

Date of Hearing: July 11, 2023

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION

Jesse Gabriel, Chair

SB 680 (Skinner) – As Amended June 12, 2023

**SENATE VOTE:** Not relevant

**SUBJECT:** Features that harm child users: civil penalty

**SYNOPSIS**

*This bill is intended to protect child users, defined as those under the age of 16, from large social media platforms' deployment of designs, algorithms, or features that a platform knows or reasonably should have known causes a child user to inflict harm on themselves or others, develop an eating disorder, or experience addiction to the social media platform. It seeks to achieve this goal in two ways. First, and most importantly, it would provide social media companies a liability shield if they institute and maintain a program of at least quarterly audits of designs, algorithms, and features that have the potential to cause child users to suffer the enumerated harms, and proceed to correct these designs, algorithms, or features.*

*Second, the bill would allow public prosecutors to seek significant financial penalties and injunctive relief against social media companies that knowingly and willfully use such features—if they cause a child user to, in fact, inflict harm on themselves or others, develop an eating disorder, or experience addiction to the platform. Requiring platforms to be responsible for the harms they intentionally cause is one of the policies consistently recommended by experts in human centered technology design. This bill seeks to implement that recommendation on the state level.*

*The bill has been introduced in response to widespread documentation, including by the United States Surgeon General, of the harms that social media can present to child users, as well as Facebook whistleblower Frances Haugen's testimony and documentation that the company knew of the harms its platforms and algorithms were causing young users—yet persisted in using them.*

*This bill is sponsored by the Office of Attorney General Rob Bonta. It is supported by California District Attorneys Association, California Federation of Teachers, NextGen California, and a number of nonprofit organizations concerned with health and children's welfare. It is opposed by the California Chamber of Commerce, TechNet, various other technology trade associations, and by privacy groups, including Electronic Frontier Foundation.*

*The bill was previously heard by the Assembly Judiciary Committee, where it passed on an 8-0-3 vote.*

**SUMMARY:** Prohibits a social media platform from using a design, algorithm, or feature that the platform knows or reasonably should have known causes a child user to inflict harm on themselves or others, develop an eating disorder, or experience addiction to the social media platform. Specifically, **this bill:**

1) Defines the following terms:

- a) “Addiction” means a use of one or more social media platforms that does both of the following:
    - i) Indicates preoccupation or obsession with, or withdrawal or difficulty to cease or reduce use of, a social media platform; and
    - ii) Causes physical, mental, emotional, developmental, or material harms to the user.
  - b) “Audit” means a good faith, written, systemic review or appraisal by a social media platform that provides reasonable assurance of monitoring compliance with this bill and that meets both of the following criteria:
    - i) The review or appraisal describes and analyzes each of the social media platform’s current and forthcoming designs, algorithms, and features with the potential to cause a violation of 2) below.
    - ii) The review or appraisal includes any plans to change designs, algorithms, and features that pose more than a minimal risk of a violation of 2) below.
  - c) “Child user” means a person who uses a social media platform and is younger than 16 years of age.
  - d) “Eating disorder” means a behavioral condition characterized by a severe and persistent disturbance in eating behaviors and associated distressing thoughts and emotions, including anorexia nervosa, bulimia nervosa, and avoidant restrictive food intake disorder.
  - e) “Harm” means physical, mental, or emotional harm.
- 2) Prohibits a social media platform from using a design, algorithm, or feature that the platform knows, or by the exercise of reasonable care should have known, causes a child user to do any of the following:
    - a) Inflict harm on themselves or others.
    - b) Develop an eating disorder.
    - c) Experience addiction to the social media platform.
  - 3) Specifies that a social media platform shall be deemed to have caused a child user to inflict harm on themselves or others if, as a result of a design, algorithm, or feature, the platform sends a child user any of the specified information:
    - a) Information about how to obtain a firearm in violation of the portion of the Penal Code that regulates control of deadly weapons, and the child user obtains or uses the firearm.
    - b) Information regarding how to obtain a controlled substance, and the child user obtains or uses the controlled substance.
    - c) Information regarding how to die by suicide, and the child user acts on this information or becomes suicidal.

- 4) Makes a social media platform that knowingly and willfully violates 2) subject to a civil penalty of up to \$250,000 per violation, an injunction, and an award of litigation costs and attorney's fees. Establishes a four year statute of limitations.
- 5) Authorizes the following entities to bring a claim under this bill:
  - a) The Attorney General.
  - b) A district attorney.
  - c) A city attorney of a city having a population in excess of 750,000.
  - d) A county counsel of any county within which a city has a population in excess of 750,000.
  - e) With the consent of the district attorney, a city prosecutor in a city that has a full-time city prosecutor.
- 6) Exempts a social media platform from liability under this bill if it demonstrates it did both of the following:
  - a) Instituted and maintained a program of at least quarterly audits of its designs, algorithms, and features that have the potential to cause a violation of 2).
  - b) Corrected, within 30 days of the completion of an audit, any design, algorithm, or feature discovered to present more than a minimal risk of a violation.
- 7) Clarifies that a social media is not liable for any of the following:
  - a) Conduct protected under Section 230 of the Communications Decency Act.
  - b) Conduct protected under the First Amendment or under Article I, Section 2 of the California Constitution.
- 8) Excludes from the bill a social media platform controlled by a business entity that generated less than \$100,000,000 in gross revenue during the preceding year.
- 9) Provides that the bill does not negate or limit a cause of action under common law or any other statute, including any cause of action that may have existed or exists against a social media platform under the law as it existed before January 1, 2024.
- 10) Declares any waiver of the provisions of the bill to be contrary to public policy, void, and unenforceable.
- 11) Includes a severability clause.

