LEGISLATIVE OFFICE BUILDING
1020 "N" STREET, SUITE 162
SACRAMENTO, CA 95814
(916) 319-2200

CHIEF CONSULTANT
JOSH TOSNEY

PRINCIPAL CONSULTANT
JULIE SALLEY

COMMITTEE SECRETARY
MIMI HOLTKAMP

Assembly
California Legislature

ASSEMBLY COMMITTEE ON
PRIVACY AND CONSUMER PROTECTION
REBECCA BAUER-KAHAN, CHAIR
ASSEMBLYMEMBER, SIXTEENTH DISTRICT

INFORMATIONAL HEARING

Tuesday, March 17, 2026
1:30 p.m., State Capitol, Room 437

Online Safety Controls: What They Are, Why They Can Fail, And What We Can Do About It

BACKGROUND PAPER

I. INTRODUCTION

Nearly every child in the United States has internet access by the time they are school age, and by early adolescence, most are active on social media, gaming platforms, artificial intelligence (AI) chatbots, and video-sharing apps. This informational hearing will examine the state of online safety controls for minors, including the types of safety controls available to families, key challenges and failure points that limit the effectiveness of these controls, and a range of potential solutions available to industry, families, and policymakers.

The evidence presented in this background paper reflects a troubling gap: the tools that exist to protect children online are underused, difficult to navigate, often ineffective, and siloed across dozens of platforms. Meanwhile, some commercial incentives can lead to design decisions that work against child safety. To help determine whether that gap can be filled, the Committee will hear testimony from panelists representing the perspectives of parents, children, safety advocates, and industry.

II. ONLINE HARMS

The Family Online Safety Institute's (FOSI) 2025 national survey of parents and kids age 10-17 found that families identified the following online harms in order of perceived risk:

	Total	Parents	Children
Predatory Behavior	70%	**73%**	66%
Giving away personal information	66%	66%	65%
Cyberbullying	63%	63%	63%
Seeing age-inappropriate content	60%	**68%**	52%
Accidentally downloading viruses	49%	41%	**57%**
Data breaches and privacy risks	42%	38%	**46%**
Getting scammed out of money	42%	36%	**47%**
Spending too much time online	39%	**44%**	34%
Spending money without permission	29%	28%	30%
Encountering mis/dis-information	28%	**30%**	25%
Plagiarism related to AI tools	13%	13%	14%

Source: Family Online Safety Institute[1]

Research suggests that parents tend to underestimate both the scope and risks of their children's online activity. Microsoft's 2023 Global Online Safety Survey found that 74% of teens reported experiencing an online risk in the past year – a figure 12 points higher than the 62% of parents who believed their child had done so – and that children reported engaging in a wide range of online activities at far higher rates than their parents perceived.[2] A brief examination of common harms associated with different types of platforms follows.

a. Social media

A 2025 survey of 10,092 11-to-15-year-old adolescents found that 69.5% had at least one social media account. Among social media users, the most common platforms were TikTok (67.1%), YouTube (64.7%), and Instagram (66.0%). A majority (63.8%) of participants under 13 years – the minimum age requirement for these platforms – reported social media use.[3]

---

[1] Family Online Safety Institute, "Connected and Protected: Insights from FOSI's 2025 Online Safety Survey," (May 28, 2025), https://fosi.org/wp-content/uploads/2025/05/Connected-and-Protected-Insights-from-FOSIs-2025-Online-Safety-Survey.pdf.

[2] Microsoft, "New Microsoft Research Illustrates the Online Risks and Value of Safety Tools to Keep Kids Safer in the Digital Environment," Microsoft On the Issues (Feb. 6, 2023), https://blogs.microsoft.com/on-the-issues/2023/02/06/safer-internet-day-global-online-safety-survey-2023/ .

[3] Jason M. Nagata et al., "Prevalence and Patterns of Social Media Use in Early Adolescents," *Academic Pediatrics,* vol. 25(4), May-June 2025, https://www.academicpedsjnl.net/article/S1876-2859(25)00009-9/fulltext.

In May 2023, former Surgeon General Vivek Murthy issued an advisory warning of the potential mental health impacts of social media on young people.[4] The advisory recognizes the benefits of social media for some users but concludes "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."[5] 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, such as compulsive or uncontrollable use. Harmful 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."[6]
- Bullying and harassment: roughly two-thirds of adolescents are "often" or "sometimes" exposed to hate-based content, with nearly 75% of adolescents stating that social media sites do a fair to poor job of addressing online harassment and bullying.[7]
- 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.[8]

The advisory also cites studies showing that on a typical weekday, nearly one in three adolescents report using screens – most commonly, social media – until midnight or later.[9] 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.[10] Poor sleep, in turn, is linked with neurological development issues, depression, and suicidality.[11]

Excessive use is driven in part by systems that are optimized to maximize user engagement through design features, such as personalized recommendation algorithms, likes, push notifications, auto-play, and endless scroll.[12] According to a former social media company

---

[4] "Social Media and Youth Mental Health: The U.S. Surgeon General's Advisory" (May 23, 2023), https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf. ("Surgeon General's Advisory").
[5] *Id.* At p. 4.
[6] *Id*. at p. 8.
[7] 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/.
[8] Nesi, et al., "Teens and mental health: How girls really feel about social media" Common Sense Media (2023), https://www.commonsensemedia.org/research/teens-and-mental-health-how-girls-really-feel-about-social-media.
[9] 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.
[10] Surgeon General's Advisory, *supra*, at p. 10.
[11] *Ibid.*
[12] 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.

executive's statements, such features were designed intentionally to increase time spent through features that "give you a little dopamine hit every once in a while."[13] These features "can trigger pathways comparable to addiction."[14] Young people with still-developing pre-frontal cortexes who crave social reward and lack inhibition are especially susceptible.[15] As of 2024, the average daily social media usage for US adolescents was 4.8 hours.[16]

b. AI chatbots

Generative AI chatbots that can convincingly mimic human conversation have exploded in popularity. Many of these characters purport to provide companionship, offering friendship or romantic or erotic relationships, as well as life-coaching and even therapeutic services. Roughly half of teens report using chatbots, with 24% using them at least weekly and 11% daily.[17]

Privacy is a key concern when it comes to the use of companion chatbots. A Stanford HAI study found that AI companies "employ users' chat data by default to train their models, and some developers keep this information in their systems indefinitely."[18] Replika was fined €5 million by Italy's Garante for allowing children under 13 to access the platform without safeguards.[19]

