interesting content. Thus, the few platforms subject to this bill would likely curtail only the practices most obviously harmful to children; those most obviously exposing them to liability.

Second, the bill requires that a platform knew it was causing harm or departed from basic and ordinary care in its practices. Neither of these is easy to prove, and the burden is on the person claiming the law has been broken to do so. Plus, platforms can take cover in the bill's safe harbor. This means that platforms only have to behave with a minimum of responsibility entirely to avoid liability; that is, unless the platforms are behaving in ways worse than we suspect.

Third, the root cause of the harms to children is the immoral but profitable business decision to maximize their "user engagement" — the amount of time a child user spends on the platform — no matter the known harms caused to children as a result of this decision. Again, this strategy is not technologically inevitable. It is simply a business decision. When humans give artificial intelligence technologies the profit-driven instruction to maximize what platforms call "user engagement," AI does this at fantastic speeds by testing what recommendations work best across billions of users. Then they deliver content based on not just what a child says she wants but also what is statistically most likely to keep children like her glued to the screen. Too often, the answer is whatever exploits her fears and anxieties. This means that with disturbing frequency, depressed teens are offered suicide tips, body-image-anxious girls get content promoting eating disorders, and drug-curious youths get opportunities to buy pills laced with lethal fentanyl. Moreover, platforms use neuroscientifically tailored gimmicks such as auto-scrolling, constant reminders to return to the platform and dopamine-firing "likes" that can be addictive to children. Often, children who earnestly want to turn off the platform can't; their brains just aren't old enough to resist addictions to the same degree as adults.

As one very familiar and massively profitable example shows, it doesn't have to be this way. The Google search engine's method of prioritizing content served to viewers, designed to be based on websites' expertise, authoritativeness, and trustworthiness, shows there are profitable ways of deciding who sees what that are far less risky to children — and everyone else.

Fourth, to underscore how much it would take for the platforms to fear liability so much they would over-correct and fail to send nonharmful, appealing content to children to avoid liability, consider this: when the FTC fined Facebook a record \$5 billion in 2019, the platform's stock went up the next day.[53]



Isn't this bill federally preempted by Section 230 of the Communications Decency Act?

No, this bill would not be preempted by Section 230 because this bill holds companies liable for their own product's practices and features, *not for content that others post on their platforms*. Section 230 of the federal Communications Decency Act of 1996 immunizes social media companies, among others, from liability for content that others post on their platforms or services.[54]

This bill, on the other hand, expressly states that it shall not be construed as imposing liability for content created by persons or parties other than the social media platform.[55] Instead, this bill targets the harms that result from social media companies' practices and product features that cause children's addiction to the platform.

As the Solicitor General has recently written: "Under the principles articulated above, the Backpage.com decision was erroneous. The plaintiffs did not seek to treat Backpage.com as the publisher or speaker of the ads through which others carried out sex trafficking. Rather, they sought to hold the operator liable for its own policies and platform-design choices that facilitated sex trafficking. Where a website operator's conduct in furthering unlawful activities goes well beyond failing to block or remove objectionable third-party content from its platform, holding the operator liable does not 'treat' it 'as the publisher or speaker of' the third party posts." [56]

Recent court precedent clarifies that Section 230 does not immunize social media companies from harms that their own practices or product features cause. Last year, the Ninth Circuit ruled that Section 230 does not immunize social media platforms from liability for "unreasonable and negligent" design decisions, including product features that encourage users to take harmful actions.[57] "Section 230 does not protect companies when a claim is premised on harm that flows from the company's own speech or actions, independent from the speech of other users." [58]

Moreover, Section 230 states that it shall not "be construed to prevent any State from enforcing any State law that is consistent with this section." [59] This bill is consistent with Section 230, as interpreted by the Ninth Circuit, and would not be preempted by it.

Finally, the bill itself instructs courts not to interpret it in a way that would violate Section 230.



Who can bring a lawsuit under the bill?

Those harmed as well as public prosecutors.



Does the bill violate the First Amendment rights of social media platforms or children?

The answer is no for four reasons.

First, the Supreme Court has long distinguished between speech, which is protected by the First Amendment, and conduct, which is not. For instance, taking photographs is a form of protected speech, but entering onto someone else's property to take photographs is not speech; it is conduct.

