

Date of Hearing: April 1, 2025

Fiscal: No

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION

Rebecca Bauer-Kahan, Chair

AB 56 (Bauer-Kahan) – As Amended February 25, 2025

PROPOSED AMENDMENTS

SUBJECT: Social media: warning labels

SYNOPSIS

According to U.S. Surgeon General Vivek Murthy, “The mental health crisis among young people is an emergency – and social media has emerged as an important contributor. Adolescents who spend more than three hours a day on social media face double the risk of anxiety and depression symptoms, and the average daily use in this age group, as of the summer of 2023, was 4.8 hours. Additionally, nearly half of adolescents say social media makes them feel worse about their bodies.” As a result, he called for “a surgeon general’s warning label on social media platforms, stating that social media is associated with significant mental health harms for adolescents,” which “would regularly remind parents and adolescents that social media has not been proved safe.”¹

This bill, the Social Media Warning Law, answers that call. Sponsored by Attorney General Rob Bonta and Common Sense Media, the bill would require platforms to display the following message, based on the Surgeon General’s 2023 Advisory on Social Media and Youth Mental Health: “The Surgeon General has advised that there are ample indicators that social media can have a profound risk of harm to the mental health and well-being of children and adolescents.” The warning – which occupies at least 75 percent of the screen for 90 seconds – must be displayed each calendar day when the user first logs on to the platform, again after three hours of cumulative active use, and every hour thereafter, until the end of the calendar day.

The bill is supported by, among others, the Academy of Pediatrics, California, Children’s Advocacy Institute, California Initiative for Technology & Democracy (CITED), and Professor Jonathan Haidt, author of The Anxious Generation. It is opposed by a coalition of associations, including TechNet and California Chamber of Commerce, as well as Electronic Frontier Foundation. Oakland Privacy take an oppose-unless-amended position.

Amendments described in Comment #9 would reword the warning and acknowledge social media’s benefits for some users. The amendments would also allow for the Department of Public Health to update the warning label to remain consistent with the state of the science.

If passed by this Committee, this bill will next be heard by the Assembly Judiciary Committee.

THIS BILL:

¹ Dr. Vivek Murthy, “Surgeon General: Why I’m Calling for a Warning Label on Social Media Platforms” *New York Times* (Jun. 17, 2024), <https://www.nytimes.com/2024/06/17/opinion/social-media-health-warning.html>.

- 1) Makes findings and declarations regarding the association between social media use and adolescent mental health impacts.
- 2) Defines “social media platform” as an internet website or internet medium that:
 - a. Permits a person to become a registered user, establish an account, or create a profile for the purpose of allowing the user to create, share, and view user-generated content through that account or profile.
 - b. Enables on or more users to generate content that can be viewed by other users of the internet website or internet medium.
 - c. Primarily serves as a medium for users to interact with content generated by other users of the internet website or internet medium.
- 3) Requires, for each calendar day in which a user uses a social media platform, that the platform display a black box warning to the user when the user initially accesses the platform, again after three hours of cumulative active use, and thereafter at least once per hour of cumulative active use.
- 4) Requires the warning to be displayed clearly and continuously for at least 90 seconds, without providing the ability to bypass or click through the warning, in a manner that is clear and legible and that occupies between 75 and 100 percent of the screen or window.
- 5) Requires that the warning reads: “The Surgeon General has advised that there are ample indicators that social media can have a profound risk of harm to the mental health and well-being of children and adolescents.”
- 6) States that provision of notice under the bill does not waive, release, otherwise limit, or serve as a defense to any claim, including a warning, other than a claim premised on a violation of this section.

EXISTING LAW:

- 1) Under federal law, defines “social media platform” as an internet website or internet medium that:
 - a) Permits a person to become a registered user, establish an account, or create a profile for the purpose of allowing the user to create, share, and view user-generated content through that account or profile.
 - b) Enables on or more users to generate content that can be viewed by other users of the internet website or internet medium.
 - c) Primarily serves as a medium for users to interact with content generated by other users of the internet website or internet medium. (42 U.S.C. § 1862w.)
- 2) Under state law, defines “Social media platform” means 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 users to interact socially with each other within the service or application. A service or application that provides email or direct messaging services is considered to meet this criterion on the basis of that function alone.
- 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(f).)
- 3) Under the Protecting Our Kids from Social Media Addiction Act, prohibits an operator of an addictive internet-based service or application, including a social media platform, from providing an addictive feed, as defined, to a minor user, except as prescribed. (Health & Saf. Code § 27000 et seq.)

COMMENTS:

1) **Author's statement.** The author writes:

We are in the midst of a global mental health crisis. In the last decade or so, young people around the world, have experienced a dramatic spike rates of depression, anxiety, self-harm, and suicide. The crisis is real, urgent, and getting worse. Mounting evidence shows that online addiction in children is a widespread problem, with troubling implications for their mental health and well-being. More time on social media tends to be correlated with depression, anxiety, eating disorders, and interference with daily life, including establishing healthy sleep patterns. With AB 56 we're responding to the U.S. Surgeon General Vivek Murthy's call that all users of social media be warned about the risks associated with heavy social media use. This education is critical for both children and parents alike.

2) **The recent, international spike in youth mental health disorders.** The early 2010s saw a major upsurge in adolescent depression and anxiety, self-harm, and suicide. The trend is concentrated in Gen Z, and girls are more impacted than boys:

% U.S. Anxiety Prevalence

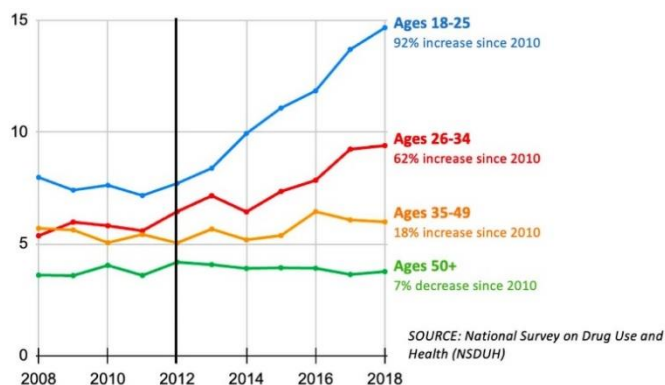


Figure 1. Percent US anxiety prevalence. National Survey on Drug Use and Health (NSDUH).²

As of 2021, relative rates of depression teen girls and boys had increased by roughly 150% compared to 2010:

% US Teens with Major Depression

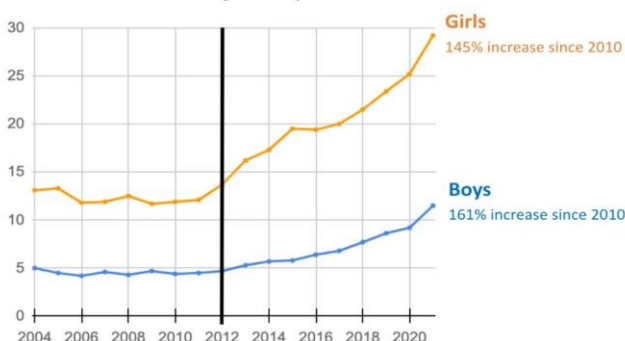


Figure 2. US teen depression prevalence. NSDUH data.³

The trend is reflected in objective measures, including hospitalizations from self-harm. In 2020, young teenage girls were hospitalized for self-harm, primarily from cutting, at three times the rate they were in 2010:

² Haidt, “The Teen Mental Illness Began Around 2012” *After Babel* (2023), <https://www.afterbabel.com/p/the-teen-mental-illness-epidemic>, summarizing Haidt et al, “Adolescent mood disorders since 2010: A collaborative review” (ongoing) available at https://docs.google.com/document/d/1diMvsMeRphUH7E6D1d_J7R6WbDdgnzFHDHPx9HXzR5o/edit?tab=t.0#.