Additionally, according to a recent report from the Minnesota Attorney General, the widespread use of chatbots by teens "has not been accompanied by corresponding safeguards."[20] These products can be "'extremely addictive'" and "researchers have documented that over-usage and addiction are primary risks of personalized chatbots. Several studies have shown that aggregate positive benefits of chatbots are possible, but investigations by journalists and clinicians suggest that these products are not robust in terms of the quality and safety of their responses."[21] The Minnesota Attorney General's Report concludes:

> Despite in-product reminders that chatbots are not real, the design features of these products are intended to convey a misleading sense of "humanness" such that even trained engineers confuse them with actual humans, especially when these products are trained to state unequivocally that they are indeed people. Given the epidemic of loneliness in society, care needs to be taken in introducing vulnerable youth and adults to products that may appear to fulfill an immediate social need, but where acute harms have already begun to surface and

---

[13] 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.
[14] Surgeon General's Advisory, *supra*, at p. 9.
[15] *Ibid.*
[16] 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.
[17] "Minnesota Attorney General's Report on Emerging Technology and Its Effects on Youth Well-Being" (Feb. 2025), p. 28, https://www.ag.state.mn.us/Office/Reports/EmergingTechnology_2025.pdf. ("Minnesota Attorney General's Report").
[18] Stanford Human-Centered Artificial Intelligence, "Be Careful What You Tell Your AI Chatbot," (Oct. 15, 2025), https://hai.stanford.edu/news/be-careful-what-you-tell-your-ai-chatbot.
[19] https://iapp.org/news/a/italy-s-dpa-reaffirms-ban-on-replika-over-ai-and-children-s-privacy-concerns.
[20] *Ibid.*
[21] Minnesota Attorney General's Report, *supra,* p. 28.

where long-term negative impacts, such as social deskilling and demotivation resulting from substitution for in-person socialization, may arise.[22]

In extreme cases, bots have reinforced delusions or even encouraged dangerous behavior. OpenAI, the company behind ChatGPT, is facing at least 11 lawsuits claiming that the chatbot caused psychological harm, leading to personal injury or wrongful death,[23] including the death by suicide of 16-year-old Adam Raine.[24] OpenAI has since announced the rollout of features designed to protect minors.[25] Character.AI implemented age verification and eventually banned minors from open-ended chats only after the death by suicide of 14-year-old Sewell Setzer III and other lawsuits.[26]

c. Gaming

Online gaming now reaches an estimated 3.42 billion players worldwide, with nearly 80% of children ages 2 to 18 participating globally and 85% of U.S. teenagers reporting they play video games.[27] Within these environments, young people face pervasive harassment, misogyny, and identity-based abuse. The Anti-Defamation League found that 75% of players ages 10 to 17 experienced harassment in online multiplayer games in 2023, with girls and children of color bearing a disproportionate share of targeting.[28] Pew Research similarly found that 41% of teen video game players had been called an offensive name while playing, and 8% had received unwanted sexually explicit material.[29] Los Angeles County and the Georgia Attorney General both launched proceedings against Roblox in February 2026, alleging that the platform – where over 40% of its 151 million daily users are under age 13 – knowingly enabled grooming, sexual exploitation, and contact from predators.[30] Additionally, UNICEF's research indicates that violent extremist groups and criminal cartels use gaming sites to recruit children into organized violence.[31]

---

[22] Minnesota Attorney General's Report, *supra,* p. 29.
[23] Jennifer Valentino-DeVries and Kashmir Hill, "How Bad Are A.I. Delusions? We Asked People Treating Them," *New York Times* (Jan. 26, 2026), https://www.nytimes.com/2026/01/26/us/chatgpt-delusions-psychosis.html?smid=nytcore-ios-share.
[24] Kashmir Hill. "A Teen Was Suicidal. ChatGPT Was the Friend He Confided In," *New York Times* (Aug. 26, 2025), https://www.nytimes.com/2025/08/26/technology/chatgpt-openai-suicide.html.
[25] Angela Yang, "Mom who sued Character.AI over son's suicide says the platform's new teen policy comes 'too late,'" *NBC News* (Oct. 30, 2025), https://www.nbcnews.com/tech/tech-news/characterai-bans-minors-response-megan-garcia-parent-suing-company-rcna240985.
[26] Lisa Eadicicco, "After a wave of lawsuits, Character.ai will no longer let teens chat with its chatbots" *CNN* (Oct. 29, 2025), https://www.cnn.com/2025/10/29/tech/character-ai-teens-under-18-app-changes.
[27] Galen Lamphere-Englund, "Protecting Children in Online Gaming: Mitigating Risks from Organized Violence," UNICEF Innocenti Working Paper (Oct. 2025), p. 2, https://www.unicef.org/innocenti.
[28] Anti-Defamation League, "Hate is No Game: Hate and Harassment in Online Games" (2023), cited in Lamphere-Englund, *supra*, p. 10.
[29] Jeffrey Gottfried and Olivia Sidoti, "Teens and Video Games Today," Pew Research Center (May 9, 2024), p. 4, https://www.pewresearch.org/internet/2024/05/09/teens-and-video-games-today/.
[30] County of Los Angeles, "LA County Sues Roblox for Unfair and Deceptive Business Practices that Endanger and Exploit Children," Press Release (Feb. 19, 2026), https://lacounty.gov/2026/02/19/la-county-sues-roblox-for-unfair-and-deceptive-business-practices-that-endanger-and-exploit-children/; Office of the Georgia Attorney General, "Carr Investigates Roblox for Reports of Child Exploitation," Press Release (Feb. 17, 2026), https://law.georgia.gov/press-releases/2026-02-17/carr-investigates-roblox-reports-child-exploitation.
[31] Lamphere-Englund, *supra*, pp. 11–14.

III. TYPES OF ONLINE SAFETY CONTROLS

Online safety controls, broadly speaking, are tools and features used to manage a child's digital access at the device or app level, or through third-party apps. Key categories are described below.

Device and operating system controls

- App approval and blocking: features that require parental permission before downloading.
- Global time limits: shutoff times for device or specific categories.
- Content ratings: app restrictions by age-rating in the App store.
- Location tracking: real-time GPS tracking of the device.

App-level controls

- Content filtering: algorithms that filter out certain content, websites, hashtags, or keywords.
- Privacy and visibility: sets accounts to private mode or restricts discoverability.
- Direct messaging (DM) regulation: settings to disable DMs entirely or restrict them to friends.
- Family pairing/supervision: linking a parent's account to the child's to monitor the child's activities and control safety settings.
- In-app time management: "take a break" reminders and sleep mode modifications.
- Spending limits: sets limits to the amount of money that can be spent in-game or in-app.
- Engagement feature limitations: restricts engagement-optimizing features like algorithmically-curated feeds, likes, and notifications.