**EXISTING FEDERAL LAW:**

- 1) Establishes, under Section 230 of the Communications Decency Act, that no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider. (47 U.S.C. § 230(c)(1).)

- 2) Exempts from Section 230 protection violations of federal criminal law; intellectual property law; state law that is consistent with Section 230; communications privacy law; and sex trafficking law. (47 U.S.C. § 230(e).)
- 3) Establishes the Children’s Online Privacy Protection Act of 1998 (COPPA), which regulates the collection of personal information from a child under the age of 13. (15 U.S.C. §§ 6501-6506.)

#### **EXISTING STATE LAW:**

- 1) Defines “social media platform” as a public or semipublic internet-based service or application that has users in California and that meets both of the following criteria:
  - a) A substantial function of the service or application is to connect users in order to allow them to interact socially with each other within the service or application. (A service or application that provides email or direct messaging services does not meet this criterion based solely on that function.)
  - b) The service or application allows users to do all of the following:
    - i) Construct a public or semipublic profile for purposes of signing into and using the service or application.
    - ii) Populate a list of other users with whom an individual shares a social connection within the system.
    - iii) Create or post content viewable by other users, including, but not limited to, on message boards, in chat rooms, or through a landing page or main feed that presents the user with content generated by other users. (Bus. & Prof. Code § 22675(e).)
- 2) Defines “social media company” as a person or entity that owns or operates one or more social media platforms. (Bus. & Prof. Code § 22675(d).)
- 3) Establishes the California Consumer Privacy Act (CCPA). (Civ. Code §§ 1798.100-1798.199.100.)
- 4) Prohibits a business from selling or sharing the personal information of consumers if the business has actual knowledge that the consumer is less than 16 years of age, unless the consumer, in the case of consumers at least 13 years of age and less than 16 years of age, or the consumer’s parent or guardian, in the case of consumers who are less than 13 years of age, has affirmatively authorized the sale or sharing of the consumer’s personal information. A business that willfully disregards the consumer’s age is deemed to have had actual knowledge of the consumer’s age. (Civ. Code § 1798.120.)
- 5) Establishes the California Age-Appropriate Design Code Act, which places a series of obligations and restrictions on businesses that provide online services, products, or features likely to be accessed by children. (Civ. Code §§ 1798.99.28-1798.99.40.)
- 6) Establishes the “substantial factor” test as the general causation standard under California tort law. (*Rutherford v. Owens-Illinois, Inc.* (1997) 16 Cal. 4th 953, 968-9.)

**FISCAL EFFECT:** As currently in print this bill is keyed fiscal.

**COMMENTS:**

1) **Background.** This bill, sponsored by Attorney General Rob Bonta, seeks to ensure that social media companies regularly audit and update their platforms to ensure that they do not cause children to harm themselves or others, develop eating disorders, or experience addiction to these platforms. By so doing, the bill would help ensure that large and profitable social media companies consider the substantial harm their platforms are causing children and weigh those considerations against their profits.

2) **Author's statement.** According to the author:

It is well documented that features built into social media platforms are designed to maximize the time users spend on the platform and direct users to specific content independent of whether the user requested or searched for that content. Children are particularly vulnerable to the addictive design features of these platforms with the result that kids and teens stay on these products for far longer than is healthy. Additionally, content directed to children has been found to promote dangerous and harmful practices and products that may cause the child user to inflict harm on themselves or others. Examples of this content direction include research that documented teen girls on the platform TikTok were recommended suicide content within 8 minutes, and directed to eating disorders content every 39 seconds. [...] SB 680 will enable the Attorney General and other prosecutors to hold social media platforms accountable for knowingly or carelessly addicting children to their platforms...or assisting children in inflicting harm on themselves or others.

3) **What this bill does.** This bill would authorize public prosecutors to bring civil claims against large social media platforms (those that generated \$100 million or more in gross revenue during the prior calendar year) that use a design, algorithm, or feature that the platform knew or should have known causes a user under the age of 16 to do *any* of the following:

1. Inflict harm on themselves or others.
2. Develop an eating disorder.
3. Experience addiction to the social media platform.

The definitions of the terms “harm,” “addiction,” and “eating disorder” are set forth in paragraph 1) of the **SUMMARY** above.

The bill also implements a safe harbor provision to allow compliant platforms to avoid liability. In order to benefit from the safe harbor, platforms must do *both* of the following:

1. Institute a program of quarterly audits of their designs, algorithms, and features that have the potential to cause a violation.
2. Correct a design, algorithm, or feature discovered by the audit that presents more than a minimal risk of violation within 30 days of completing the audit.

Finally, the bill would authorize a court to award a civil penalty of up to \$250,000 per violation, grant injunctive relief, and/or award attorney's costs and fees for a knowing and willful violation of the bill's prohibitions.