³ “The Teen Mental Illness Began Around 2012,” *supra*.

US Teens Admitted to Hospitals for Nonfatal Self-harm (Ages 10-14)

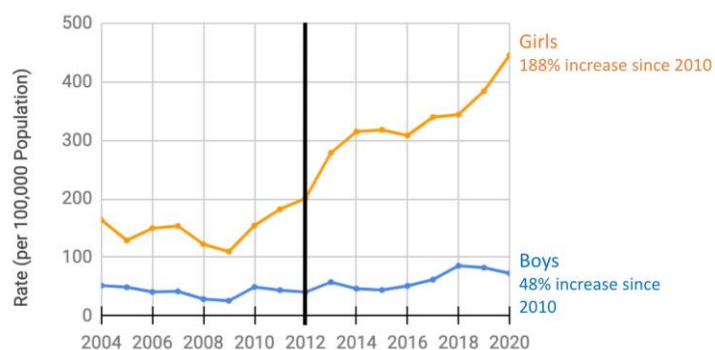


Figure 3. Hospital admissions for self-harm, ages 10-14. CDC data.⁴

Young teen suicide more than doubled in this timeframe:

US Teens, Suicides (Ages 10 – 14)

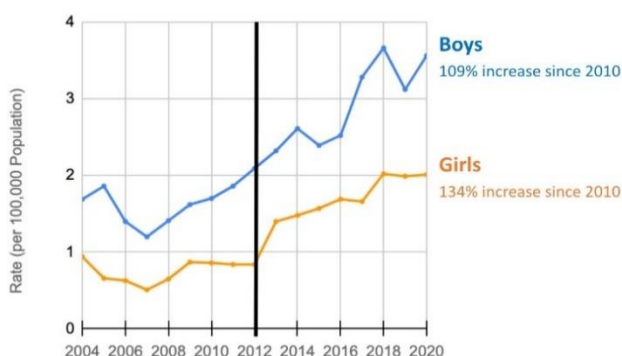


Figure 4. Suicide rate per 100,000 of US population, ages 10-14.⁵

Similar trends have been observed in several western counties.⁶ These trends track “the years when adolescents in rich countries traded their flip phones for smartphones and moved much more of their social lives online – particularly onto social-media platforms designed for virality and addiction.”⁷

3) Social media use is associated with significant mental health harms. In May 2023, Surgeon General Murthy issued an advisory warning of the potential mental health impacts of social media on young people. The advisory recognizes the benefits of social media for some users:

⁴ *Id.* For older teens, the increase for girls was 48%; for boys, 37%.

⁵ *Id.* For older teens, the increase for girls was 63.9%; for boys, 35%.

⁶ A series of articles from Haidt and Rausch addresses this issue under the header “The International Mental Health Crisis” on Haidt’s Substack, *After Babel*, <https://www.afterbabel.com/t/the-international-mental-health-crisis>.

⁷ Haidt, “End the Phone-Based Childhood Now” *The Atlantic* (March 13, 2024), <https://www.theatlantic.com/technology/archive/2024/03/teen-childhood-smartphone-use-mental-health-effects/677722/>.

Social media can provide benefits for some youth by providing positive community and connection with others who share identities, abilities, and interests. It can provide access to important information and create a space for self-expression. The ability to form and maintain friendships online and develop social connections are among the positive effects of social media use for youth. These relationships can afford opportunities to have positive interactions with more diverse peer groups than are available to them offline and can provide important social support to youth. The buffering effects against stress that online social support from peers may provide can be especially important for youth who are often marginalized, including racial, ethnic, and sexual and gender minorities.⁸

The advisory calls for more research and concludes that while “the 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.”⁹ A large and growing body of evidence, highlighted below, suggests the conclusion is well founded.

Evidence of correlation. A vast body of research, summarized in large literature reviews, establishes a correlation between social media use and poor mental health outcomes for youth.¹⁰ For example, a 2020 review of 13 studies synthesized evidence on the influence of social media on depression, anxiety, and psychological distress in a total of 21,231 adolescents in various countries across four domains of social media: time spent, activity, investment, and addiction. The authors concluded: “All domains correlated with depression, anxiety and psychological distress,” but noted there was complexity and need for further study.¹¹ In an open-source review of studies of social media and mental health, Haidt et al conclude that, although the research was not sufficiently clear just a few years ago, today nearly all researchers agree that there is a relationship between time spent using social media and mental health problems, although there is considerable disagreement as to the size and significance of this effect.¹²

Some researchers have found the relationship between social media and mental health negligible. Most prominently, Orben & Przybylski (2019) concluded the risk of social media use was comparable to “eating potatoes” or wearing eyeglasses.¹³ However, the study applied to all screen-based activities, including watching television¹⁴ – an activity that pre-dated the recent spike in youth mental health disorders. Professor Orben subsequently concluded that the

⁸ “Social Media and Youth Mental Health: The U.S. Surgeon General’s Advisory” (May 23, 2023) p. 6 (emphasis added), <https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf>. (“Surgeon General’s Advisory”)

⁹ *Id.* at p. 4.

¹⁰ Costello et. al, “Algorithms, Addiction, and Adolescent Mental Health: An Interdisciplinary Study to Inform State-level Policy Action to Protect Youth from the Dangers of Social Media” (2023) 49 Am. J. L. and Med. 135, fn. 60 (listing studies).

¹¹ Betul Keles et al., “A Systematic Review: The Influence of Social Media on Depression, Anxiety and Psychological Distress in Adolescents” (2020) 25 Int’l J. Adolescence & Youth 79, 84-86, <https://www.tandfonline.com/doi/full/10.1080/02673843.2019.1590851#d1e1197>.

¹² Haidt, et al., “Social media and mental health: A collaborative review” (Ongoing) Unpublished manuscript, New York University, p. 307, <https://docs.google.com/document/d/1w-HOfseF2wF9YIpXwUUtP65-olnkPyWcgF5BiAtBEy0/edit>. (“Social media and mental health.”)

¹³ Orben & Przybylski, “The association between adolescent well-being and digital technology use” *Nature* (2019), <https://www.nature.com/articles/s41562-018-0506-1>.