Third-Party controls

- Semantic monitoring: scanning text messages, emails, and saved photos for dangerous content.
- Granular web blocking: blocking specific websites, game servers, or URLs that the OS filters might miss.
- Cross-platform reporting: aggregating activity logs from multiple apps into a single dashboard for parents.
- Alert-dased detection: proactive notifications sent to parents when a potential threat is detected, rather than constant monitoring.

IV. CHALLENGES AND FAILURE POINTS

Online safety controls can fail to ensure a child's safety for a variety of reasons. This section examines practical barriers to effective implementation of controls, including challenges that parents and kids face, as well as structural reasons across platforms that can lead to failures.

a. Practical barriers

Low adoption rates. Many online platforms have seen strikingly low rates of adoption of parental monitoring tools. On Discord and Snapchat, for instance, fewer than 1% of minors have a parent

using the available tools to monitor them.[32] Fewer than 10% of teen accounts on Instagram had enabled supervision features by the end of 2022.[33] Adoption rates are somewhat higher for device-level controls than for platform-specific tools, but remain far from universal: FOSI's 2025 national survey found that 47% of parents use smartphone screentime controls, 46% use controls on desktops, and only 35% use controls on game consoles.[34] Disparities in economic, social, and cultural capital affect adoption rates, leading to "digital parenting divides" with more well-resourced parents better positioned to use the most effective strategies.[35]

Number of apps. A University of Michigan and Common Sense Media study found that teens use an average of 40 different apps, receiving more than 200 notifications per day,[36] while 46% of teens report being online "almost constantly."[37] Setting up safety controls across that landscape can be overwhelming: Dr. Jean Twenge described configuring Mac safety controls as "a day-long exercise in frustration."[38] Cross-platform interactions compound this: a child's risky interaction may, for example, begin in a game chat, continue on Discord, and escalate over Instagram DMs – yet each platform's parental control tool operates in its own silo, with little cross-platform interoperability.

Difficult to use and ineffective controls. Child safety reviewers have noted that parental settings on most social media apps are "buried or hard to use."[39] Even if a parent figures out how to use them, some protections fail when tested. A recent study found that 64% of Meta's safety tools were ineffective – users of teen accounts were shown in appropriate content, including posts involving self-harm, highly sexualized materials, and offensive or misogynistic comments and messages – and just 17% worked as described.[40] Moreover, safety controls often fail to address the harms parents are most worried about. A recent literature review "found no studies that matched the reasons for use with the outcome measures, nor any that compared measures of children's online experiences before and after using parental controls, with one exception that

---

[32] Kat Tenbarge, "Fewer than 1% of parents use social media tools to monitor their children's accounts, tech companies say," NBC News (Mar. 29, 2024), https://www.nbcnews.com/tech/social-media/fewer-1-parents-use-social-media-tools-monitor-childrens-accounts-tech-rcna145592.

[33] Naomi Nix, "Meta says parental controls protect kids, but most don't use them," Washington Post (Jan. 30, 2024), https://www.washingtonpost.com/technology/2024/01/30/parental-controls-tiktok-instagram-use/.

[34] Family Online Safety Institute, "Parental Controls for Online Safety Are Underutilized, New Study Finds," Press Release (May 28, 2025), https://fosi.org/parental-controls-for-online-safety-are-underutilized-new-study-finds/.

[35] Pengfei Zhao, Natalie N. Bazarova, and Natercia Valle, "Digital Parenting Divides: The Role of Parental Capital and Digital Parenting Readiness in Parental Digital Mediation," Journal of Computer-Mediated Communication, vol. 28, no. 5 (Aug. 2023), https://doi.org/10.1093/jcmc/zmad032.

[36] Jenny Radesky et al., "Constant Companion: A Week in the Life of a Young Person's Smartphone Use," University of Michigan C.S. Mott Children's Hospital and Common Sense Media (Sept. 2023), https://www.michiganmedicine.org/health-lab/study-average-teen-received-more-200-app-notifications-day.

[37] Monica Anderson et al., "Teens, Social Media and Technology 2024," Pew Research Center (Dec. 12, 2024), https://www.pewresearch.org/internet/2024/12/12/teens-social-media-and-technology-2024/.

[38] Jean M. Twenge, "I tried to protect my kids from the internet. Here's what happened." *Washington Post* (Sept. 23, 2025), https://www.washingtonpost.com/opinions/interactive/2025/parental-controls-tech-companies-social-media-verify-age/.

[39] BrightCanary, "Instagram vs. Snapchat for Kids" (Apr. 16, 2025), https://www.brightcanary.io/which-is-better-for-kids-instagram-or-snapchat/.

[40] Fairplay "Instagram Teen Accounts fail to protect children, safety tools," (Sept 24, 2025) https://fairplayforkids.org/instagram-teen-accounts-fail-to-protect-children-safety-tools-testing-reveals/.

reported a null effect . . ." [41] As a result, "it cannot be concluded that the evidence supports the claims of tool efficacy. Instead, most research relies on parental perception or satisfaction with improvements following the use of parental controls, and even that shows a mixed pattern of beneficial, null and even adverse results."[42] Pew Research found that parents' top concerns are online predators, sexually explicit content, and cyberbullying, while the dominant tools on the market offer little more than time limits and website blocking, none of which can detect interpersonal threats.[43]

The "whack-a-mole" effect. Age-based platform restrictions have triggered what critics call a "whack-a-mole" effect, pushing children toward less-regulated alternatives rather than off the internet. Signup for virtual private networks, which mask a device's internet activity, surged 1,400% within hours of the UK Online Safety Act taking effect and 1,150% after Florida's law; Australia's social media ban produced a 170% spike in VPN traffic on day one.[44] Some Australian teens have circumvented the ban via exempt platforms that offer fewer safety controls.[45] Some of these apps have no parental controls, allow inappropriate images to appear on a child's device without consent, and operate age verification as nothing more than an easily bypassed checkbox.[46]

Schools and families drive app usage. Restrictions also collide with practical realities of children's daily lives. YouTube – used by 90% of U.S. teens and visited daily by 73% – is the top app accessed on school devices outside core productivity sites.[47] Schools routinely rely on age-restricted apps such as Discord for classroom communication.[48] Parents themselves often enable access: Pew Research's 2025 survey found that 92% allow smartphone use primarily to stay in contact with their child, while 85% allow smartphone usage for entertainment.[49] Additionally, 60% of parents report their child under 12 uses or interacts with a smartphone –

---

[41] Mariya Stoilova, Monica Bulger, and Sonia Livingstone, "Do Parental Control Tools Fulfil Family Expectations for Child Protection? A Rapid Evidence Review of the Contexts and Outcomes of Use," Journal of Children and Media (Oct. 29, 2023), https://doi.org/10.1080/17482798.2023.2265512.

