TO: OfCom

FROM: Common Sense Media

DATE: 17 May, 2023

RE: Call for Evidence: Second phase of online safety regulation - Protection of children

Submission Details:

Full name: Jenna Khanna

Contact phone number: +44 (0)7738 016 969

Organisation name / representing: Common Sense Media

Email address: jkhanna@commonsense.org

Confidentiality

Your details: We will keep your contact number and email address confidential. Is there anything else you want to keep confidential?

No. The responses below are not confidential.

Your response: Please indicate how much of your response you want to keep confidential
N/A

For confidential responses, can Ofcom publish a reference to the contents of your response?
N/A

Responses are included below.

Question 1: To assist us in categorising responses, please provide a description of your organisation, service or interest in protection of children online.

Common Sense Media is an independent nonprofit organisation dedicated to helping children thrive in a rapidly changing world. The organisation is based in San Francisco with regional offices across the U.S. We launched our first international office in the UK in 2019, a registered charity (also Common Sense Media, 1188840), through which we rate, educate, investigate, and advocate for the well-being, safety, and privacy of children in the UK and EU. Our research and free resources provide parents, teachers, and policymakers with reliable, independent data on children's use of media and technology and the impact it has on their physical, emotional, social, and intellectual development.

Common Sense Media achieves its aim to create a more healthy, equitable, and empowering future for children in the following ways:
RATE: Through our parent platform, Common Sense Media, we provide independent ratings and reviews of various forms of media.
EDUCATE: Through Common Sense Education, we share our Digital Citizenship curriculum with millions of teachers across the world to increase media literacy and shape digital citizens, as well as offer tips to families and communities as they navigate media and technology.
ADVOCATE: We advocate and raise awareness to drive policy and industry changes that protect the safety, wellbeing, and privacy of children in the digital world, including the UK, EU, and USA.
INVESTIGATE: Our team conducts independent research about children's use of media and technology and its impact on their development.
We are greatly encouraged by the OfCom and the British government's efforts to address online safety and specifically to protect children. Thank you for this opportunity to respond to the second phase of online safety regulation consultation around protecting children from legal content that is harmful to them.

Common Sense Media was best equipped to submit responses to the following questions, which leverage our organisation's research, policy expertise, and evidence around protecting children online.

Question 2: Can you identify factors which might indicate that a service is likely to attract child users?

When we consider the factors that might attract child users to media platforms, it is important to examine the types of platforms children are using as well as the amount of time young people spend engaging with various types of screen media.

A nationally representative tracking survey[1] of 1,306 8- to 18-year-olds in the United States found that media use is rapidly growing with online videos reported as the top daily media activity among teens ages 13- to 18-years-old (77%) and second largest media activity for tweens ages 8- to 12-years old (64%). Additionally, watching online videos on sites such as YouTube or TikTok is the media activity young people enjoy doing the most with 61% of tweens and 62% of teens saying they enjoy watching videos online "a lot," far more than any other media activity. Furthermore, among the 79% of teens who report being regular users of both social media and online videos, 32% report that they "wouldn't want to live without" YouTube, highlighting the important role of online video sites in adolescent lives.

[1] The Common Sense Census: Media Use by Teens and Tweens, 2021.

There is no "one size fits all" for understanding why certain children are attracted to specific platforms or content. Rather, it is likely that children seek out content based on their specific interests and motivations. Examining the types of platforms that children are using as well as the amount of time they spend on and their enjoyment of these platforms is a necessary first step for understanding the role of specific factors in attracting children.

Question 3: What information do services have about the age of users on different platforms (including children)?

Firms collect billions of data points of personal information about users that go far beyond their birthdate, home address, and phone number. As soon as users are on a platform, it is recording their online activity: what they are searching, what websites they visit, what content they click on, how long they view the content, their location information, user demographics, login information, credit card purchase data, and more.[2] Some firms also track and analyse public record data from data brokers, such as court records, or records on births, marriage, divorce, and bankruptcy.[3] After collecting this data, firms apply predictive data analytics to draw inferences and create a profile on each user, including children.[4] For example, based on their online activity, a firm may profile a child as a gamer or impulsive purchaser.[5] With the use of machine learning algorithms, firms then curate each user's online experience by showing content the algorithm determines to be interesting or relevant based on the user's profile.[6]

This mass data collection and resulting profiling enables firms to infer the age of users. For example, if a user watches many toy unboxing videos or children's cartoons, the online service can infer that the user is likely a child. We have testified about how firms know there are children and teens on their platform, but turn a blind eye to them. In the most prominent example, YouTube claimed to the Federal Trade Commission (FTC) that their services are directed to a general audience, while touting to advertisers that they can target children. In 2019, YouTube settled with the FTC for collecting data in violation of COPPA.[7] The FTC had to waste time and money to get evidence confirming the obvious: YouTube knew kids were on its platform and collected information from them in violation of COPPA anyway.