¹⁴ See response paper from Twenge et al, “Underestimating digital media harm,” *Nature* (2019), <https://www.nature.com/articles/s41562-020-0839-4.epdf>.

associations between social media use and a decline in well-being ranged from $r = 0.10$ to $r = 0.15$, where $r = 1.0$ indicates a perfect relationship and $r = 0$ indicates no relationship.¹⁵ For comparison, the correlation for lead and child IQ was $r = 0.11$, enough to justify a public health campaign.¹⁶

The correlation appears to be strongest for heavy users. While some light users – roughly an hour a day – may have better mental health outcomes than non-users,¹⁷ several studies indicate that heavy users – around three or more hours a day – experience a sharp increase in the risk of mental health outcomes, with girls more impacted than boys:

- A study of 6,595 adolescents found that those who spend more than three hours a day on social media face double the risk of experiencing symptoms of depression and anxiety.¹⁸
- A study of over 10,904 14-year-olds in the UK found that compared to those who used social media less than three hours a day, the increase in depressive symptoms for girls and boys, respectively, was as follows:
 - 3-5 hours of daily use: 26% and 21%
 - Over five hours of daily use: 50% and 35%.¹⁹
- A review of three large surveys of a total of 221,096 adolescents in two countries found that, compared to those who use social media less than one hour a day, those who used social media three hours a day were 28% more likely to report they were low in wellbeing, a figure that increased to 64% for those who used social media more than five hours a day.²⁰
- A meta-analysis of 26 studies with a total of 55,340 adolescents found that time spent on social media was significantly associated with a higher risk of depression symptoms, with a stronger association for girls.²¹ The results show a substantial uptick for both genders beginning at three hours of use:

¹⁵ *Ibid.*

¹⁶ The range was from $r = .10$ to $r = .15$, where $r = 1.0$ indicates a perfect relationship and $r = 0$ indicates no relationship. (See *ibid.*) The correlation for lead and child IQ is $r = .11$. (Reuben et al, “Association of Childhood Blood Lead Levels With Cognitive Function and Socioeconomic Status at Age 38 Years and With IQ Change and Socioeconomic Mobility Between Childhood and Adulthood” (2017), <https://jamanetwork.com/journals/jama/fullarticle/2613157>.)

¹⁷ *Id.* p. 311.

¹⁸ Riehm, K. E., et al., “Associations Between Time Spent Using Social Media and Internalizing and Externalizing Problems Among US Youth” *JAMA psychiatry*, 76 (12), 1266–1273 (2019), <https://pubmed.ncbi.nlm.nih.gov/31509167/>.

¹⁹ Kelly et al, “Social Media Use and Adolescent Mental Health: Findings From the UK Millennium Cohort Study” *EClinicalMedicine* (2019), <https://pmc.ncbi.nlm.nih.gov/articles/PMC6537508/>

²⁰ Twenge & Campbell “Digital media use is linked to lower psychological well-being: Evidence from Three Datasets,” *Psychiatric Quarterly* (Mar. 2019), Tables 1 & 2, <https://pubmed.ncbi.nlm.nih.gov/30859387/>.

²¹ Liu et al, “Time Spent on Social Media and Risk of Depression in Adolescents: A Dose-Response Meta-Analysis” *PubMed* (2022), <https://pubmed.ncbi.nlm.nih.gov/35564559/#:~:text=The%20risk%20of%20depression%20increased,of%20adolescent%20social%20media%20use.>

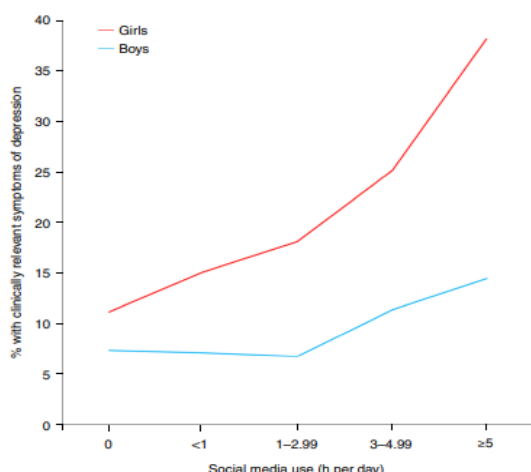


Figure 5: Percentage of clinically relevant depressive symptoms as a function of social media use.²²

As of 2024, the average daily social media usage for US adolescents was 4.8 hours.²³

Evidence of causation. Causal evidence for social media’s effect on adolescent mental health is difficult to ascertain due to the ethical and logistical limits on experimentation. As a result, some scholars, such as UC Irvine professor of psychology Candice Odgers, assert that the data is at most correlative and does not show causation.²⁴ By contrast, the Strategic Training Initiative for the Prevention of Eating Disorders, a research program based at the Harvard School of Public Health and Boston Children’s Hospital, recently concluded “rigorous experimental and longitudinal public health studies of social media effects strongly suggest social media has a harmful impact on the mental health of young users.”²⁵

Haidt et al list 24 experimental and quasi-experimental studies, 19 of which found significant evidence of harm.²⁶ For example, Facebook’s adoption on college campuses across the US was staggered, providing a natural experiment with control groups on campuses where Facebook had not yet been adopted. One study found that the roll-out of the platform was associated with a 9% increase in depression and a 12% increase in anxiety.²⁷ Conversely, a Canadian university performed a randomized study on undergraduates with symptoms of anxiety and depression and found that the group assigned to limited social media activity for one month had improvements

²² Twenge et al, “Underestimating digital media harm” *Nature* (2019), <https://www.nature.com/articles/s41562-020-0839-4.epdf>.

²³ “Surgeon General: Why I’m Calling for a Warning Label on Social Media Platforms,” *supra*.

²⁴ Odgers, “The great rewiring: is social media really behind an epidemic of teenage mental illness?” *Nature* (2024), <https://www.nature.com/articles/d41586-024-00902-2>.

²⁵ Costello et. al, *supra*, p. 142.

²⁶ “Social media and mental health,” *supra*, at pp. 181-186, 309.

²⁷ Braghieri et al, “Social Media and Mental Health” *American Economic Review* (2022), <https://www.aeaweb.org/articles?id=10.1257/aer.20211218>.

in subjective mental wellbeing.²⁸ Similarly, studies show mental health improvements when schools go phone free.²⁹

Longitudinal cohort studies that track mental health impacts on groups across time also suggest a causal link between time spent on social media and mental health problems. Professor Haidt discusses 33 such studies that looked at intervals longer than a month, 25 of which found evidence indicating causation.³⁰ One study of US high school students found that the use of image-related social media was associated with the emergence of depressive symptoms one year later.³¹ Another study found that adolescent girls who used social media two or three hours a day early in adolescence, and subsequently increased use over time, had increased suicide risk ten years later.³²

Internal research from social media companies. Internal studies from social media obtained from leaks or in litigation provide additional evidence for the connection between social media use and mental health impacts. For example, leaked research from Meta, Instagram’s parent company, revealed that Instagram made “body image issues worse for one in three teen girls.”³³ More alarmingly, the report also found that of users who reported having suicidal thoughts, 13% of UK and 6% of US teenagers suggested that they were the result of Instagram use.³⁴ Their internal research goes on to state that: “Teens blame Instagram for increases in the rate of anxiety and depression . . . This reaction was unprompted and consistent across all groups.”³⁵

Absence of compelling alternative explanations. Some scholars have argued that the apparent spike in youth mental health issues can be attributed to changing diagnostic criteria and the willingness of Gen Z to report distress as stigma around mental illness declined.³⁶ Others, such as Professor Odgers, have argued that a broader set of social ills, including “access to guns, exposure to violence, structural discrimination and racism, sexual abuse, the opioid epidemic, economic hardship, and social isolation,” better explain the rapid onset of anxiety and

²⁸ Davis & Goldfield, “Limiting social media use decreases depression, anxiety, and fear of missing out in youth with emotional distress: A randomized controlled trial,” *Psychology of Popular Media* (2024), <https://psycnet.apa.org/doiLanding?doi=10.1037%2Fppm0000536>.

²⁹ Abrahmasson, “Smartphone Bans, Student Outcomes and Mental Health,” Discussion paper, <https://openaccess.nhh.no/nhh-xmlui/handle/11250/3119200>.