4) **The infinite feed and the quest for “likes.”** There are two key technological inventions that fundamentally altered the way people engage with social media: the infinite scroll and the “like” button.

In 2006, engineer Aza Raskin, developed the infinite scroll. At the time, he viewed it as a way for people on the internet to more easily use search engines or platforms like Yelp. The idea behind the infinite scroll was that users would no longer be limited to only 10 items per page while viewing search results. Rather than having to click through to the next page, they would be able to simply scroll through all of the results. While the infinite scroll was created for one type of internet use, it was quickly adapted for another, the social media infinite feed. Raskin, now the co-founder of the Center for Humane Technology, has publicly stated that this creation is one of his greatest regrets. In hindsight, Raskin notes that the infinite scroll has been used for mindlessness and for “weaponizing mindlessness at scale” as social media companies, like Facebook, Twitter, and TikTok, compete to get users’ attention and keep them on their sites longer. According to Raskin, people spend more time on these social media platforms because they are infinitely scrolling than if they did not have that technology. It equates to approximately 200,000 human lifetimes a day that are being wasted, based on his calculations. (Center for Humane Technology, *The Three Rules of Humane Tech*, Your Undivided Attention Podcast (Apr. 6, 2023), transcript available at [https://assets-global.website-files.com/5f0e1294f002b1bb26e1f304/642f41e223acbcf73d0eadf6\\_spotlight-your-undivided-attention-three-rules-of-humane-tech-transcript.pdf](https://assets-global.website-files.com/5f0e1294f002b1bb26e1f304/642f41e223acbcf73d0eadf6_spotlight-your-undivided-attention-three-rules-of-humane-tech-transcript.pdf) )

The second, and perhaps more profound, key innovation in social media that keeps billions of people glued to their feeds every day is the introduction of the “like” button on Facebook in 2009. Leah Pearlman, a former Facebook employee who is one of the people credited with the implementation of the feature, has said that it came about as a way of solving a redundancy issue where users would express their approval of a post with one word comments, which would then have the impact of causing more meaningful comments to get lost in the sea of one word approvals. Using a “like” button allowed users to easily express approval while reserving comments for fuller reactions and statements. It quickly became clear that beyond simply being an efficient way to express approval, the button fed people’s deep need to receive social validation. According to Pearlman, “The stats went up so fast—all the stats we thought would be affected, but 50 comments became 150 likes, almost immediately. Those people would start making more status updates, so there was way more content and it all just worked.” (Morgans, *The Inventor of the ‘Like’ Button Wants You to Stop Worrying About Likes*, Vice Magazine (Jul. 5, 2017) available at <https://www.vice.com/en/article/mbag3a/the-inventor-of-the-like-button-wants-you-to-stop-worrying-about-likes>.)

The “like” button on Facebook became the driving force in an increasingly potent and complex news-feed ranking algorithm: the more likes a post got, the more people Facebook would show it to. Likes became an asset for brands and advertisers, and a rich source of data for Facebook itself, telling the company about each user’s preferences and browsing habits. Additionally, likes became a powerful incentive for social media users. The like counter on each post became an explicit measure of its popularity, and an implicit measure of its value. Consciously or otherwise, users learned what sorts of posts would rack up likes and wrote more like that, while learning to

avoid those that garnered only a few. (Will Oremus, *How Facebook designed the like button—and made social media into a popularity contest*, Fast Company (Nov. 15, 2022) available at <https://www.fastcompany.com/90780140/the-inside-story-of-how-facebook-designed-the-like-button-and-made-social-media-into-a-popularity-contest>.)

Over time, people have come to use the number of likes that a post receives to judge others and themselves. Receiving more likes for a post than someone else receives for their post often provides a person with an increased sense of self-worth and superiority. Conversely, receiving too few likes can negatively impact someone's self-esteem, sometimes in very significant ways if the user relies only on social media likes to measure their worth. After posting a picture on any platform, users often crave the assurance it provides. Users become dependent on likes and positive comments as signs of validation rather than developing an internal sense of worth that is not dependent on the opinions of other social media users.

**5) Social media addiction.** Intentional product decisions by social media companies have resulted in products rife with features designed to engage and, in some cases, addict users through well-known psychological mechanisms. For example, social media sites often contain variable reward features, colloquially known as “slot machine” features. These features are predicated on the well-established psychological finding that the most effective way of maintaining a behavior is not with a consistent, predictable reward, but with unpredictable rewards that vary in their frequency or magnitude. Variable reward features on social media apps include “pull-to-refresh” affordances and unexpected notifications and animations encouraging a user to engage with new content.