[42] *Ibid.*

[43] Pew Research Center, "Parenting Children in the Age of Screens" (July 28, 2020), https://www.pewresearch.org/internet/2020/07/28/parenting-children-in-the-age-of-screens/.

[44] The Register, "UK VPN Demand Soars After Debut of Online Safety Act" (July 28, 2025), https://www.theregister.com/2025/07/28/uk_vpn_demand_soars/; UK Tech News, "Proton VPN and the UK Online Safety Act" (Aug. 6, 2025), https://www.uktech.news/cybersecurity/proton-vpn-uk-online-safety-act-20250806; VPN Super, "Australia Social Media Ban: December 2025," https://www.vpnsuper.com/vpn-observatory/australia-social-media-ban-under-16-december-2025.

[45] NPR, "Social Media Ban: Children in Australia" (Dec. 10, 2025), https://www.npr.org/2025/12/10/nx-s1-5639694/social-media-ban-children-australia; TechLicious, "Teens Circumvent Social Media Restrictions in US and Australia," https://www.techlicious.com/blog/teens-circumvent-social-media-restrictions-us-australia/; CNBC, "Australia Teens Bypass Social Media Ban via Yope, Lemon8, Discord, VPNs," https://www.cnbc.com/.

[46] Safer Schools, "Is Locket Safe for My Child?" https://saferschools.co.uk/is-locket-safe-for-my-child/; Screen Rant, "Locket Widget App: What Is It and Is It Safe?" https://screenrant.com/locket-widget-app-what-is-it-safety/.

[47] Monica Anderson et al., "Teens, Social Media and Technology 2024," Pew Research Center (Dec. 12, 2024), https://www.pewresearch.org/internet/2024/12/12/teens-social-media-and-technology-2024/; Pew Research Center, "Teens and Social Media Fact Sheet" (2024), https://www.pewresearch.org/internet/fact-sheet/teens-and-social-media-fact-sheet/.

[48] Lightspeed Systems, "Making YouTube Safe and Productive for Students," EdTech App Report (2024), https://www.lightspeedsystems.com/blog/making-youtube-safe-and-productive-for-students/ .

[49] Pew Research Center, "How Parents Manage Screen Time for Kids" (Oct. 8, 2025), https://www.pewresearch.org/internet/2025/10/08/how-parents-manage-screen-time-for-kids/.

most often a parent's own device, meaning that children frequently access the internet through devices that are not configured for their protection.[50]

Privacy tradeoffs. Parental monitoring tools can undermine privacy, leading to family conflict and erosion of trust in the parent-child relationship.[51] Kids respond accordingly: 76% gave parental control apps one star in app store reviews, describing them as "stalking," and a UK regulator's 2024 report found that 40% of teens had taken active steps to evade parental supervision, with 68% aware of circumvention methods.[52] Researchers have found that parental control apps were "no different from stalkerware," concluding that tools designed to hide their presence on a child's device are fundamentally inappropriate for child safety.[53] Impacts can be especially pronounced for LGBTQ+ youth who rely on online support networks.[54]

Beyond concerns with parental surveillance, privacy protections can undermine the utility of third-party monitoring tools. For example, Apple's iOS privacy architecture prevents third-party apps from monitoring cellular data entirely – one tap by a child disables WiFi and suspends monitoring – while end-to-end encryption on iMessage and WhatsApp means such apps can only scan those services through periodic backups on separate hardware.[55]

b. Structural challenges

Incentives to maximize engagement. A Harvard T.H. Chan School of Public Health study concluded that in 2022 alone, six major platforms collectively generated $11 billion in U.S. advertising revenue from users under 18, including $2 billion from children 12 and under,

---

[50] Research Center, "How Parents Manage Screen Time for Kids" (Oct. 8, 2025), https://www.pewresearch.org/internet/2025/10/08/how-parents-manage-screen-time-for-kids/.

[51] Mariya Stoilova, Monica Bulger, and Sonia Livingstone, "Do Parental Control Tools Fulfil Family Expectations for Child Protection? A Rapid Evidence Review of the Contexts and Outcomes of Use," Journal of Children and Media (Oct. 29, 2023), https://doi.org/10.1080/17482798.2023.2265512; University of Central Florida, "Apps That Keep Children Safe Online May Be Counterproductive," UCF News, https://www.ucf.edu/news/apps-keep-children-safe-online-may-counterproductive/.

[52] Ruba Abu-Salma et al., "Children's App Store Reviews of Parental Control Apps," ACM CHI Conference on Human Factors in Computing Systems (2018), https://dl.acm.org/doi/10.1145/3173574.3173698; Ofcom, "Children and Parents: Media Use and Attitudes Report 2024," https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-children/children-and-parents-media-use-and-attitudes-report-2024.

[53] Jonah Rauber et al., "Parental Control or Stalkerware? An Analysis of Commercially Available Parental Monitoring Apps," Proceedings on Privacy Enhancing Technologies (2025), https://petsymposium.org/popets/2025/popets-2025-0052.php; UCL News, "Unofficial Parental Control Apps Put Children's Safety and Privacy at Risk" (Mar. 2025), https://www.ucl.ac.uk/news/2025/mar/unofficial-parental-control-apps-put-childrens-safety-and-privacy-risk.

[54] Kirsten Weir, "Social media brings benefits and risks to teens. Psychology can help identify a path forward," American Journal of Psychology, Vol. 54, No. 6 (Sep. 2023), https://www.apa.org/monitor/2023/09/protecting-teens-on-social-media.

[55] Sarah Mitchell, "Bark Failed on My iPhone – Here's What Actually Works (2026)," WhitelistVideo Blog (updated Feb. 6, 2026), https://whitelist.video/blog/bark-doesnt-work-on-ios; Safety Detectives, "Bark Parental Control Review," https://www.safetydetectives.com/best-parental-control/bark/; Bark Technologies, "Frequently Asked Questions," https://www.bark.us/faq/; Bark Technologies, "How Bark Works," https://www.bark.us/bark-app/; Cybernews, "Bark Parental Control App Review," https://cybernews.com/best-parental-control-apps/bark-review/.

creating substantial commercial pressure against reform.[56] One of the lead authors, Professor Bryn Austin, concluded: "Although social media platforms may claim that they can self-regulate their practices to reduce the harms to young people, they have yet to do so, and our study suggests they have overwhelming financial incentives to continue to delay taking meaningful steps to protect children."[57]

Race to the bottom. Platform safety investments carry real competitive costs – including friction in sign-up flows, reduced engagement metrics, foregone data collection – while competitors that forgo those investments face no equivalent penalty. This dynamic can create a structural disincentive for platforms to move first on safety. As researchers Zach Rausch and Jonathan Haidt have argued, social media companies are caught in a classic collective action problem: due to intense competition and limited regulation, they are engaged in a race to the bottom, and any company that drops out of the race risks losing its users to platforms that stay in it.[58] The effect of such incentives can arguably be seen in Meta's alleged former policy of allowing 17 strikes before it suspended accounts engaged in the "trafficking of humans for sex."[59] Similarly, the FTC's 2024 review of nine major platforms' data privacy practice concluded that "self-regulation is failing" and called on Congress to enact comprehensive federal privacy legislation – a recognition that without a regulatory floor that raises the standard for all competitors simultaneously, market incentives may continue to work against child safety.