³⁰ Haidt, “Social Media is a Major Cause of the Mental Illness Epidemic in Teen Girls. Here’s the Evidence” *After Babel* (Feb. 22, 2023), <https://www.afterbabel.com/p/social-media-mental-illness-epidemic>. By contrast, for studies in which the reduction in social media use was a week or less, just one in seven showed an improvement in mental health, suggesting that “[g]oing cold turkey doesn’t make you happy” (*Ibid.*)

³¹ Nesi, et al, “Emotional responses to social media experiences among adolescents: Longitudinal associations with depressive symptoms” *Journal of Clinical Child and Adolescent Psychology* (2021), <https://pubmed.ncbi.nlm.nih.gov/34424131/>.

³² Coyne et al., “Suicide Risk in Emerging Adulthood: Associations with Screen Time over 10 Years” (2021) 50 *J. Youth & Adolescence* 2324, 2326-27.

³³ Wells et al. “Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show” *Wall Street Journal* (Sept. 14, 2021), https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739?mod=hp_lead_pos7&mod=article_inline.

³⁴ *Ibid.*

³⁵ *Ibid.*

³⁶ See discussion in Rausch, “The Girls Are Not Alright: Responses to Three Claims that the Youth Mental Health Crisis Is Exaggerated” *After Babel* (Apr. 11, 2024), https://www.afterbabel.com/p/the-girls-are-not-alright-responses?utm_source=publication-search.

depression.³⁷ But such explanations fail to account for the fact that the spike in mental health disorders has been observed in numerous other countries, is largely concentrated in Gen Z, has impacted girls more than boys, and began suddenly in the early 2010s – around the time that smartphones became widely adopted.

4) How does social media contribute to mental health harms? According to the Surgeon General, adolescents, in a critical formative period of brain development, are especially vulnerable to potential mental health impacts of social media.³⁸ While noting that several complex factors shape social media’s influence on children and adolescents, the Surgeon General points to two primary risk factors: 1) harmful content, and 2) excessive and problematic use.

Harmful content. According to the Surgeon General, “extreme, inappropriate, and harmful content continues to be easily and widely accessible by children and adolescents” and is “spread through direct pushes, unwanted content exchanges, and algorithmic designs.”³⁹ Such content includes:

- Extreme content such as live depictions of self-harm acts, like asphyxiation or cutting, “which can normalize such behaviors, including through the formation of suicide pacts and posing of self-harm models for others to follow.”⁴⁰
- Bullying and harassment: roughly two-thirds of adolescents are “often” or “sometimes” exposed to hate-based content, with nearly 75% adolescents stating that social media sites do a fair to poor job of addressing online harassment and bullying.⁴¹
- Predatory behaviors, including financial or sexual exploitation of children and adolescents; nearly 6-in-10 adolescent girls surveyed had received unwanted advances from strangers on social media platforms.⁴²

Leaked internal platform studies indicate that youth exposure to unwanted, disturbing, graphic, or sexual content is common and facilitated by platform design.⁴³ According to documents

³⁷ “The Great Rewiring: Is Social Media Really Behind an Epidemic of Teenage Mental Illness?” *supra*. For a response to this article, see Haidt, “Yes, Social Media Really Is a Cause of the Epidemic of Teenage Mental Illness: Two problem with a review in Nature” *After Babel* (Apr. 9, 2024), <https://www.afterbabel.com/p/phone-based-childhood-cause-epidemic>.

³⁸ “Extractive Technology is Damaging our Attention and Mental Health,” Center for Humane Technology, <https://www.humanetech.com/attention-mental-health>.

³⁹ Surgeon General’s Advisory, *supra*, at p. 8.

⁴⁰ *Ibid*.

⁴¹ Alhajji et al., “Cyberbullying, Mental Health, and Violence in Adolescents and Associations With Sex and Race: Data From the 2015 Youth Risk Behavior Survey” *Global pediatric health* (2019), <https://journals.sagepub.com/doi/10.1177/2333794X19868887>; Vogels, “Teens and Cyberbullying,” Pew Research Center: *Internet, Science & Tech* (2022), <https://www.pewresearch.org/internet/2022/12/15/teens-and-cyberbullying-2022/>.

⁴² Nesi, et al., “Teens and mental health: How girls really feel about social media” Common Sense Media (2023), <https://www.common Sense Media.org/research/teens-and-mental-health-how-girls-really-feel-about-social-media>.

⁴³ “Minnesota Attorney General’s Report on Emerging Technology and Its Effects on Youth Well-Being” (Feb. 2025), p. 10-11. https://www.ag.state.mn.us/Office/Reports/EmergingTechnology_2025.pdf. (“Minnesota Attorney General’s Report”)

obtained by the *Wall Street Journal*, one in eight users under the age of 16 experienced unwanted sexual advances on Instagram, facilitated by lax privacy settings.⁴⁴

Additionally, the advisory cites a synthesis of 20 studies demonstrating that many users, especially adolescent girls, experience envy and social comparison, leading to body dissatisfaction, disordered eating behaviors, and low self-esteem. “When asked about the impact of social media on their body image, nearly half (46%) of adolescents aged 13–17 said social media makes them feel worse, 40% said it makes them feel neither better nor worse, and only 14% said it makes them feel better.”⁴⁵ Internal studies by platforms also indicate similar patterns of social comparison, with negative effects on wellbeing.⁴⁶ In an internal Meta study, younger and female users reported much greater rates of feeling “worse about yourself because of other peoples’ posts on Instagram,” with 27.4% of 13-15 year old females reporting this experience over a 7-day period, compared to 14.6% of males in the same age group.⁴⁷

Excessive and problematic use. The advisory cites studies showing that on a typical weekday, nearly one in three adolescents report using screens – most commonly, social media – until midnight or later.⁴⁸ One third or more of girls aged 11-15 feel “addicted” to certain platforms. Excessive use correlates with attention problems, feelings of exclusion, and sleep problems.⁴⁹ Poor sleep, in turn, is linked with neurological development issues, depression, and suicidality.⁵⁰ These findings are borne out by the observations of platforms themselves: internal Meta research detailed in a recent lawsuit concluded that “when social media use displaces sleep in adolescents, it is negatively correlated to indicators of mental health.”⁵¹

Excessive use is driven in part by systems that are optimized to maximize user engagement through design features, such as recommendation algorithms, likes, push notifications, auto-play, and endless scroll.⁵² According to a former social media company executive’s statements, such features were designed intentionally to increase time spent through features that “give you a little dopamine hit every once in awhile.”⁵³ These features “can trigger pathways comparable to

⁴⁴ Jeff Horwitz, “His Job Was to Make Instagram Safe for Teens. His 14-Year-Old Showed Him What the App Was Really Like” *The Wall Street Journal* (Nov. 2, 2023), https://www.wsj.com/tech/instagram-facebook-teens-harassment-safety-5d991be1?mod=hp_featst_pos3.

⁴⁵ Bickham et al., “Adolescent Media Use: Attitudes, Effects, and Online Experiences” Boston Children’s Hospital Digital Wellness Lab (2022), https://digitalwellnesslab.org/wpcontent/uploads/Pulse-Survey_Adolescent-Attitudes-Effectsand-Experiences.pdf.

⁴⁶ Minnesota Attorney General’s Report, *supra*, pp. 11-12.

⁴⁷ *Arizona et al. v. Meta Platforms, Inc., et al.*, Case No. 4:23-cv-05448, Complaint (N.D. Cal. Oct. 24, 2023), <https://storage.courtlistener.com/recap/gov.uscourts.nmd.496039/gov.uscourts.nmd.496039.36.2.pdf>.