Social media sites often build engagement and, in turn, addict users, through features that capitalize on human desires for social validation and social reciprocity. The encouragement to publicly “like” or favorite another user's content or message provides a sense of validation while also nudging the receiver of a “like” to “like” content as well, generating a mutually-reinforcing network of engagement. Snapchat's “snap streaks” feature capitalizes on the desire for social reciprocity by encouraging users to exchange content daily. The feature employs a system of emoji badges that indicate how many days the streak has lasted and when the streak is about to expire. Finally, social media products tend to addict users by omitting natural stopping cues from products. Nearly all social media products contain a near-infinite feed of content with no logical end. Moreover, the content in this feed is often only partly displayed on the screen, which is designed to encourage users to continue to scroll to see the content. (Bhargava and Velazquez, *Ethics of the Attention Economy: The Problem of Social Media Addiction*, Business Ethics Quarterly (July 2021) available at <https://www.cambridge.org/core/journals/business-ethics-quarterly/article/ethics-of-the-attention-economy-the-problem-of-social-media-addiction/1CC67609A12E9A912BB8A291FDFFE799/share/08cfe97de12fef45b5175836cfd00d3941a74b78>.)

The U.S. Surgeon General recently stated:

Social media platforms are often designed to maximize user engagement, which has the potential to encourage excessive use and behavioral dysregulation. Push notifications, autoplay, infinite scroll, quantifying and displaying popularity (i.e., ‘likes’), and algorithms that leverage user data to serve content recommendations are some examples of these features that maximize engagement. According to one recent model, nearly a third (31%) of social media use may be attributable to self-control challenges magnified by habit formation.

Further, some researchers believe that social media exposure can overstimulate the reward center in the brain and, when the stimulation becomes excessive, can trigger pathways comparable to addiction. Small studies have shown that people with frequent and problematic social media use can experience changes in brain structure similar to changes seen in individuals with substance use or gambling addictions. (*Social Media and Youth Mental Health: The U.S. Surgeon General's Advisory* (May 23, 2023) p. 9, available at <https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf>.)

Social media reinforces certain negative patterns by triggering the release of dopamine, a chemical in the brain that makes you feel good. The constant novelty and endless scrolling content create an addictive cycle of refreshing, much like someone in Las Vegas playing a slot machine. The unpredictability, coupled with social validation features, keeps you coming back. Platforms take advantage of our instinct to belong. Throughout human history, belonging to a group was necessary for our survival; our brains are wired to care what people think of us. When we receive social validation, the reward centers in our brain light up, triggering a release of more dopamine. Even the expectation of a positive reward releases more of that feel-good chemical which keeps us coming back during times of anxiety and stress. We become conditioned to seek regular and constant feedback with likes, comments, and shares. This creates a cycle of increased social media use and a decrease in mental health. (Center for Humane Technology, *Extractive Technology is Damaging our Attention and Mental Health*, available at <https://www.humanetech.com/attention-mental-health>.)

According to the Wall Street Journal's reporting on the internal documents leaked by a Facebook whistleblower, Frances Haugen, Facebook knew that the features built into the platform had the potential to addict users. Internal Facebook researchers found that 1 in 8 of its users report engaging in compulsive use of social media that impacts their sleep, work, parenting or relationships. These patterns, which Facebook refers to as "problematic use," mirror what is perhaps more appropriately known as internet addiction. (Wells, et al., *Is Facebook Bad for You? It Is for About 360 Million Users, Company Surveys Suggest*, Wall Street Journal (Nov. 5, 2021), available at <https://www.wsj.com/articles/facebook-bad-for-you-360-million-users-say-yes-company-documents-facebook-files-11636124681>.)

**6) Impact of social media on children.** According to the recent U.S. Surgeon General's advisory on the impact of social media on children's mental health, social media use by youth is nearly universal. Up to 95% of youth ages 13–17 report using a social media platform, with more than a third saying they use social media "almost constantly." Although age 13 is commonly the required minimum age used by social media platforms in the U.S., nearly 40% of children ages 8–12 use social media. As of 2021, the Surgeon General notes that 8th and 10th graders now spend an average of 3.5 hours per day on social media. (*Social Media and Youth Mental Health*, supra, p. 7.)

Writing in support of this bill, the California Academy of Family Physicians provides the following information:

The Centers for Disease Control and Prevention (CDC) reports that children in the U.S. ages eight to 10 spend an average of six hours per day in front of a screen; kids ages 11 to 14 spend an average of nine hours per day in front of a screen; and youth ages 15 to 18 spend an average of seven-and-a-half hours per day in front of a screen, with much of that time viewing social media content. Research has revealed that the intentional design of these



social media platforms—design that uses artificial intelligence to maximize “user engagement”—causes addiction to the platform, particularly for children.

In his advisory, the Surgeon General warns, “[T]he current body of evidence indicates that while social media may have benefits for some children and adolescents, there are ample indicators that social media can also have a profound risk of harm to the mental health and well-being of children and adolescents. At this time, we do not yet have enough evidence to determine if social media is sufficiently safe for children and adolescents.” (*Social Media and Youth Mental Health*, supra, p. 4.) The advisory goes on to describe why young people are at greater risk than adults:

Brain development is a critical factor to consider when assessing the risk for harm. Adolescents, ages 10 to 19, are undergoing a highly sensitive period of brain development. This is a period when risk-taking behaviors reach their peak, when well-being experiences the greatest fluctuations, and when mental health challenges such as depression typically emerge. Furthermore, in early adolescence, when identities and sense of self-worth are forming, brain development is especially susceptible to social pressures, peer opinions, and peer comparison. Frequent social media use may be associated with distinct changes in the developing brain in the amygdala (important for emotional learning and behavior) and the prefrontal cortex (important for impulse control, emotional regulation, and moderating social behavior), and could increase sensitivity to social rewards and punishments. As such, adolescents may experience heightened emotional sensitivity to the communicative and interactive nature of social media. Adolescent social media use is predictive of a subsequent decrease in life satisfaction for certain developmental stages including for girls 11–13 years old and boys 14–15 years old. Because adolescence is a vulnerable period of brain development, social media exposure during this period warrants additional scrutiny. (*Ibid.*, p. 5.)