Would *The New York Times* be able to argue it has a First Amendment right to instruct its delivery employees to insert rocks into its papers and throw Sunday editions at windows? Not every law involving "content" is a lawsuit about that content alone. For instance, laws proscribing subliminal advertising, workplace harassment, cyberbullying, all make actionable or unlawful the context of speech, not the speech itself. It is context and means of delivery that result in the categorical exclusion of true threats, fighting words, and incitement from First Amendment protection. Even Justice Holmes' classic example of unprotected speech — falsely shouting fire in a crowded theater (*Schenck v. United States*, 249 U.S. 47 (1919)) — relies on context, where and how the utterance is delivered, and whether harms are caused as a result. Thus, statutes such as the ones proposed by SB 287 differentiate between false shouts of "fire!" if someone is harmed and shouts of "fire!" that are transmitted to law enforcement. [60]

Put another way, if a platform is going to create a machine that uses vast databases and AI to send to users whatever will keep them online no matter what, using neuroscientifically designed addictive interfaces, and the platform knows or should know this machine hurts children, then requiring platforms to alter these operations to the extent needed to prevent harm is what we require of every business every day in the operations of their machines.

Second, courts have acknowledged that the First Amendment rights of adults cannot be used as a rationale for endangering children. In *New York v.*

Ferber, 458 U.S. 747 (1982), a unanimous Court held that the First Amendment does not protect child pornography. This is in part because there is a longstanding recognition that, while there is of course a foundational First Amendment right of freedom of expression, it also "is evident beyond the need for elaboration that a State's interest in 'safeguarding the physical and psychological well-being of a minor' is 'compelling.'" *Globe Newspaper Co. v. Superior Court*, 457 U.S. 596, 607 (1982). "A democratic society rests, for its continuance, upon the healthy, well-rounded growth of young people into full maturity as citizens." *Prince v. Massachusetts*, 321 U.S. 158, 168 (1944).

Third, a so-called "facial" First Amendment challenge to SB 287 before it is actually applied to any company in an actual lawsuit will run afoul of the doctrine that claims that a statute is too broad "do[] not apply to commercial speech." *Village of Hoffman Estates v. Flipside*, 455 U.S. 489, 490 (2008). Moreover, recent Supreme Court authority has signaled a retreat from embracing such challenges, expressing a preference to rule on such matters in so-called "as-applied" cases.[61]

Fourth, AI has no rights and it is AI machine learning algorithms that are making who-sees-what decisions. It is really challenging to describe in words just how *truly and autonomously intelligent* the AI machines used by social media platforms are, but here is one way of describing it:

Cicero, released last week [by Facebook], was able to trick humans into thinking it was real ... and can invite players to join alliances, craft invasion plans and negotiate peace deals when needed. The model's mastery of language surprised some scientists and its creators, who thought this level of sophistication was years away. But experts said its ability to withhold information, think multiple steps ahead of opponents and outsmart human competitors sparks broader concerns. ... "It's a great example of just how much we can fool other human beings," said Kentaro Toyama, a professor and artificial intelligence expert at the University of Michigan[.] "These things are super scary ... [and] could be used for evil."

So, if an AI developed by platforms is given the instruction "maximize user engagement!" it will do so by testing what recommendations work best and then re-writing and re-writing its own algorithms, at fantastic speeds, adjusting in real-time, based not a "neutral" criteria or one that is solely or even mostly determined by what the user thinks it wants.

It will serve up content that fulfills the goal of “maximizing engagement” no matter how foreseeable the harms to children ***precisely because AI does not foresee harm unless programmed to avoid it.*** If a lawsuit alleges that someone was harmed in whole or in part because of an AI-driven recommendation, some form of reasonable human intervention at some phase of the content-gathering and delivery-to-user process is required to minimize or eliminate these harms, if the company that issued the instruction is to satisfy its duty of ordinary care in negligence law. That’s what SB 287 would require – that human intervention mitigating the foreseeable harms of a company’s own business decision.