⁴⁸ Rideout, V., & Robb, M. B. “Social media, social life: Teens reveal their experiences” Common Sense Media (2018), <https://www.commonsensemedia.org/sites/default/files/research/report/2018-social-mediasocial-life-executive-summary-web.pdf>.

⁴⁹ Surgeon General’s Advisory, *supra*, at p. 10.

⁵⁰ *Ibid.*

⁵¹ *Arizona et al. v. Meta Platforms, Inc., supra.*

⁵² Burhan & Moradzadeh, “Neurotransmitter Dopamine and its Role in the Development of Social Media Addiction” 11 Journal of Neurology & Neurophysiology 507 (2020), <https://www.iomcworld.org/open-access/neurotransmitter-dopamine-da-and-its-role-in-the-development-of-social-mediaaddiction.pdf>.

⁵³ Alex Hern, “Never get high on your own supply” – why social media bosses don’t use social media,” *The Guardian* (Jan. 23, 2018), <https://www.theguardian.com/media/2018/jan/23/never-get-high-on-your-own-supply-why-social-media-bosses-dont-use-social-media>.

addiction.”⁵⁴ Young people with still-developing pre-frontal cortexes who crave social reward and lack inhibition are especially susceptible.⁵⁵

5) The Surgeon General’s call for a social media mental health warning label. In June of 2024, the Surgeon General called for safety warning labels – akin to those a predecessor called for on cigarettes in 1964 – on social media platforms in order to remind teens and parents that social media has not been proven safe.⁵⁶ The Surgeon General wrote:

The mental health crisis among young people is an emergency — and social media has emerged as an important contributor. Adolescents who spend more than three hours a day on social media face double the risk of anxiety and depression symptoms, and the average daily use in this age group, as of the summer of 2023, was 4.8 hours. Additionally, nearly half of adolescents say social media makes them feel worse about their bodies.

It is time to require a surgeon general’s warning label on social media platforms, stating that social media is associated with significant mental health harms for adolescents. A surgeon general’s warning label . . . would regularly remind parents and adolescents that social media has not been proved safe. Evidence from tobacco studies show that warning labels can increase awareness and change behavior.⁵⁷

6) This bill mandates a mental health warning label for social media platforms. This bill responds to the Surgeon General’s call for a warning label. Adopting the definition of “social media platform” from section 124 of the federal Trafficking Victims Prevention and Protection Reauthorization Act of 2022⁵⁸ – the definition used in Senators Britt and Fetterman’s “Stop the Scroll Act,” a federal social media warning label bill – AB 56 requires platforms to display the following message, based on the Surgeon General’s advisory:

“The Surgeon General has advised that there are ample indicators that social media can have a profound risk of harm to the mental health and well-being of children and adolescents.”

The message – which occupies at least 75 percent of the screen for 90 seconds – must be displayed each calendar day when the user first logs on to the platform, again after three hours of cumulative active use, and every hour thereafter, until the end of the calendar day, when the process resets.

7) How much certainty is needed to justify a warning label? In calling for social media mental health warning labels, the Surgeon General stated: “One of the most important lessons I learned in medical school was that in an emergency, you don’t have the luxury to wait for perfect information. You assess the available facts, you use your best judgment, and you act quickly.”⁵⁹ Although the science has not definitively established a causal relationship between social media

⁵⁴ Surgeon General’s Advisory, *supra*, at p. 9.

⁵⁵ *Ibid.*

⁵⁶ Dr. Vivek Murthy, “Surgeon General: Why I’m Calling for a Warning Label on Social Media Platforms” *New York Times* (Jun. 17, 2024), <https://www.nytimes.com/2024/06/17/opinion/social-media-health-warning.html>. (“Surgeon General’s Call for a Warning Label”).

⁵⁷ *Ibid.*

⁵⁸ 42 U.S.C. § 1862w.

⁵⁹ Surgeon General’s Call for a Warning Label, *supra*.

use and youth mental health harms, the best available evidence strongly suggests a significant correlative relationship, especially for heavy users and girls. The question for this Committee is whether this suffices to justify a public health campaign using warning labels on social media.

Comparison to the public health debate about tobacco. The discussion about whether social media in fact contributes to bad health outcomes tracks a similar debate in the 1950s and 1960s about tobacco. Professor Anna Lembke, Chief of the Stanford Addiction Medicine Dual Diagnosis Clinic, writes:

For much of the 20th century, the scientific literature reported an “association” (correlation) between exposure to cigarettes and the occurrence of lung cancer, that is, lung cancer was found to have occurred more frequently among smokers. However, cigarette manufacturers denied their products “caused” the increased number of lung cancer cases and attributed the association instead to other factors—third variables like innate propensity for cancer and exposure to other environmental toxins, such as asbestos and air pollution.

In 1956, the noted British epidemiologists Sir Austin Bradford Hill and Sir Richard Doll published an influential study of smoking and lung cancer among physicians in Britain. This article rejected alternative explanations for the increase in lung cancer, e.g., the claim “that smoking does not produce cancer in a person in whom cancer would not otherwise have occurred at all, but merely determines the primary site of a growth that is destined to appear in some part of the body,” and that “atmospheric pollution” might explain the increased risk. Hill and Doll observed a higher mortality rate in smokers than in non-smokers, a higher mortality rate in heavy smokers than in light smokers, and a higher mortality rate in those who continued to smoke than in those who gave it up. In 1964, their study became part of the data set that resulted in the 1964 Report of the United States Surgeon General that “cigarette smoking is causally related to lung cancer in men; the magnitude of the effect of cigarette smoking far outweighs other factors. The data for women, though less extensive, point in the same direction.”⁶⁰

Like tobacco, worse outcomes from social media use are associated with heavy users, associations begin to reverse when use is reduced, and there does not appear to be a strong alternative explanation for the observed higher rate of harms in users. These factors suggest a sufficient connection between social media use and youth mental health outcomes to warrant a public health campaign.

The comparison only goes so far. Unlike tobacco, for which no amount of consumption is healthy, social media consumption can, as the Surgeon General acknowledges, be beneficial. A report from Common Sense Media frames social media as a “double-edged sword,” finding that “[m]any groups of youth turn to these digital technologies for emotional support, connection with friends and family, learning about ways to support their own mental health, and just decompressing by having fun.”⁶¹ Additionally, the complexity and variation among social media

⁶⁰ Lembke, “Applying The Bradford Hill Criteria To Social Media Use and Adolescent Mental Health” *After Babel* (2024), https://www.afterbabel.com/p/bradford-hill-social-media?utm_source=publication-search.

⁶¹ “A Double-Edged Sword: How Diverse Communities of Young People Think About the Multifaceted Relationship Between Social Media and Mental Health” Common Sense Media (2024), https://www.commonsensemedia.org/sites/default/files/research/report/2024-double-edged-sword-hopelab-report_final-release-for-web-v2.pdf.

platforms, users, and their individual experiences make it more difficult to ascertain the effects of social media on average for users. Nevertheless, these differences do not undermine the Surgeon General's ultimate conclusion that the available evidence is sufficient to justify a mental health warning label.

Comparison to other products that may pose a health risk. The Surgeon General argues we should treat social media as we do other consumer products that may prove defective:

It doesn't have to be this way. Faced with high levels of car-accident-related deaths in the mid- to late 20th century, lawmakers successfully demanded seatbelts, airbags, crash testing and a host of other measures that ultimately made cars safer. This January the F.A.A. grounded about 170 planes when a door plug came off one Boeing 737 Max 9 while the plane was in the air. And the following month, a massive recall of dairy products was conducted because of a listeria contamination that claimed two lives.