Approaches to assessing the age of users are evolving, and online services have more options than just asking users to self-report their age. For example, they can use facial recognition technology like Yoti, which estimates age to verify whether a user is old enough to use the online service. Last summer, Instagram started testing Yoti as a method for users under 18 years old to verify their age.[8] However, firms do not need to use age verification mechanisms to

---

[2] Common Sense Media, Behavioral Advertising Harms: Kids and Teens (Feb. 2022).
[3] *Id.*
[4] *See* Common Sense Media Comments, Before the Federal Trade Commission on Trade Regulation Rule on Commercial Surveillance and Data Security 12 (Nov. 21, 2022).
[5] Common Sense Media, AdTech and Kids: Behavioural Ads Need a Time Out 5 (May 13, 2021).
[6] Common Sense Media Comments, Before the Federal Trade Commission on Trade Regulation Rule on Commercial Surveillance and Data Security 12.
[7] Press Release, Federal Trade Commission, Google and YouTube Will Pay Record $170 Million for Alleged Violations of Children's Privacy Law (Sept. 4, 2019).
[8] Press Release, Meta, Introducing New Ways to Verify Age on Instagram (June 23, 2022), https://about.fb.com/news/2022/06/new-ways-to-verify-age-on-instagram/. However, Instagram only required users who were trying to edit their birthdate from under the age of 18 to 18 and over to verify their age through using Yoti, uploading their ID, or asking a friend to vouch for them.

ascertain users' ages because of all the data they already hold about users that they can draw inferences from.

Question 4: How can services ensure that children cannot access a service, or a part of it?

To help ensure children cannot access the parts of the service where they are exposed to harmful content, online platforms directed at or likely to be accessed by children, such as social media platforms and video streaming platforms, should take a two-pronged approach to meet age-appropriate needs. They should (1) adopt data minimisation practices, and (2) design platforms with children and teens' well-being in mind, such as by ceasing use of design features that optimise young users' time on the platform.

Social media platforms have taken some steps to tailor their service to meet age-appropriate needs, but they have fallen short. For example, last summer, Instagram updated its Sensitive Content Control to default new users under 16 years old to the "less" sensitive control option.[9] Users will see less content the platform considers "sensitive" but does not violate their community guidelines, such as content that may depict violence, is sexually explicit or suggestive, or content that may be attempting to sell products or services based on health-related claims such as to help a person lose weight.[10] It is unclear whether some types of content, such as content promoting eating disorders, would be categorised as "sensitive." While this was a more substantial move that could help reduce the amount of harmful content young users will see, they can still change the default setting to see the regular amount, or even "more" sensitive content, making the setting less effective.

Our recommended two-prong approach would address many social media platforms' invasive data privacy practices and manipulative design features that lead children and teens to the parts of the service where users can receive harmful targeted ads or recommended content.

Data minimization, where companies only collect and use the data that is directly relevant and necessary to accomplish the specified purpose of providing the core features of the application or service, helps address the harmful targeted advertising and algorithmic amplification problem by cutting off the unrestricted access to users' data. Without so much data, companies will not be able to curate ads and content with as much precision to each individual user. This would help prevent young users from falling down rabbit holes where they consume a large amount of the same type of harmful content. For example, if companies cannot use online activity of users, such as what videos a user watches and for how long, to curate recommendations, a user will be less likely to start receiving an endless stream of videos promoting disordered eating simply because they watch one or two videos about how to eat healthy.[11]

In addition to adopting data minimization, companies should stop using design features that optimise a teen user's time on the platform, like autoplay and endless scroll. Such features make it easy for users to continually consume content. Facebook whistleblower Frances Haugen has explained that in the first ten minutes on Instagram, people will see content from

---

[9] Taylor Hatmaker, *Instagram's 'Sensitive Content' Controls Will Soon Filter All Recommended Content*, Tech Crunch (Jun. 7, 2022), https://techcrunch.com/2022/06/07/instagrams-sensitive-content-controls/.
[10] *Id.*
[11] *See* Avani Dias et. al, *The TikTok Spiral*, ABC News Australia (Jul. 25, 2021) (discussing how TikTok led a teen to develop an eating disorder).

their friends or pages they follow.[12] By the time a user is an hour or two into their Instagram session, Facebook's algorithm becomes the main thing choosing what the user sees, increasing the likelihood of getting exposed to harmful content.[13] This creates a feedback cycle where people stay on the platform to "self-soothe" and find the next good piece of content, but instead start to feel anxious.[14] Prohibiting companies from utilising design features like autoplay and endless scroll reduces a user's time on the platform, which helps reduce exposure to harmful content.

Online platforms likely to be accessed by children and teens should both adapt data minimization practices and stop utilising design features that optimise time to ensure children and teens receive a more age-appropriate experience where they are not exposed to so much harmful content.

Question 6: Can you provide any evidence relating to the presence of content that is harmful to children on user-to-user and search services?

A demographically representative survey[15] of 1,358 American teenagers (age 13 to 17), found that 73% of teenagers had consumed online pornography with most (54%) indicating that they first saw pornography when they were 13 or younger. Among teens who have intentionally viewed pornography, 38% report viewing pornography at least once a week on social media sites (e.g., TikTok, Instagram) and 34% report viewing weekly on video sites/platforms (e.g., YouTube). While 23% of all teens responding to the survey said they have accidentally seen pornography as a result of a friend or classmate they know having shown it to them, a majority of all teens (51%) said they have accidentally encountered pornography via clicking a link, a search engine result, an online ad, or on social media (note - no single social media platform stood out as an overwhelming source of pornography exposure). These findings suggest that online pornography is intentionally being accessed across a variety of platforms.