Apps designed for adults. Rather than building apps specifically intended for kids, the dominant industry approach has been to add safety features and content restrictions to adult-focused products, resulting in a degraded experience that children and teens often reject. The early version of YouTube Kids is illustrative: despite being designed for children under 13, only 23% of 8-to-12-year-olds reported using it, compared to 76% who used regular YouTube.[60] The alternative approach of building genuinely engaging, age-appropriate products that are safe by design is costly and rare. A small number of newer entrants, such as Coverstar, have attempted this model, incorporating privacy protections, AI-assisted moderation, and the elimination of direct messaging; independent reviewers have rated it among the safer options available to tweens.[61] YouTube has since introduced more nuanced age-based content tiers and supervised account options, but child reviewers consistently describe the app as engaging mainly for very

---

[56] Harvard T.H. Chan School of Public Health, "Social Media Platforms Generate Billions in Annual Ad Revenue from U.S. Youth" (2023), https://hsph.harvard.edu/news/social-media-platforms-generate-billions-in-annual-ad-revenue-from-u-s-youth/.

[57] *Ibid.*

[58] Zach Rausch and Jonathan Haidt, "Solving the Social Dilemma: Many Paths to Social Media Reform," *After Babel* (Nov. 28, 2023), https://www.afterbabel.com/p/solving-the-social-dilemma.

[59] Jonathan Limehouse, "Meta had 17-strikes policy for sex trafficking posts, lawsuit alleges," *USA Today* (Nov. 22, 2025), https://www.usatoday.com/story/tech/2025/11/22/meta-strike-policy-sex-trafficking-violations-testimony/87425612007/.

[60] Common Sense Media, "The Common Sense Census: Media Use by Tweens and Teens" (2019), https://www.commonsensemedia.org/press-releases/the-common-sense-census-media-use-by-tweens-and-teens-new-research-finds-youtube-videos-beat-out-tv-and.

[61] Bark Technologies, "Is Coverstar Safe? A Coverstar Review for Parents" (Nov. 21, 2023), https://www.bark.us/app-reviews/apps/coverstar-app-review/; BrightCanary, "5 TikTok Alternatives for Kids" (Sept. 3, 2025), https://www.brightcanary.io/tiktk-alternatives-for-kids/.

young children, with older users finding it too limited.[62] Youth-focused online experiences are discussed in more detail below.

V. CALIFORNIA'S EFFORTS TO PROTECT CHILDREN ONLINE

California has enacted a number of measures to proactively protect children online, including:

Safe-by-design provisions. AB 2273 (Wicks, 2022) establishes the California Age-Appropriate Design Code Act (AADC), which imposes obligations and restrictions on businesses that provide online services, products, or features likely to be accessed by children. Key provisions include a prohibition on profiling and the collection of a child's precise geolocation; data minimization and purpose limitations; age estimation; prohibitions on dark patterns; and a requirement that covered entities prepare data protection impact assessments (DPIA) before offering online services likely to be accessed by children. NetChoice, a trade association of online businesses, sued on several grounds. Some portions of the AADC – including the provisions relating to DPIAs, dark patterns, profiling, and data minimization – have been blocked from implementation, while others have been allowed to go into effect, including age estimation provisions and limits on geolocation data.[63] Litigation is ongoing.

Addictive feeds. SB 976 (Skinner, 2024), the "Protecting Our Kids from Social Media Addiction Act" regulates how internet platforms allow minors to access personalized recommendation algorithms. The Act restricts minors' access to algorithmic feeds, requires certain default settings, including restricting notifications, hiding like counts, and making accounts private, and mandates that the Attorney General adopt regulations governing age assurance by 2027. The Act also requires covered companies to annually disclose the number of minors that use their services. NetChoice has challenged the Act on First Amendment grounds. Provisions governing notifications, like-counts, and disclosures have been blocked pending the ruling on the merits, while the rest of the Act has been allowed to stand.[64] Litigation is ongoing.

Social media warning labels. AB 56 (Bauer-Kahan, 2025) responds to former Surgeon General Vivek Murthy's call for safety warning labels on social media platforms. Beginning in 2027, social media platforms must display to minors mental health warning labels about the harms associated with social media when the child logs on to the platform and after extended use.

Device-based age assurance. AB 1043 (Wicks, 2025) creates a device-based age assurance system. Operating system providers and covered app stores must transmit to app developers a non-identifying age bracket signal at setup. App developers who receive such signals are deemed to have "actual knowledge" of the user's age.

Companion AI restrictions. SB 243 (Padilla, 2025) requires companion chatbot operators to disclose that the chatbot is AI in certain circumstances, implement protocols for instances when a

---

[62] Google Support, "Content Settings on YouTube Kids," https://support.google.com/youtubekids/answer/7554914?hl=en; Common Sense Media, user reviews, YouTube Kids app, https://www.commonsensemedia.org/app-reviews/youtube-kids/user-reviews/child.
[63] *NetChoice, LLC v. Bonta* (March 12, 2026) No. 25-2366 D.C. No. 5:22-cv-08861- BLF.
[64] *NetChoice, LLC v. Bonta* (9th Cir. 2025) 152 F.4th 1002, 1025.

user expresses suicidal ideation or intent to self-harm, and disclose related information to the Office of Suicide Prevention.

The Legislature is currently considering additional child online safety measures, including:

- The creation of an eSafety Commission to oversee implementation and enforcement of online safety legislation (AB 1700, Lowenthal).
- Prohibitions on social media accounts for users under age 16 (AB 1709, Lowenthal).
- Enhanced restrictions on companion chatbots (AB 1988, Pellerin; AB 2023, Wicks & Bauer-Kahan; SB 1119, Padilla; SB 300, Padilla).