Why is it that we have failed to respond to the harms of social media when they are no less urgent or widespread than those posed by unsafe cars, planes or food? These harms are not a failure of willpower and parenting; they are the consequence of unleashing powerful technology without adequate safety measures, transparency or accountability.⁶²

The available evidence warrants action to protect kids. Even among the most ardent skeptics of the hypothesis that social media causes mental health problems, there is consensus that: kids and teens are experiencing a mental health crisis; there is strong correlative evidence linking excessive social media use to youth mental health issues; and interventions are needed to address this crisis.^{63,64} Professor Odgers argues that the data more likely suggests that poor mental health leads to excessive social media use, rather than reverse.⁶⁵ Assuming for the sake of argument that this is the case, intervention would be warranted nonetheless. As over 90% of teens are on social media, the most efficient strategy to educate teens on this mental health crisis would be to meet teens where they are at.⁶⁶

Indeed, there is now a broad, bipartisan consensus among policymakers that major action is warranted. Last year, the notoriously polarized and paralyzed United States Senate passed the Kids Online Safety Act – which would have imposed on social media platforms a duty of care towards minors and restrictions on the use of addictive features – by a vote of 91-3. Other jurisdictions have taken more aggressive action, with Florida banning social media for children under the age of 14 and Australia recently banning it for children under the age of 16. Congress is considering a ban for younger children, as are other states. Developments such as these suggest a warning label is well justified.

As the Surgeon General writes, in concluding his call to action:

⁶² Surgeon General's Call for a Warning Label, *supra*.

⁶³ Odgers and Hayes. "Let's Stop Shaming Teens About Social Media Use" (Feb. 1, 2024), <https://ascd.org/el/articles/lets-stop-shaming-teens-about-social-media-use>.

⁶⁴ Odgers and Haidt. "Making Sense of the Research on Social Media & Youth Mental Health," Youth-Nex. (Oct. 8, 2024), <https://www.youtube.com/watch?v=Ewx4pWOH-I&t=1s>.

⁶⁵ Odgers, "The Great Rewiring: Is Social Media Really Behind an Epidemic of Teenage Mental Illness?" *Nature*, (2024), <https://www.nature.com/articles/d41586-024-00902-2>.

⁶⁶ Vogels et al, "Teens, Social Media and Technology 2022" (2022), <https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/>.

The moral test of any society is how well it protects its children. . . . We have the expertise, resources and tools to make social media safe for our kids. Now is the time to summon the will to act. Our children’s well-being is at stake.⁶⁷

8) **Compelled speech.** “The First Amendment’s guarantee of freedom of speech makes no distinction of ‘constitutional significance’ ‘between compelled speech and compelled silence.’”⁶⁸ By requiring social media platforms to warn users of potential mental health impacts, this bill implicates the First Amendment.

A long-standing and widespread staple of consumer protection, health and safety warnings have been applied to numerous commercial products, including on nutritional labels for foods, health inspection results for restaurants, and warnings on products containing tobacco, alcohol, pharmaceuticals, toxins, flammable or corrosive substances, and carcinogens.⁶⁹ Such warnings enjoy a much more lenient standard of judicial scrutiny – “akin to a rational basis test”⁷⁰ – than other forms of First Amendment infringements. Under this standard, the warning label must contain “‘purely factual and uncontroversial information’”⁷¹ and must not be “unduly burdensome.”⁷²

Purely factual and uncontroversial? A compelled statement, though factually true, may be deemed controversial if it is “hotly disputed.”⁷³ Relevant to this analysis is the effect on the speaker. In *Nat’l Inst. of Fam. and Life Advocs. v. Becerra*, the Supreme Court in a 5-4 decision written by Justice Clarence Thomas held that California could not require anti-abortion “crisis pregnancy centers” to post a notice giving factual information about state-provided services, including abortion, which is “anything but an ‘uncontroversial’ topic.”⁷⁴ But the fact that a “purely factual statement . . . can be tied in some way to a controversial issue” does not make the statement controversial “for that reason alone.”⁷⁵ Rather, the controversy stemmed from the fact that “the compelled statement took sides in a heated political controversy, forcing the clinic to convey a message fundamentally at odds with its mission.”⁷⁶

⁶⁷ Surgeon General’s Call for a Warning Label, *supra*.

⁶⁸ *X Corp. v. Bonta* (9th Cir. 2024) 116 F.4th 888, 900.

⁶⁹ See Symposium: *Compelled Speech: The Cutting Edge of First Amendment Jurisprudence: Compelled Speech and the Regulatory State* (2022) 97 Ind. L.J. 881, 894-895.

⁷⁰ *X Corp. v. Bonta*, *supra*, 116 F.4th at p. 900, quoting *Nat’l Ass’n of Wheat Growers v. Bonta* (9th Cir. 2023) 85 F.4th 1263, 1266 (*Wheat Growers*).

⁷¹ *CTIA - The Wireless Ass’n v. City of Berkeley* (9th Cir. 2019) 928 F.3d 832, 842 (*CTIA*), quoting *Zauderer v. Off. of Disciplinary Couns. of Sup. Ct. of Ohio* (1985) 471 U.S. 626, 651 (*Zauderer*); *Am. Bev. Ass’n v. City & Cty. of San Francisco* (9th Cir. 2019) 916 F.3d 749, 756 (“*Zauderer* provides the appropriate framework to analyze a First Amendment claim involving compelled commercial speech . . . when the government requires health and safety warnings”); *Nat’l Inst. of Fam. and Life Advocs. v. Becerra* (2018) 585 U.S. 755, 775 (*NIFLA*) (stating that “we do not question the legality of health and safety warnings long considered permissible, or purely factual and uncontroversial disclosures about commercial products”); *X Corp. v. Bonta* (9th Cir. 2024) 116 F.4th 888, 901 (“retail product warnings” are “characterized . . . as commercial speech” even though they are “not a clear fit” with the general rule that commercial speech involves speech that proposes a commercial transaction); see also *Chamber of Commerce of United States v. United States SEC* (5th Cir. 2023) 85 F.4th 760, 768 (“[s]tates may require commercial enterprises to disclose ‘purely factual and uncontroversial information’ about their services”).

⁷² *Zauderer*, *supra*, 471 U.S. at p. 651.

⁷³ *Wheat Growers*, *supra*, 85 F.4th at p. 1279.

⁷⁴ *NIFLA*, *supra*, 585 U.S. at p. 769.

⁷⁵ *CTIA*, *supra*, 928 F.3d at p. 845.

⁷⁶ *Ibid*.