Endnotes

- [1] Quote on Page 1: Facebook Whistleblower Frances Haugen Testifies on Children and Social Media Use: Full Senate Hearing Transcript, <https://www.rev.com/blog/transcripts/facebook-whistleblower-frances-haugen-testifies-on-children-social-media-use-full-senate-hearing-transcript>.
- [2] Quote on Page 2: Whistleblower tells Congress that Facebook products harm kids and democracy. NPR (October 5, 2021), <https://www.npr.org/2021/10/05/1043207218/whistleblower-to-congress-facebook-products-harm-children-and-weaken-democracy>
- [3] Dave Davies, Users Beware: Apps are Using a Loophole in Privacy Law to Track Kids’ Phones, NPR (Jun. 16, 2021), <https://www.npr.org/2022/06/16/1105212701/users-beware-apps-are-using-a-loophole-in-privacy-law-to-track-kids-phones>.
- [4] Effects of Social Media on Children, Cleveland Clinic (Dec. 3, 2021), <https://health.clevelandclinic.org/dangers-of-social-media-for-youth/>.
- [5] Melinda Wenner Moyer, Kids as Young as 8 Are Using Social Media More Than Ever, Study Finds, N.Y. Times (Mar. 24, 2022), <https://www.nytimes.com/2022/03/24/well/family/child-social-media-use.html>.
- [6] Social Media and Adolescent Mental Health: Pediatric Mental Health Minute Series, American Academy of Pediatrics, <https://www.aap.org/en/patient-care/mental-health-minute/social-media-and-adolescent-mental-health/> (last visited Nov. 19, 2022).
- [7] All quotes and data from Greg McFarlane, How Facebook (Meta), Twitter, Social Media Make Money From You, Investopedia (Dec. 2, 2022), <https://tinyurl.com/yeyndyuj>.
- [8] Salvador Rodriguez, Facebook has had numerous scandals, So why does user engagement keep growing?, Milwaukee J. Sentinel (Jul 22, 2019), <https://tinyurl.com/y66mbvpt>.
- [9] Facebook’s research “suggests that no matter where we draw the lines for what is allowed, as a piece of content gets close to that line, people will engage with it more on average – even when they tell us afterwards they don’t like the content.” Mark Zuckerberg, A Blueprint for Content Governance and Enforcement, Facebook, Nov. 15, 2018, <https://www.facebook.com/notes/751449002072082/>