Among the impacts of social media on mental health are increased neuroticism and anxiety, higher rates of depression, lower self-esteem, decreased attention spans, poor impulsivity, and brain patterns that look like attention-deficit hyperactivity disorder (ADHD). (Center for Humane Technology, *Extractive Technology is Damaging our Attention and Mental Health*, available at <https://www.humanetech.com/attention-mental-health>.)

The studies reviewed by the Surgeon General’s Office point to a higher risk of harm in adolescent girls and those already experiencing poor mental health. “For example, a study conducted among 14-year-olds found that greater social media use predicted poor sleep, online harassment, poor body image, low self-esteem, and higher depressive symptom scores with a larger association for girls than boys.” (*Social Media and Youth Mental Health*, supra, p. 7.)

What is more concerning are the issues that this bill seeks to address: the fact that social media companies have known for some time that social media use can be harmful to young users, and despite that knowledge, have continued to use algorithms and other design features to capture and hold their attention. In the fall of 2021, much of the country was enthralled by Frances Haugen’s testimony regarding Facebook’s internal research regarding the impact of social media on users, particularly young users. (Note: Facebook’s parent company renamed itself “Meta” shortly after Haugen’s testimony, an apparently unrelated event.) Haugen’s testimony claimed that social media companies are part of the reason for the child and teen mental health crisis—and that Facebook was aware of the correlation. Haugen, a former Facebook employee, testified, “Facebook chooses to mislead and misdirect. Facebook has not earned our blind faith.”

According to a Facebook study leaked by Haugen, more than 13% of teen girls in the U.K. reported their suicidal thoughts became more frequent after beginning to use Instagram; another leaked study found 17% of teen girls reported their eating disorders worsened after using the platform. (Allyn, *Here are 4 key points from the Facebook whistleblower's testimony on Capitol Hill*, National Public Radio (Oct. 5, 2021), available at <https://www.npr.org/2021/10/05/1043377310/facebook-whistleblower-frances-haugen-congress>.)

Haugen's testimony portrays the extent to which Facebook knew of its products' effects on the mental health of children, as well as the company's efforts to nonetheless recruit children to their platforms. According to Haugen, internal Facebook studies have confirmed anecdotal accounts that its amplification algorithms, such as engagement-based rankings on Instagram, "can lead children from very innocuous topics like health recipes [...] to anorexia promoting content over a very short period of time. [...] So Facebook knows that they are leading young users to anorexia content." (Haugen, *Facebook Whistleblower Frances Haugen Testifies on Children & Social Media Use: Full Senate Hearing Transcript* (Oct. 5, 2021), available at <https://www.rev.com/blog/transcripts/facebook-whistleblower-frances-haugen-testifies-on-children-social-media-use-full-senate-hearing-transcript>.)

Reports compiled within Facebook also detail the severe harm to body image visited upon teens and young adults, especially women, as a result of social comparison on these platforms. Young, impressionable users are particularly susceptible to the long-term mental and physical health impacts of this type of troublesome content. (Wells, et al, *The Facebook Files: Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show*, Wall Street Journal, (Sep. 14, 2021) available at <https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739>.) Haugen testified:

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 it is. It says 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. Five to [six percent] of 14 year olds have the self-awareness to admit both those questions. It is likely that far more than 5 to [six percent] of 14 year olds are addicted to Instagram. (Haugen, *Facebook Whistleblower*, supra.)

Haugen also referenced internal reports from Facebook indicating that Instagram makes "body image issues worse for one in three teen girls," and that "13% of British users and 6% of American users trace their desire to kill themselves to Instagram." According to Haugen, these negative effects are exacerbated by the inability of most parents to understand and address these unfamiliar problems, and the ensuing isolation that children experience as a result. (*Ibid.*)

These revelations underscored the culpability of some social media companies in propagating features detrimental to the wellbeing of youth. Studies corroborate that certain patterns of internet use can be harmful to the mental and physical wellbeing of children. Such problematic patterns of use are often encouraged by intentional design choices that maximize engagement with profit-motivated online services.

#### **6) Attention-extractive business models and the importance of requiring ethical design.**

Before the advent of social media and smart phones, technology, including the personal computer, was considered a tool: an object that is passive, awaiting a person to use it. As an

example, a blender is a tool that waits in a kitchen cabinet until it is needed. It does not take anything from the user. Something is not a tool, in this sense, if it demands information from the user and seduces and manipulates the user to extract information and divert the user's attention from other tasks. Many of the applications on smartphones are designed to capture and monopolize the user's attention through the use of notifications, the infinite feed, the constant need for social validation, and the ability to increase users' fears of missing out on information. As a result of this business model, the industry has moved away from a tools-based technology environment to an addiction-based environment. It has its own goals and pursues them relentlessly. That goal is largely a quest for the profits that come from the monetization of human attention and engagement by selling over a billion people's attention to advertisers, regardless of the cost to humanity.