VI. POTENTIAL SOLUTIONS FOR INDUSTRY, FAMILIES, AND POLICYMAKERS

There are a range of protective options available to help mitigate some of the risks that children online face. This section briefly highlights many of these options that currently exist, as well as possible options that could be developed.

a. What's available?

i. Education and media literacy

For decades, internet platforms have argued that the internet is not a place designed for children. However, according to the National Center for Education Statistics, some 97 percent of children aged three to 18 had access to the Internet in 2021.[65] By 2024, caregivers reported that children eight years and younger spent roughly two and a half hours with screen media.[66] By age two, 40 percent of children have their own tablet, rising to nearly 60 percent of children by age four.[67] As young children continue to seek out online media, parents look for guidance from educators and industry on best practices for teaching internet safety.

Google's Be Internet Awesome. In 2017, Google started the Be Internet Awesome program, designed to empower children and their families to safely use the internet to make smart decisions. The program offers separate curriculums for both families and educators about how to teach good internet stewardship to children with the tagline "Smart, Alert, Strong, Kind, Brave."[68] The program also has free games that children can play such as Interland where "kids will help their fellow Internauts combat the badly behaved hackers, phishers, oversharers and bullies by practicing the skills they need to be good digital citizens."[69] Google's curriculum for Be Internet Awesome has five fundamental topics of digital citizenship and safety:

---

[65] National Center for Educational Statistics, "Children's Internet Access at Home," U.S. Department of Education, (Aug 2023), https://nces.ed.gov/programs/coe/indicator/cch/home-internet-access.
[66] "The Common Sense Census: Media Use by Kids Zero to Eight," *Common Sense Media,* (2025), https://www.commonsensemedia.org/sites/default/files/research/report/2025-common-sense-census-web-2.pdf.
[67] *Ibid*.
[68] Google Be Internet Awesome, "Digital Safety Resources," https://beinternetawesome.withgoogle.com/en_us/families.
[69] *Ibid.*

- *Share with Care*: Digital Footprint and Responsible Communication
- *Don't Fall for Fake*: Phishing, Scams, and Credible Sources
- *Secure Your Secrets*: Online Security and Passwords
- *It's Cool to Be Kind*: Combatting Negative Online Behavior
- *When in Doubt, Talk It Out*: Questionable Content and Scenarios

The educator curriculum offers worksheets and extensive discussions on various aspects of internet safety, whereas the family version is a shorter, more digestible version that offers scenarios and activities families can talk through with their children to establish internet best practices. Both guidelines are free to download.

Common Sense Media. Common Sense Media is a nonprofit that reviews and rates media and technology for families. Along with reviewing new movies and tv shows, Common Sense Media offers a variety of resources for parents such as guides to understanding AI, tips for setting up parental controls across a variety of apps and websites, instructions for teaching children about data privacy, and more.[70] Common Sense Media's website also offers a 'Tips & FAQs' tab where parents can sort articles by topic and find useful advice from experts and parents alike.[71]

Family Online Safety Institute (FOSI). Another nonprofit, FOSI, works with industry and parents alike to develop and teach best practices for family internet use. FOSI offers online safety programs through their Digital Parenting Program that provides "expert-backed support, practical guidance, and easy how-to's on all things online safety."[72] The program is in partnership with the Digital Citizenship Initiative, created by Discovery Education, that provides free resources on teaching media literacy, digital footprints, digital safety, and the impact of the internet on the brain.[73] Parents using FOSI's website can sort information by device (e.g., laptop, smartphone, etc.), platform (e.g., Amazon, Google, YouTube, etc.), and topic (e.g., cyberbullying tips, education aides, etc.). Along with practical guidelines, FOSI offers blog posts about emerging topics in children's internet use. FOSI also has a research branch that provides detailed briefs and surveys of children and parents about a range of topics spanning from generative AI to the federal Children's Online Privacy Protection Act (COPPA).[74]

Third-party blogs. Several third-party parental control apps also offer blogs and websites with resources for parents. Aura, an app that monitors children's website and app activity for parents, started Digital Parenthood, a website that provides "resources, community, and conversations necessary to foster a safer, more connected generation, focusing on promoting online safety for kids."[75] The website hosts discussion forums about parenting in the age of the internet and even has an "ask an expert" tab that allows caregivers to post a question to be answered by one of the

[70] Common Sense Media, "Parents' Ultimate Guides," https://www.commonsensemedia.org/parents-ultimate-guides.
[71] Common Sense Media, "Parenting, Media, and Everything in Between," https://www.commonsensemedia.org/articles.
[72] Family Online Safety Institute, "Online Safety Programs," https://fosi.org/parenting/.
[73] Discovery Education, "Digital Citizenship Initiative," https://digitalcitizenship.discoveryeducation.com/.
[74] Family Online Safety Institute, "Research," https://fosi.org/research/.
[75] Aura Digital Parenthood, "About Us," https://www.digitalparenthood.com/category/welcome/discussions/about-us.

experts partnering with Aura.[76] Another parental control app, Bark, offers a blog that breaks down digital technology and trends for parents.[77] The blog posts cover topics like internet slang, common emojis, "rage-baiting" content, and how technology is changing classroom education.

ii. Youth-focused online spaces

Children-centered media entertainment has become a multi-billion-dollar industry.[78] With growing numbers of young children on social media and media platforms, some companies have turned toward making child-friendly apps and websites tailored for younger audiences.

Coverstar. Coverstar brands itself as "the safe TikTok alternative." The app is intended for children aged nine and up and has strong community guidelines that users must agree to when signing up, including not posting videos in underwear or bathing suits, along with not posting sexually explicit content. Any user under 13 must have guardian verification to create an account. Additionally, there are no direct messaging (DM) options, which is advertised on the app's website as a way of "protect[ing] users from predators and unwanted contact."[79] While comments are still allowed, every post is moderated by both AI-assisted technology and trained human moderators to prevent cyberbullying or negative comments. According to the app developers, Coverstar's feed is intended to be "positive and age-appropriate – not addictive."[80] Profiles can also be made private, which means only approved users can see the private-profiles' posts. According to Coverstar's Terms of Service, the app collects targeting and advertising cookies that "are used to deliver advertisements that are relevant to you based on your interests. They may be used to track your browsing behavior across different websites[.]"[81]

Coverstar promotes itself as the positive social media app, with the tagline "go viral, not toxic."[82] The app is primarily filled with tweens doing viral dances, lip syncing challenges, and other common internet trends. When an account is set up, the user is assigned a customizable avatar that represents them in the app. Users can change the avatar's appearance and purchase new clothing using Starcoins, which are a form of digital currency that must be purchased with real money. Gifts come in a range of forms and have different Starcoin costs associated with them. Gifts can then be given to users who go "Live" (a real-time video streaming option common to many social media platforms). Any gifts a live streamer – a verified user with 20,000 followers[83] – receives are then converted into real money, allowing users to profit from Coverstar.