- [10] Facebook’s own research has proven that “people reporting problematic use report the site as more valuable to them, highlighting the complex relationship between [Facebook] use and well-being.” Justin Cheng et al., Understanding Perceptions of Problematic Facebook Use, in Chi Conference On Human Factors In Computing Systems Proceedings (CHI 2019), May 4-9, 2019, Glasgow, Scotland, <https://research.facebook.com/publications/understanding-perceptions-of-problematic-facebook-use/>.
- [11] In an internal Facebook memo leaked as part of the Facebook Files, Mark Zuckerberg told internal hate speech researchers that a proposed fix “wouldn’t launch if there was a material tradeoff with MSI impact” (MSI, or “meaningful social interactions,” is Facebook’s internal metric for user engagement). Elizabeth Dwoskin et al., The Case Against Mark Zuckerberg: Insiders Say Facebook’s CEO Chose Growth Over Safety, Wash. Post, Oct. 25, 2021, <https://www.adn.com/nation-world/2021/10/25/the-case-against-mark-zuckerberg-insiders-say-facebooks-ceo-chose-growth-over-safety/>.
- [12] Peter J. Adams, Addiction Industry Studies: Understanding How Proconsumption Influences Block Effective Interventions, American Journal of Public Health (2012).
- [13] Id.
- [14] Facebook Whistleblower, supra note 1.
- [15] <https://prcp.psychiatryonline.org/doi/full/10.1176/appi.prcp.20190015>.
- [16] <https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739>.
- [17] Elizabeth J. Ivie et al., A Meta-Analysis of the Association Between Adolescent Social Media Use and Depressive Symptoms, 275 J. of Affective Disorders 165, 165-174 (2020), <https://tinyurl.com/bdzu6h8h>, Alan Mozes, As Social Media Time Rises, So Does Teen Girls’ Suicide Risk, U.S. News (Feb. 16, 2021), <https://tinyurl.com/49hzmm9v>.
- [18] Monica Anderson, A Majority of Teens Have Experienced Some Form of Cyberbullying, Pew Rsch. Ctr. (Sept. 27, 2018); Responding to Online Threats: Minors’ Perspectives on Disclosing, Reporting, and Blocking, Thorn (May 2021).
- [19] Elia Abi-Jaoude et al., Smartphones, Social Media Use and Youth Mental Health, 192 Can. Med. Ass’n. J. E136 (Feb. 2020); RSPH, Instagram Ranked Worst for Young People’s Mental Health, <https://www.rsph.org.uk/about-us/news/instagram-ranked-worst-for-young-people-s-mental-health.html> (last visited Feb. 27, 2022).
- [20] <https://www.theatlantic.com/ideas/archive/2021/11/facebooks-dangerous-experiment-teen-girls/620767/>.
- [21] Roberto Mosquera et al., The Economic Effects of Facebook, 23 Exp. Econ. 575 (Jun. 2020). Melissa G. Hunt et al., No More FOMO: Limiting Social Media Decreases Loneliness and Depression, 37 J. Soc. Clinical Psych. 751 (Guilford Publications Inc. Nov. 2018). Hunt Allcott et al., The Welfare Effects of Social Media, 110 Am. Ec. Rev. 629 (Mar. 2020).
- [22] U.S. SURGEON GEN., ADVISORY: PROTECTING YOUTH MENTAL HEALTH 8 (2021).
- [23] Fatal Injury Reports, National, Regional and State, 1981– 2020, CTRS. FOR DISEASE CONTROL: WEB-BASED STAT. QUERY & REPORTING SYS., <https://wisqars.cdc.gov/fatal-reports> (last visited Feb. 22, 2023).
- [24] Id. at 25.
- [25] Id. (emphasis in original).
- [26] <https://www.cbsnews.com/news/facebook-instagram-dangerous-content-60-minutes-2022-12-11/>.
- [27] https://counterhate.com/wp-content/uploads/2022/12/CCDH-Deadly-by-Design_120922.pdf.
- [28] Jennifer Wallace, Instagram is Even Worse than We Thought for Kids. What Do We Do about It?, Washington Post, <https://www.washingtonpost.com/lifestyle/2021/09/17/instagram-teens-parent-advice/>.
- [29] Id.
- [30] Elia Abi-Jaoude, Karlina Treurnicht Naylor & Antonio Pignatiello, Smartphones, Social Media Use and Youth Mental Health 192 Canadian Med. Assoc. J. 136, 137 (2020).
- [31] Facebook Whistleblower, supra note 1.
- [32] Teen Mental Health Deep Dive. (2021, Sept. 29). The Wall Street Journal. <https://s.wsj.net/public/resources/documents/teen-mental-health-deep-dive.pdf>.
- [33] <https://www.latimes.com/california/story/2022-11-12/more-teenagers-are-dying-from-fentanyl>.
- [34] Christopher J. Hammond et al., Neurobiology of Adolescent Substance Use and Addictive Behaviors: Prevention and Treatment Implications, 25 ADOLESC. MED. STATE ART. REV. 15 (Apr. 2014), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4446977/>.
- [35] Daniel Kruger, Social Media Copies Gambling Methods ‘To Create Psychological Cravings’, Univ. of Mich. Inst. For Healthcare Pol’y & Innovation (May 8, 2018), <https://ihpi.umich.edu/news/social-media-copies-gambling-methods-create-psychological-cravings>.

Endnotes (cont.)