Humane technologists, many of whom were formerly tech executives who developed and profited from the creation of Facebook, Google, Twitter, and Pinterest, to name a few, have sounded the alarm about the destructive nature of social media and are calling for a fundamental shift from the extraction model of technology development to a human centered, ethical model. (For more information on their concerns, see Netflix's 2020 documentary *The Social Dilemma*.) Aza Raskin and Tristan Harris, the co-founders of the Center for Humane Technology, for example, argue that the foundational operating models and incentives of current technologies need to change so that it is aligned with humanity's best interests, rather than prioritizing profits. Causing harm to individuals and society, in their view, is not a "cost of doing business; we do not need to accept the current, negative effects we are facing. Technologies like social media and artificial intelligence can and should increase our well-being, strengthen our democracies, and improve our shared information environment." (Center for Humane Technology, *Key Issues Overview*, available at <https://www.humanetech.com/key-issues#moving-forward>.)

Many humane technologists are asking government to step in to regulate social media, which is now operating in a largely lawless environment. Recommended policy changes include everything from restoring the same level of protections for kids and advertising, such as those once used to regulate advertising during Saturday morning cartoons, to banning behavioral advertising and microtargeting/hyper-personalization. These technologists also caution against approaches that involve industry self-regulation, content moderation, uncoordinated action, and laws that address individual problems without fixing their source. (Center for Human Technology, *The Dark Side of Social Media*, available at <https://www.humanetech.com/infographic-dark-side-social-media>.)

7) **Analysis.** The first question before this Committee is whether or not there is a need to protect consumers—in this case, children under the age of 16—from a potentially-harmful product: social media platforms. As discussed in detail above, there is an abundance of evidence demonstrating that social media can be harmful to children. The Surgeon General's report discusses the potential harms to children from social media use generally. The evidence from Facebook's internal documents suggests that the damage, at least for that platform, is knowing and willful. It would appear, then, to be sound policy to protect children from identifiable harms.

Recognizing the need to protect children, the question then becomes how to go about doing so. This bill proposes to do so in two ways. First, and most importantly, it would provide social media companies a liability shield if they (i) "institute[] and maintain[] a program of at least quarterly audits of...designs, algorithms, and features that have the potential to cause" child users to inflict harm on themselves or others, develop an eating disorder, or experience addiction

to the platform; and (ii) correct these designs, algorithms, or features. *This should be an uncontroversial undertaking.* There does not seem to be any sound policy reason for businesses to use designs, algorithms, and features that they know increase the risk of children under the age of 16 suffering these harms. The argument that social media platforms will simply bar younger users rather than be subject to this bill is a tacit admission that these platforms are unsafe for children to use; it is also perhaps the best argument that they ought to be redesigned with child safety in mind.

Second, the bill would allow public prosecutors to seek significant financial penalties and injunctive relief against social media companies that knowingly and willfully use such features—if they cause a child user to, in fact, inflict harm on themselves or others, develop an eating disorder, or experience addiction to the platform. Causation under this bill is a high bar. Under California tort law, it requires a showing that the defendant’s act was a substantial factor in causing the harm. (*Rutherford v. Owens-Illinois, Inc.* (1997) 16 Cal. 4th 953, 968-9.) The causation cannot be merely theoretical. (*Ibid.*) Requiring platforms to be responsible for the harms they intentionally cause is one of the policies consistently recommended by experts in human centered technology design. (See, e.g., Center for Human Technology, *Policy Reform Tool Kit: Accountability for Harms*, available at <https://www.humanetech.com/policy-reforms>.) This bill seeks to implement that recommendation on the state level.

Questions of whether or not the policies in this bill would violate the First Amendment or are preempted by federal law fall outside of this Committee’s jurisdiction, but are discussed in detail in in the Assembly Judiciary Committee’s analysis of the bill. The following are relevant excerpts from that analysis:

*First Amendment concerns.* The opposition, consisting of a coalition of tech industry organizations, raises numerous concerns including the potential that SB 680 runs afoul of the First Amendment . . . As a threshold issue, in order to fall under the scope of the First Amendment’s speech protections, a bill or law must limit or promote speech. Courts have yet to answer the question of whether an algorithm in particular would constitute speech. Moreover, it is plausible that when faced with the question, a court would find that rather than the speech itself, an algorithm is the mechanism by which speech is delivered to certain consumers. [On the other hand,] if algorithms are determined to be a tool for dissemination of speech, attempts such as this to regulate their use and development would not be subject to restrictions stemming from the First Amendment.

*Section 230 of the Communications Decency Act.* The federal Communications Decency Act (CDA) provides that “[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider,” and affords broad protection from civil liability for the good faith content moderation decisions of interactive computer services. (47 U.S.C. Sec. 230(c)(1) and (2).) Though Section 230 was originally passed in response to judicial inconsistency with respect to the liability of internet service providers under statutes pertaining to “publishers” of content created by others, it has since been interpreted to confer operators of social media platforms and other online services with broad immunity from liability for content posted on their platforms by others. [...]