---

[76] Aura Digital Parenthood, "Ask an Expert," https://www.digitalparenthood.com/category/get-help/discussions/ask-an-expert .
[77] Bark, "The Bark Blog," https://www.bark.us/blog/?srsltid=AfmBOopDoS55DVLgfgP5TFItUsfhHdnqKClCRLeyeU8kyK-D8mI6HMSq.
[78] Barbara Ortutay & Haleluya Hadero, "Social media companies made $11 billion in US ad revenue from minors, Harvard study finds," *Associated Press,* (Dec 27, 2023), https://apnews.com/article/tiktok-meta-instagram-revenue-teens-harvard-cc9bf875d6f7259ba2aee8805ccdaf3d.
[79] Coverstar, "The social platform for the next generation," https://coverstar.app/.
[80] *Ibid.*
[81] Coverstar Help Center, "Cookie Policy," https://help.coverstar.app/hc/en-us/articles/34168340613652-Cookie-Policy. For $7.99 a month, users can access the Premium app without ads.
[82] Apple App Store, "Coverstar – Positive Social," https://apps.apple.com/us/app/coverstar-positive-social/id1219890480.
[83] Coverstar Help Center, "How do I get verified?" https://help.coverstar.app/hc/en-us/articles/33465620090644-How-do-I-get-verified; Reddit, "Coverstar App: Recent Changes, Monetization & Information Not Available To

YouTube Kids. In 2015, the video streaming platform YouTube released YouTube Kids, a filtered version of YouTube that hosts family-friendly videos for children under 13. YouTube Kids uses three content settings: "Preschool," aimed at children four and under; "Younger," aimed at children aged five to eight; and "Older," aimed at children nine to twelve years old.[84] These settings determine what videos children will see across a range of filtered content such as sexual content, violence, weapons, dangerous content, language, health and beauty, sensitive topics, and music videos. For example, under the Preschool setting, YouTube Kids allows video with non-romantic expressions of love such as a kiss on the cheek or holding hands, whereas the Older setting allows content related to non-sexual romance like dating and first kisses, as well as age-appropriate videos on sex education like puberty.[85] Parents can set up individual accounts for each child. YouTube Kids also allows parents to control a variety of features on the app such as introducing time limits, turning off sound and background music, and removing the search feature.[86] These controls were expanded in 2018 to allow parents to limit content to human-reviewed channels as well as a system for manually whitelisting approved videos.[87] However, critics have noted that the app still allows product placement ads.[88]

iii. Third-party tools

With so many new apps and websites all vying for children's attention (and parents' money), third-party tools meant to provide parents both technical assistance and guidance have risen in popularity. Some of these tools block access to apps while others allow parents to see what their children are doing on their devices and to set limits on the kinds of content they can view. Many of these tools have a one-time or monthly subscription fee.

Brick. Launched in 2023, Brick is a physical device intended to block distracting apps and websites on phones. Purchasers can receive a small, gray square block (about a third the size of a cell phone) that will block any app or website they select. Brick operates through a specialized chip that works by tapping a phone to the Brick to activate the blocking. Users can set custom settings (e.g., study, workout) that can block different apps depending on the time of day. Unlike app limits or app blockers, the only way to end the blocks is to be in physical proximity to the

---

The Public (Live, Gifts, & Starcoins)," https://www.reddit.com/r/apps/comments/1nx42b7/coverstar_app_recent_changes_monetization/.

[84] YouTube For Families Help, "Content policies for YouTube Kids," https://support.google.com/youtubekids/answer/10938174#zippy=%2Cpreschool-content-setting%2Cyounger-content-setting%2Colder-content-setting

[85] *Ibid.*

[86] Shimrit Ben-Yair, "Introducing the newest member of our family, the YouTube Kids app—available on Google Play and the App Store," *YouTube Blog,* (Feb 23, 2015), https://blog.youtube/news-and-events/youtube-kids/

[87] Abner Li, "YouTube Kids now allows parents to whitelist any video or channel," *9to5 Google,* (Sep 13, 2018), https://9to5google.com/2018/09/13/youtube-kids-parents-approved/.

[88] Center for Digital Democracy, "FTC complaints filed by Campaign for a Commercial-Free Childhood and Center for Digital Democracy lead to major changes for kids on YouTube," (Oct 8, 2019), https://democraticmedia.org/filings/ftc-complaints-filed-campaign-commercial-free-childhood-and-center-digital-democracy-lead.

Brick or to use limited "emergency" unblocking measures.[89] Proponents argue Brick helps reduce the temptation to scroll mindlessly.[90]

Bark. Bark is an app that allows parents to monitor their child's activity on a range of apps (27 apps on Android, 15 apps on Apple) and allows parents to get alerts if their child's location moves.[91] Bark offers two versions of their app: Bark Jr. and Bark Premium. Bark Jr. allows parents to control screen time, filters websites and apps, and monitor for inappropriate content. Bark Premium adds social media, email, and YouTube monitoring. Bark also offers a Bark Phone that integrates the app's features into a child-friendly smartphone with automatic monitoring capabilities, giving families an option for greater controls over apps, social media, or internet browsers for their younger kids. Bark has also released the Bark Watch, a smartwatch with text monitoring and location tracking, and Bark Home, a device that monitors screen time, content, and apps for every device in the home.[92] Along with allowing parents to set time limits on apps and block apps or broad topics in general (i.e., sexual material), Bark monitors for problematic content such as bullying, sexual content, self-harm, drug or alcohol references, hate speech, etc. and sends an alert only if such content is detected.[93]

Aura. Another app aimed at monitoring children's device use, Aura lets parents monitor and limit their child's internet activities. Parents can set "bedtimes," which automatically disconnect their child's device from the internet at a set time, set time limits for apps and websites that are customizable for each family member, and can even pause the internet on their child's device at the touch of a button. Along with parental controls, Aura provides online history and usage reports that also claim to track shifts in social interactions, daytime activity, and sleeping habits.[94] As part of their digital wellbeing metrics, Aura assigns each child a social persona, one of six personas that highlight how a child behaves online based on their engagement patterns like session length, app diversity, and messaging frequency on social apps.[95] Aura claims to track overall patterns and not individual behaviors so that "[p]arents see changes in social engagement levels, not private conversations, ensuring parents are aware of important changes to spot early warning signs to check in, while kids maintain their privacy."[96] Aura also tracks AI apps and will notify parents if their child is starting a conversation with an AI app, as well as the number of