- [36] Julia Jargon, TikTok Brain Explained: Why Some Kids Seem Hooked on Social Video Feeds, Wall. St. J. (Apr. 2, 2022), <https://tinyurl.com/2p8xydhw>
- [37] Julia Jargon, TikTok Brain Explained: Why Some Kids Seem Hooked on Social Video Feeds, Wall. St. J. (Apr. 2, 2022), <https://tinyurl.com/2p8xydhw>.
- [38] Georgia Wells et al., Facebook Knows Instagram Is Toxic For Teen Girls, Company Documents Show, Wall St. J. (Sept. 14, 2021), <https://tinyurl.com/2p85fpud>. See also Cecilie Schou Andreassen et al., The Relationship Between Addictive Use of Social Media, Narcissism, and Self-Esteem: Findings From a Large National Survey, 64 Addictive Behav. 287, 289-92 (2017), <https://tinyurl.com/yy7v2hxx> (employing general diagnostic addiction criteria in finding relationships between basic demographic variables and personality traits and social media addiction).
- [39] Social Media and Mental Health, Senate Republican Pol'y Ctr. (Oct. 5, 2021), <https://tinyurl.com/3yn5vv57> (citing Amy Wang, Former Facebook VP Says Social Media Is Destroying Society With Dopamine-Driven Feedback Loops, Wash. Post (Dec. 12, 2017), <https://www.washingtonpost.com/news/the-switch/wp/2017/12/12/former-facebook-vp-says-social-media-is-destroying-society-with-dopamine-driven-feedback-loops/>).
- [40] <https://www.addictioncenter.com/drugs/social-media-addiction/>
- [41] Christopher J. Hammond et al., *supra* note 7.
- [42] Wells et al., *supra* note 2.
- [43] Nir Eyal & Ryan Hoover, Hooked: How To Build Habit-Forming Products (2014).
- [44] Chris Nodder, Evil By Design: Interaction Design To Lead Us Into Temptation (2013).
- [45] Jonathan Shieber, Meet the Tech Company that Wants to Make You Even More Addicted to Your Phone, TechCrunch, <https://social.techcrunch.com/2017/09/08/meet-the-tech-company-that-wants-to-make-you-even-more-addicted-to-your-phone/>.
- [46] Vikram Bhargava & Manuel Velasquez, Ethics of the Attention Economy: The Problem of Social Media Addiction, 31 Cambridge University Press 321 (Jul. 2021).
- [47] *Id.*
- [48] *Id.*
- [49] Catherine Price, Trapped – the secret ways social media is built to be addictive (and what you can do to fight back) (October 29, 2018), <https://www.sciencefocus.com/future-technology/trapped-the-secret-ways-social-media-is-built-to-be-addictive-and-what-you-can-do-to-fight-back/>.
- [50] Hilary Andersson, Social Media Apps “deliberately” Addictive to Users, BBC News (Jul. 3, 2018), <https://www.bbc.com/news/technology-44640959>.
- [51] Olivia Solon, Ex-Facebook President Sean Parker: Site Made to Exploit Human “Vulnerability,” The Guardian, <http://www.theguardian.com/technology/2017/nov/09/facebook-sean-parker-vulnerability-brain-psychology> (Nov. 9, 2017).
- [52] Andersson, *supra* note 32.
- [53] <https://www.theverge.com/2019/7/12/20692524/facebook-five-billion-ftc-fine-embarrassing-joke>.
- [54] 47 U.S.C. § 230(c)(1), 110 Stat. 137 (1996).
- [55] See proposed Civil Code section 1714.48(f) .
- [56] https://www.supremecourt.gov/DocketPDF/21/21-1333/249441/20221207203557042_21-1333tsacUnitedStates.pdf at p. 19.
- [57] Lemmon v. Snap, 995 F.3d 1085, 1094 (9th Cir. 2021) (“In short, Snap is being sued for the predictable consequences of designing Snapchat in such a way that it allegedly encourages dangerous behavior. [Section 230] does not shield Snap from liability for such claims.”).
- [58] Sophia Cope, Lawsuit Against Snapchat Rightfully Goes Forward Based on “Speed Filter,” Not User Speech, Electronic Frontier Foundation, <https://www EFF.org/deeplinks/2021/05/lawsuit-against-snapchat-rightfully-goes-forward-based-speed-filter-not-user> (May 18, 2021).
- [59] 47 U.S.C. § 230(e)(3).
- [60] See, e.g., Col. Rev. State. § 18-8-111. Waller v. Osborne, 763 F. Supp. 1144, 1148 (M.D. Ga. 1991) (subliminal speech would be “akin to false and misleading commercial advertising” and “worthy of little, if any, first amendment constitutional protection.”)
- [61] See, e.g., Wash. State Grange v. Wash. State Republican Party, 552 U.S. 442 (2008), where the Court noted that facial challenges “often rest on speculation,” and it asserted that invalidating a statute before it takes effect could “short circuit the democratic process.”

CONTACT CAI

Ed Howard, *Senior Counsel, Senior Policy Advocate*

eh4@sbcglobal.net

Jessica Heldman, *Fellmeth-Peterson Associate Professor in Child Rights*

jheldman@sandiego.edu

Robert Fellmeth, *Executive Director, Children's Advocacy Institute*

cpil@sandiego.edu

SCAN HERE