The bill explicitly exempts from liability 1) any conduct protected by Section 230, and 2) any conduct protected by the First Amendment to the United States Constitution or Section 2 of

Article I of the California Constitution. These prohibitions seemingly attempt to foreclose the risk that the bill as a whole and claims brought under its provisions would violate Section 230 and the Constitution.

To summarize this extensive case law, courts have seemingly distinguished between a feature of a platform developed and implemented by the platform which seemed reasonably likely to cause harm (such as the SnapChat filter in *Lemmon* [v. *Snap, Inc.* (2021) 995 F.3d 1085]) and a facially neutral tool used to disseminate information (such as the algorithm in *Twitter* [v. *Taamneh* (2023) 598 U.S. \_\_\_, 143 S. Ct. 1206]). If challenged, the survival of this measure may turn on the proponent's ability to argue that imposing liability for the platform's designs, features, or affordances is more akin to the facts of *Lemmon* than those of *Twitter*. In the context of individual claims against platforms brought under this bill's provisions, plaintiffs would likely need to make a similar showing in order to avoid being dismissed under a Section 230 argument, which would necessitate a case by case determination of the facts.

8) **Committee amendments.** All of the proposed amendments set forth below are to provisions of the Civil Code that would be enacted by this bill.

**1714.48.** [...] (b) A social media platform is not in violation of this section if it demonstrates it did both of the following:

(1) The social media platform instituted and maintained a program of at least quarterly audits of its designs, algorithms, and features that have the potential to cause violations of subdivision (a).

(2) The social media platform corrected, within ~~30~~ **60** days of the completion of an audit described in paragraph (1), any design, algorithm, or feature discovered by the audit to present more than a de minimis risk of violating subdivision (a).

This amendment would provide social media platforms an additional 30 days to, after an audit, correct designs, algorithms, and features that place child users at a more-than-minimal risk of causing children to inflict harm on themselves or others, develop an eating disorder, or experience addiction to the platform. This additional time is arguably necessary given the complexity of rewriting and testing software.

**1714.48.** [...] (g) For purposes of this section: [...]

(7) (A) "Social media platform" has the same meaning as defined in Section 22675 of the Business and Professions Code.

***(B) "Social media platform" does not include a standalone direct messaging service that provides end-to-end encrypted communication or the portion of a multi-service platform that uses end-to-end encrypted communication.***

This amendment serves several purposes. As noted in the **SUMMARY** above, under this bill, a social media platform would be deemed to have caused child users harm if, as a result of its design, algorithm, or feature, the platform sends a child user any of the following:

1. Information about how to obtain a firearm in violation of the portion of the Penal Code that regulates control of deadly weapons, and the child user obtains or uses the firearm.
2. Information regarding how to obtain a controlled substance, and the child user obtains or uses the controlled substance.
3. Information regarding how to die by suicide, and the child user acts on this information or becomes suicidal.

It follows that, in order to minimize their potential liability, social media platforms must screen the content they send child users so as to avoid providing them such information. However, as noted in analyses of other bills heard this year by the Privacy Committee, such as AB 1394 (Wicks), “applications that provide 100% end-to-end encrypted direct messaging...play an important role in ensuring the privacy of sensitive communications, such as political organizing and obtaining information about reproductive healthcare.” It would subvert important policy goals of the state to require social media platforms to scan users’ end-to-end encrypted communications for harmful content. This amendment would solve the dilemma by exempting such communication services from the definition of “social media platform.” In other words, even if a child user were to obtain information about how to illegally purchase a gun through encrypted communications and then purchased the gun, the platform would not be liable. The importance of maintaining the privacy of these communications is too important, particularly given the assault on reproductive and other freedoms at the U.S. Supreme Court and in many states.

**1714.48.** [...] (i) This section does not negate or limit a cause of action under common law or any other statute, including any cause of action that may have existed or *will* exists against a social media platform under the law as ~~it existed~~ *enacted* before January 1, 2024.

This amendment is meant to ensure that the bill is not construed to interfere in any way with provisions of the Age Appropriate Design Code, enacted by AB 2273 (Wicks, Chap. 320, Stats. 2022) that take effect on July 1, 2024.

9) **Related legislation.** SB 287 (Skinner, 2023), a substantially similar bill, would prohibit a social media platform from using a design, algorithm, practice, affordance, or feature that the platform knows or should have known causes child users to experience specified harms, including receiving content that facilitates purchase of a controlled substance or the development of an eating disorder. Status: Held, Senate Floor.

AB 1282 (Lowenthal, 2023) would require the state Mental Health Services Oversight and Accountability Commission to report to the Legislature a statewide strategy to understand, communicate, and mitigate mental health risks associated with the use of social media by children and youth. Status: Senate Appropriations Committee.

AB 1394 (Wicks, 2023) would require social media platforms to provide a mechanism for users to report child sexual abuse material in which they are depicted, and provides platforms 30-60 days after receiving a report to verify the content of the material and block it from reappearing. The bill would also provide victims of commercial sexual exploitation the right to sue social media platforms for deploying features that were a substantial factor in causing their exploitation. Status: Senate Judiciary Committee.