In the consumer protection context – where “every warning of a product’s risk to consumers bears a message at odds with the manufacturer’s mission to sell more products”⁷⁷ – courts also look to whether the warning requirement is “controversial from an objective scientific standpoint.”⁷⁸ In this context, “‘uncontroversial’ does not mean ‘unanimous.’”⁷⁹ Thus, a panel of the Ninth Circuit upheld Berkeley’s ordinance requiring manufacturers to provide prospective purchasers a summary of FCC-required disclosures regarding risks associated with radiation, even though cell phone manufacturers “stoutly maintain[ed] that cell phones present no danger whatsoever.”⁸⁰ The court concluded that “[t]he FCC’s required disclosure is no more and no less than a safety warning, and Berkeley’s required disclosure is a short-hand description of the warning the FCC already requires cell phone manufacturers to include in their user manuals.”⁸¹

By contrast, a Ninth Circuit panel found Proposition 65 warning requirements controversial where there was “robust disagreement by reputable scientific sources,” rendering the warning “controversial because it elevates one side of a legitimately unresolved scientific debate about whether eating foods and drinks containing acrylamide increases the risk of cancer.”⁸² More recently, a Ninth Circuit panel found controversial the conclusion that glyphosate – the active ingredient in the Roundup, an herbicide manufactured by Monsanto Company – was “‘probably carcinogenic to humans.’”⁸³ That conclusion had been drawn by the International Agency for Research on Cancer (IARC) but was “not shared by a consensus of the scientific community.”⁸⁴ The court stated that the IARC “stands essentially alone in its determination that glyphosate is probably carcinogenic to humans.”⁸⁵ Thus, the statement that glyphosate was a “known” carcinogen, while true, was misleading, as it implied that this conclusion was the broadly-shared position of the state.⁸⁶

Scientific certainty is not necessarily required so long as the warning is not misleading. In *CTIA—The Wireless Ass’n v. City & County of San Francisco*,⁸⁷ San Francisco’s ordinance requiring cell phone retailers to distribute a factsheet regarding radio frequency (RF) radiation stated “Although studies continue to assess potential health effects of mobile phone use, the World Health Organization has classified RF energy as a possible carcinogen.”⁸⁸ The factsheet also had recommendations for reducing exposure to RF. The district court concluded the factsheet gave the “overall impression . . . that cell phones are dangerous and that they have somehow escaped the regulatory process.”⁸⁹ The court stated the misimpression could be corrected by adding: “Although all cell phones sold in the United States must comply with RF safety limits set by the FCC, no safety study has ever *ruled out* the possibility of human harm

⁷⁷ *Wheat Growers, supra*, 85 F.4th at p. 1284 (dis. opn. of Schroeder, J.).

⁷⁸ *Id.* at p. 1278.

⁷⁹ *Ibid.*

⁸⁰ *CTIA, supra*, 928 F.3d at p. 848.

⁸¹ *Ibid.*

⁸² *California Chamber of Commerce v. Council for Education and Research on Toxics* (9th Cir. 2022) 29 F.4th 468, 478.

⁸³ *Id.* at p. 473.

⁸⁴ *Wheat Growers, supra*, 85 F.4th at p. 1266.

⁸⁵ *Id.* at p. 1278.

⁸⁶ *See ibid.*

⁸⁷ (N.D. Cal. 2011) 827 F.Supp.2d 1054.

⁸⁸ *Id.* at p. 1058.

⁸⁹ *Id.* at pp. 1062-63.

from RF exposure.”⁹⁰ San Francisco ultimately lost on appeal, but that was because it had “concede[d] that there is no evidence of cancer caused by cell phones.”⁹¹

This can be compared with a recent case – cited by Electronic Frontier Foundation in opposition to the bill – in which the Fifth Circuit Court of Appeals held Texas’s warning label for pornographic websites was controversial. The label read: “Pornography is potentially biologically addictive, is proven to harm human brain development, desensitizes brain reward circuits, increases conditioned responses, and weakens brain function.”⁹² The causal assertions in the warning were rebutted by “credentialed and persuasive experts who, on review of ‘the last several decades of research,’” had found “‘no generally accepted, peer-reviewed research studies or scientific evidence which indicate that viewing adult-oriented erotic material *causes* physical, neurological, or psychological damage such as ‘weakened brain function’ or ‘impaired brain development.’”⁹³

The warning label required by AB 56, by contrast, restates the conclusion of the nation’s top doctor based on his review of numerous generally accepted, peer-reviewed research studies and scientific studies about the ample association between social media use and youth mental health impacts. As a result, a court could readily conclude that the warning is factual and uncontroversial. Nevertheless, to further insulate the bill from challenge, the author has agreed to amend the bill to (1) acknowledge social media may have benefits for some users and (2) allow DPH to update the language of the warning to ensure it remains consistent with the state of the research.

Unduly burdensome? Additionally, a compelled disclosure requirement cannot be “unduly burdensome” such that it “might offend the First Amendment by chilling protected commercial speech.”⁹⁴ For example, the Ninth Circuit found that San Francisco’s requirement that a health warning occupy at least 20% of ads for sugar-sweetened beverages was unduly burdensome because a smaller warning could have accomplished San Francisco’s goal of informing the public of the health impacts of added sugars.⁹⁵ As a result, the overly large warning chilled protected speech.⁹⁶ Another case found that an 11x17 inch warning poster – in addition to a mandatory warning fact sheet and warning sticker – “unduly intruded on the retailers’ wall space. . . There is no reasonable cause for requiring retailers to convert their walls to billboards for the municipal message.”⁹⁷

Under AB 56, the warning label occupies at least 75% of the screen for 90 seconds. This occurs when the person first logs on, again after three hours of cumulative active use, and every hour thereafter – a process that resets daily. Given the size of screens on phones, a larger warning label appears justified, particularly in light of how briefly it lasts. And while 90 seconds may be longer than most children need to read a message, the author argues the longer duration will help signal to children the importance of actually reading and absorbing the warning. Moreover,

⁹⁰ *Id.* at p. 1063, emphasis added.

⁹¹ *CTIA - Wireless Ass’n v. City & County of San Francisco* (9th Cir. 2012) 494 F.App’x 752, 754.

⁹² *Free Speech Coal., Inc. v. Paxton* (5th Cir. 2024) 95 F.4th 263, 268, fn. 5. Another holding of that case – that Texas’s requirement that pornographic websites age-gate their sites – is on appeal before the Supreme Court.

⁹³ *Id.* at p. 282, emphasis added.

⁹⁴ *Zauderer, supra*, 471 U.S. at p. 651.

⁹⁵ *Am. Bev. Ass’n v. City & Cty. of San Francisco* (9th Cir. 2019) 916 F.3d 749, 753 (en banc).

⁹⁶ *Ibid.*

⁹⁷ *CTIA v. City & County of San Francisco* (N.D.Cal. 2011) 827 F.Supp.2d 1054, 1063.

unlike a warning that permanently occupies 1/5 of an advertisement, thereby limiting the remaining room that can be used for conveying a message, there does not appear to be any chilling effect on protected speech given the ephemeral nature of the warning. Finally, although the warning is displayed regardless of the age of the user, it is reasonable to show it to adults as well in order to help further the public awareness goals of the bill. Thus, it seems unlikely that the bill would be struck down on this basis.

9) **Amendments.** To ensure the warning label fully aligns with the current state of research, the author has agreed to amend the bill to: (1) reword the warning and include an acknowledgement that social media may have benefits for some users, and (2) allow the Director of the Department of Public Health, by regulation, to update the warning label to remain consistent with the state of research.

The reworded warning will read, in its entirety, as follows:

“The Surgeon General has warned that, while social media may have benefits for some young users, social media is associated with significant mental health harms and has not been proven safe for young users.”

DPH’s authority to update the warning will be added as paragraph (3) of Section 28002(a) and will read as follows:

(3) The director of the California Department of Public Health may, through regulation, modify the warning set forth in paragraph (2) in furtherance of the purposes of this chapter.