[89] Harry Rabinowitz, "We tested The Brick for two weeks. Did it free us from mindless scrolling?" *NBC News,* (Feb 19, 2026), https://www.nbcnews.com/select/shopping/brick-phone-app-blocker-review-rcna259740.
[90] *Ibid*; Laia Garcia-Furtado, "The Strong Case for Brick," *Vogue,* (Dec 24, 2025), https://www.vogue.com/article/you-really-should-have-used-brick#:~:text=The%20Brick%20is%20a%20small%2C%20square%20device,in%20order%20to%20access%20that%20content%20again; Elissa Sanci, "I Bricked My Phone for 2 Weeks. My Brain Feels Much Better." *WireCutter,* (Oct 17, 2025), https://www.nytimes.com/wirecutter/reviews/brick-phone-lock-review/.
[91] Mike Jennings, "Bark Review: Pros & Cons, Features, Ratings, Pricing, and more," *TechRadar,* (July 2, 2025), https://www.techradar.com/reviews/bark.
[92] *Ibid.*
[93] Anastasia Bukhtiarova, "Aura vs Bark (2026): features, pricing, and best parental controls choice for kids," *Cybernews,* (Feb 13, 2026), https://cybernews.com/best-parental-control-apps/bark-vs-aura/.
[94] Aura Parents, https://buy.aura.com/online-safety-and-wellbeing.
[95] Digital Parenthood, "From Sleep to Social: Measuring What Matters in Your Child's Digital Wellbeing," https://www.digitalparenthood.com/kb/tools-for-online-balance/from-sleep-to-social-measuring-what-matters-in-your-child%E2%80%99s-digital-wellbeing/252.
[96] *Ibid.*

messages sent. Included in this service are risk signals that flag potentially risky behaviors children are displaying with chatbots and alerts parents.[97]

Qustodio. Qustodio is a cross-platform parental control and digital wellbeing app used by over nine million parents worldwide. Qustodio allows parents to filter websites, block apps, set daily screen time limits, schedule device-free periods, monitor streaming activity, and track a child's GPS location with customizable geofencing alerts. AI-powered alerts notify parents when a child searches for content related to concerning topics such as self-harm or bullying. Premium tiers add social media monitoring, call and message tracking, and up to 30 days of detailed activity reports. A "family pause" feature cuts internet access across all connected devices instantly, while an "always allowed" setting keeps essential apps accessible even when restrictions are active. Qustodio's free tier covers only one device, however, and independent reviewers have noted that determined children can bypass its controls using a VPN.[98]

b. What's missing?

Data and transparency. Under current law, platforms face no obligation to disclose what content youth are exposed to, what harms result, or how their safety systems perform. The existing research base has been produced by academics with limited data access, advocacy organizations, and internal platform teams whose findings are selectively disclosed. Absent accurate and comprehensive data, it is difficult to fully understand the efficacy of online safety controls and policies. The AADC would have required such disclosures; however, as discussed above, these content-specific disclosures have been blocked from enforcement on First Amendment grounds. Nevertheless, there may be alternative mechanisms for ensuring standardized public reporting on underage user estimates, child safety incidents, algorithmic recommendation patterns for child-like accounts, parental control adoption rates, and response times for high-severity reports.

Independent evaluation. No independent body currently evaluates whether children's online safety systems work as advertised. Regulated industries such as the financial sector rely on third-party verification by entities with no financial stake in the outcome. Teams of approved researchers could be authorized to "red team" safety controls by testing accounts, probing age-verification and parental control systems for known circumvention methods, and assessing whether problematic content can be readily accessed. An independent evaluation body could conduct recurring, standardized usability audits of safety controls and publish the results. This would give parents comparable information across platforms and give companies competitive incentive to improve safety.

Universal controls. As discussed above, parents often navigate over 40 unique platforms to mediate their children's digital lives. No tool integrates across this landscape. A risky interaction may span game voice chat, Discord, and Instagram DMs, but each platform's controls observe

[97] Aura Parents, *supra*.
[98] Qustodio, "Features," *Qustodio.com*, https://www.qustodio.com/en/features/; Internet Matters, "What Is Qustodio? Guide for Parents," https://www.internetmatters.org/advice/apps-and-platforms/monitoring/qustodio/; Anastasia Bukhtiarova, "Qustodio Review (2026): Features, Pricing, & Real-World Testing," *Cybernews*, https://cybernews.com/best-parental-control-apps/qustodio-review/.

only its own slice. Legislation could require major platforms to provide standardized parental control application program interfaces,[99] enabling the use of cross-platform oversight tools.

Defaults. Making safety controls the default rather than requiring families to opt in to such protections would help ensure the controls are used. Opt-in safety controls create a regressive system: parents with more time, more digital literacy, and more resources are more likely to discover, configure, and maintain them. Parents working multiple jobs, with limited English, or who are less comfortable with technology are less likely to opt in. Default protections could include parental notification at account creation, content filtering calibrated to age brackets, AI-generated activity summaries for parents, screen-time awareness prompts, purchase and in-app spending alerts, and crisis detection with automatic references to resources.

Weekly Digest/Recap. Another potential tool for parents is a weekly high-level recap of what their kid saw and did, ideally in an easily digestible format similar to "year in review" recaps on Spotify, Instagram, or TikTok. Recaps could include topics seen, creators followed, sensitive topic percentage, and changes in consumption or behavioral patterns. These should be at a high level to respect child privacy, while giving parents an accurate sense of their child's overall online activity.

Even if this suite of potential controls is developed and implemented, however, it remains an open question whether they would suffice to ensure safety on platforms designed to keep kids engaged through features such as algorithmic feeds, endless scrolling, notifications, and likes. Safety experts contend that such features can be addictive to kids, undercutting the efficacy of even the most innovative of interventions.

VII. CONCLUSION

While layered online safety controls offer meaningful potential, it remains unclear whether families can realistically deploy them at scale across dozens of fragmented platforms, particularly when no independent body evaluates whether these tools work, and platforms face no obligation to disclose whether they do. As courts consider claims that product designs – such as algorithmic feeds, intermittent rewards, and sycophantic chatbots – drive compulsive use and harm adolescent mental health, governments worldwide, California included, are debating whether certain types of platforms should be off-limits to younger children, following Australia's model of social media age-gating. A question for this Committee to consider is whether, across the range of platforms where children spend their digital lives, families consistently have access to a reliable, effective, user-friendly set of safety controls – or whether stronger remedies are required.

[99] HR 2657 (Wasserman-Schultz) would require large social media platform providers to make available to third-party safety software providers a set of real-time application programming interfaces through which a child or a parent or legal guardian of a child may delegate permission to a third-party safety software provider to manage the online interactions, content, and account settings of such child on the large social media platform on the same terms as such child, and for other purposes. Text is available at https://www.congress.gov/bill/119th-congress/house-bill/2657/text.