AB 2273 (Wicks, Chap. 320, Stats. 2022) established the California Age-Appropriate Design Code.

AB 2408 (Cunningham, 2022) would have prohibited a social media platform from using a design, feature, or affordance that the platform knows, or should know by the exercise of reasonable care, causes a child user to become addicted to the platform. The bill was held in the Senate Appropriations Committee.

AB 1138 (Gallagher, 2019) would have prohibited a for-profit social media website or application from allowing a person under 16 years of age to create an account without first obtaining the consent of the person's parent or guardian. The bill was vetoed by Governor Newsom.

AB 2105 (Maienschein, Chap. 166, Stats. 2018) authorized treble damages in a civil action brought by, or on behalf of, or for the benefit of, a person who is a minor or nonminor dependent and is a victim of commercial sexual exploitation.

***ARGUMENTS IN SUPPORT.*** Writing in support of the bill, The Kennedy Forum states:

In light of the increase in teen suicides, teen fentanyl deaths, access to illegal weapons, and known social media addiction among children and youth, it can no longer be legal for social media companies to knowingly or carelessly cause harms to their child users. Thus, the Kennedy Forum wishes to express its SUPPORT of Senate Bill (SB) 680: Safeguarding Our Children. Child death and suffering caused by social platforms will continue so long as the platforms are lawfully able to invent and deploy profitable practices they know or should reasonably know harm our children. [ . . . ]

The parallel rise between the youth mental health crisis and social media usage is raising alarms everywhere, including at social media companies themselves. Facebook confirmed a "...significant, positive correlation between adolescent social media use and depressive symptoms." A 2020 analysis concluded that "Among teen users [of Instagram] who reported suicidal thoughts...6% of American [teen] users traced the desire to kill themselves to Instagram." Other research confirms that 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.

Additionally, social media platforms are knowingly causing to children extend beyond harming their general mental health. For example, social media platforms knowingly or carelessly have been documented to deluge body image anxious teen girls with pro-anorexia content and facilitate the sale of lethal fentanyl to children.

***ARGUMENTS IN OPPOSITION.*** A coalition of five organizations, including TechNet, the California Chamber of Commerce, and the Chamber of Progress, write in opposition to the bill:

[We] must respectfully oppose SB 680, which imposes such an extreme standard of liability on social media platforms for delivering user content that companies would have no choice but to limit or cease operations for kids under 16. Furthermore, SB 680 runs afoul of established first amendment principles and jurisprudence and is unambiguously preempted by federal law.



Our member companies prioritize the safety and privacy of teens that access their sites and platforms. We strongly believe children deserve a heightened level of security and privacy and there are a number of efforts within the industry to incorporate protective design features into their websites and platforms. Our companies have been at the forefront of raising the standard for teen safety and privacy across our industry by creating new features, settings, parental tools, and protections that are age-appropriate and tailored to the differing developmental needs of young people. This bill would upend those efforts in favor of an ill-conceived and likely unconstitutional private right of action designed to punish our companies rather than protect teen users.

Also writing in opposition, the Electronic Frontier Foundation notes:

We appreciate the ongoing conversation we have had with the supporters of this bill and the author's office. We recognize their genuine concern for children, and their openness to feedback. Unfortunately, we think this bill ultimately fails to accomplish its goals of protecting children and fundamentally runs afoul of both federal law and the First Amendment.

As we stated in our opposition last year to A.B. 2408, which bears several similarities to this bill, we believe S.B. 680—despite its intentions—would be preempted by a federal law protecting online speech, 47 U.S.C. § 230 (“Section 230”). It ignores that many aspects of a service are inseparable from the user-generated content that they publish. Under the logic of *Lemmon v. Snap, Inc.*, 995 F.3d 1085 (9th Cir. 2021), there is no support for the proposition that California—or any state—can avoid Section 230’s preemption by simply claiming that some or all aspects of an online service’s interactive features are defective or dangerous.

*Lemmon* held that Section 230 did not apply to a specific defective product claim based on specific allegations about a product created by a service. The Ninth Circuit’s narrow decision was the correct result in an early stage of a case, and its careful reasoning avoided undermining Section 230 in ways that would harm online speech.

## **REGISTERED SUPPORT / OPPOSITION:**

### **Support**

Attorney General Rob Bonta (sponsor)  
American Academy of Pediatrics, California  
American Association of University Women – California  
California Academy of Family Physicians  
California District Attorneys Association  
California Federation of Teachers AFL-CIO  
California Youth Empowerment Network  
Children’s Advocacy Institute  
Common Sense Media  
Consumer Federation of California  
Girl Scouts Heart of Central California  
#HalfTheStory  
Jewish Family and Children's Services of San Francisco, the Peninsula, Marin and Sonoma Counties



Mental Health America of California  
NextGen California  
Parents Television and Media Council  
Steinberg Institute  
The Kennedy Forum

**Opposition**

California Chamber of Commerce  
Chamber of Progress  
Civil Justice Association of California  
Computer and Communications Industry Association  
Electronic Frontier Foundation  
NetChoice  
Oakland Privacy  
TechNet

**Analysis Prepared by:** Jith Meganathan / P. & C.P. / (916) 319-2200, Julie Salley / P. & C.P. / (916) 319-2200