ARGUMENTS IN SUPPORT: Attorney General Rob Bonta, a co-sponsor of the bill, writes:

Social media warning labels are an equitable, effective, and transparent way to deliver public health information. By requiring warning labels on social media platforms, AB 56 provides families, children, and others information about the risks of social media use. Just as smoking rates were reduced in part through the global adoption of health warning labels that built awareness of the negative effects of smoking, this requirement can help raise public awareness and turn the tide in a growing public health crisis.

Common Sense Media, a co-sponsor of the bill, writes jointly with Children’s Advocacy Institute:

The necessity of this bill stems from consistent findings that link heavy social media usage with negative mental health outcomes, including increased risks of low self-esteem and disordered eating among adolescents. Furthermore, research and reports, including a notable op-ed by the U.S. Surgeon General, call for immediate action to mitigate these risks, including specifically, use of a warning label for social media.

AB 56 introduces a critical public health measure by making the dangers of social media use transparent, thus ensuring that users, particularly young individuals, are better informed about the risks of their excessive use. The bill does not seek to alter any other aspect of California law, but rather to empower our youth and their families with knowledge that can lead to safer online behaviors.

The American Academy of Pediatrics, California, writes:

The correlation between social media usage and mental health issues is well-documented. Research indicates that excessive use is associated with increased risks of depression, anxiety, and psychological distress. Furthermore, a 2025 study highlighted that happiness and life satisfaction among young people have significantly declined over the past decade, correlating with the rise of smartphones and social media.

AB 56 proposes a practical solution by requiring social media platforms to display a black box warning to users upon initial daily access, after three hours of cumulative active use, and at least once per hour thereafter. This measure aims to increase awareness about the potential mental health risks associated with prolonged social media engagement, empowering users to make informed decisions about their online activity.

By implementing these warnings, California can lead the nation in promoting digital well-being and safeguarding the mental health of its residents.

Professor Jonathan Haidt and researcher Zach Rausch, who together wrote *The Anxious Generation*, write:

Social media platforms, as they currently exist, are designed in ways that drive an unacceptable number of kids into crisis. As documented in our book and our substack, heavy social media use greatly increases the risk of depression, anxiety, eating disorders, addiction, and an array of other harms for adolescents. Hundreds of thousands, if not millions, of American kids are being harmed by these platforms every year. Despite the health costs of using these platforms, many kids continue to use them, in part because of design patterns and AI powered algorithms that maximize engagement. In fact, one recent study found 72% of kids report feeling manipulated by social media platforms. Children are still developing their ability to exhibit self-control in the face of immediate rewards—they know that they need help to control their online behavior. Kids' inability to control their use is no accident. The platforms are purposely designed to keep kids on screens, so that the companies can sell their attention to advertisers to increase profits.

The California Social Media Warning Law would make these harms widely known. This legislation would require tech companies to inform users of the harms of these products, with especially heavy users receiving more periodic warnings. Warning labels have been mandated on cigarettes for decades and are generally deemed effective at raising awareness of risk, though the size, use of images, and novelty of the warning all play a role in maximal effectiveness.

ARGUMENTS IN OPPOSITION: In opposition to the bill, an industry coalition that includes TechNet and California Chamber of Commerce argues:

First, the bill raises several constitutional concerns as it requires a government-mandated label on user-generated speech. The label isn't narrowly tailored to address the stated risk of harm to youth mental health. For example, the bill requires a 90 second, unskippable warning to be applied to every site that meets the bill's definition of social media, for every user regardless of whether they are actually a minor. It doesn't apply to specific content, accounts, or even platforms that are most likely to pose risks of harm to minor users.

Additionally, the bill infringes on the speech rights of minors and adults alike by creating a significant barrier to access information and communicate with others. Some studies have shown that 40% of users will give up waiting for a website to load after just 3 seconds. It is highly likely that users will navigate away from these sites or will find workarounds to prevent this label from appearing.

Second, the warning label is unlikely to be accurate in the majority of situations. It obviously is inapplicable to adults trying to access social media, but for most minors trying to communicate with friends or access useful or educational information the label tells them nothing about what kinds of content or online behaviors are most likely to impact their mental health. Courts would examine the government's interest in compelling platforms to provide an inaccurate label, one that would apply regardless of the user demographics, content, or safety features the platform had implemented. As a result, the label is clearly not the least restrictive means to improve youth mental health.

The U.S. Surgeon General's report that the bill references found positive outcomes of social media use, not just potential risks. As written, the warning label lacks clarity, fails to accurately reflect the complexity of the Surgeon General's report, and oversimplifies nuanced findings, and therefore risks being perceived as misleading or incomplete. Such miscommunication could diminish the platform's credibility and confuse users, particularly parents, educators, and policymakers.

Lastly, the fundamental policy objective of the proposed warning label is unclear. If the goal is to increase user awareness or change user behavior, it is essential to assess whether a warning label is the most effective tool to achieve that outcome. A poorly executed warning label may have limited impact, especially if it interrupts user experience or becomes repetitive and ignored over time. Research on "warning fatigue" suggests that users may disregard frequently encountered warnings, diminishing their intended effect.

Electronic Frontier Foundation similarly raises First Amendment concerns and additionally argues:

The interaction between social media and young people's mental health is far more nuanced than the generalized claim that it is harmful. In fact, there is no scientific consensus that social media is harmful to children's mental health. Social science even shows that social media can help children overcome feelings of isolation and anxiety. This is particularly true for LGBTQ+ teens. EFF conducted a survey where young people told us that online platforms are the safest spaces for them, where they can say the things they can't say in real life 'for fear of torment.' They say these spaces have improved their mental health and given them a 'haven' to talk openly and safely. Pew Research findings also find that teens are more likely to report positive than negative experiences in their social media use.

Oakland Privacy, in an oppose-unless-amended position, writes:

Firstly, as a matter of law, the delivery of engagement features by algorithms to minors without parental consent has been prohibited in the state by last year's SB 976. While it is true that an appeal is pending at the 9th Circuit in *Net Choice v. Bonta*, to date that appeal has not been successful with regard to engagement algorithms. The law has a rulemaking process prior to implementation, but nonetheless we suggest that the bill's findings be more aligned with the current state of the law.

More substantively, the delivery pattern outlined in the bill of 75-100% of the screen area, 90 seconds in duration and provided upon 1) opening the program, 2) after three hours of cumulative use, and 3) then again once an hour after that, is severely intrusive, harassing to users, and overly-broad to deliver on the bill's intent.

Moreover, the message selected . . . is not directed at nor does it have general applicability for any social media users who are not under 18 nor the guardians of children under 18.

REGISTERED SUPPORT / OPPOSITION:

Support

California Department of Justice, Attorney General, Rob Bonta (co-sponsor)
Common Sense Media (co-sponsor)
American Academy of Pediatrics, California
California Initiative for Technology & Democracy, a Project of California Common CAUSE
Childrens Advocacy Institute
County Behavioral Health Directors Association (CBHDA)
Jewish Family and Children's Services of San Francisco, the Peninsula, Marin and Sonoma Counties
Jonathan Haidt and Zach Rausch, The Anxious Generation
Kids: Keep It Digitally Safe Campaign Contest for Warning Labels on Social Media Platforms
Organization for Social Media Safety
Suicide Awareness Voices of Education (SAVE)

Opposition

Calbroadband
California Chamber of Commerce
Chamber of Progress
Computer & Communications Industry Association
Electronic Frontier Foundation
Technet

Oppose Unless Amended

Oakland Privacy

Analysis Prepared by: Josh Tosney / P. & C.P. / (916) 319-2200