Concerningly, teens in this survey reported viewing harmful content at staggering rates. A majority of teens (52%) indicated that they have viewed pornography depicting what appears to be rape, choking, or someone in pain with most teens in the survey (55%) indicating that they encounter pornography portraying Black people in a stereotypical way "often" or "sometimes" and half indicating the same about Latino/a/x people (50%).

When taken together, we see that children are using a variety of platforms to access online pornography, both intentionally and accidentally, and they are being exposed to violent and harmful content.

Question 7: Can you provide any evidence relating to the impact on children from accessing content that is harmful to them?

---

[12] Allison S. Tate, *Facebook whistleblower Frances Haugen says parents make 1 big mistake with social media*, Yahoo (Feb. 7, 2022), https://www.yahoo.com/now/facebook-whistleblower-frances-haugen-says-000630484.html?soc_src=social-sh&soc_trk=ma.
[13] *Id.*
[14] *Id.*
[15] Michael B. Robb and Supreet Mann, Common Sense Media, *Teens and Pornography*, 2022.

There is evidence of both primary priority and priority content, as defined in the Online Safety Bill July 2022 Written Ministerial Statement, negatively impacting children.

I. Primary Priority Content

A. Pornography

The ubiquity of internet access on personal devices allows young people to have access to these online spaces in ways that can be difficult to monitor and mediate. Pornographic content can be easily accessible in these online spaces and on personal media devices, such as cell phones and laptops. A demographically representative Common Sense Media survey of 1,358 American teenagers (age 13 to 17) found that 73 percent of teenagers had consumed pornography, whether intentionally or accidentally.[16] Additionally, even for those teenagers who reported only consuming pornography accidentally, a majority (63 percent) reported they had been exposed to pornography in the past week, suggesting that pornography exposure is a common occurrence for most teenagers.[17]

Teens may turn to the internet for information about sexual behaviours or advice about sexual relationships. In the same survey, many teens indicated that they were learning about sex from the pornography they consumed, with 45 percent of teens agreeing that pornography gives them "helpful" information about sex.[18] This is concerning when we consider that 52 percent of teen respondents said they had seen violent or aggressive pornography, including media that depicts rape, choking, or someone in pain.[19]

B. Content Promoting Self-harm and Suicide

The harms of seeing content promoting self-harm and suicide online is tragically illustrated by the story of 14-year-old UK teen Molly Russell. In 2017, Molly killed herself after falling into a vortex of despair on social media the last year of her life.[20] An inquest into her life concluded that she died from "an act of self-harm while suffering from depression and the negative effects of online content."[21] Of 16,300 pieces of content Molly saved, liked, or shared on Instagram in the six months before she died, 2,100 were related to suicide, self-harm, and depression.[22] The more of this content she consumed, the more the algorithm fed her similar content. The content was so disturbing that at a hearing in a London coroner court, a consultant child psychiatrist said he could not sleep well for weeks after viewing the content Molly had seen right before her death.[23]

Molly's case is far from an isolated one, and many teens are fed this content on social media to the severe detriment to their mental health. A 2019 study found that people who saw self-harm

---

[16] Michael B. Robb and Supreet Mann, Common Sense Media, *Teens and Pornography* 3, 5 (February 2023).
[17] *Id.* at 6.
[18] *Id.*
[19] *Id.* at 7.
[20] John Naughton, *Molly Russell was Trapped by the Cruel Algorithms of Pinterest and InstagramI*, The Guardian (Oct. 1, 2022), https://www.theguardian.com/commentisfree/2022/oct/01/molly-russell-was-trapped-by-the-cruel-algorithms-of-pinterest-and-instagram.
[21] *Id.*
[22] *Id.*
[23] *Id.*

content on Instagram showed "more self-harm and suicidality-related outcomes."[24] Additionally, the study found that the majority of users seem to come across Instagram's self-harm content unintentionally.[25] Most participants who saw this content indicated that they felt emotionally disturbed.[26]

C. Content Promoting Eating Disorders

A report by children's advocacy watchdog group Fairplay showed that there is a pro-eating disorder bubble on Instagram which Meta knowingly profits from that includes 90,000 unique accounts that reach 20 million unique followers.[27] At least one-third of the followers in this bubble are underage.[28] They see content like accounts that celebrate images of extremely underweight people using terms like "thinspiration" or "bonespiration" and memes about having an eating disorder.[29] The 90,000 unique accounts does not include accounts that appear to focus on recovery, or general health awareness.[30]

Consuming this kind of content has driven girls to developing eating disorders.[31] Many young people in the pro-eating disorder bubble state in their bios that they want to recover, but remain stuck in the algorithm's bubble and continue to receive such recommendations.[32]

Instagram is not the only platform pushing youth to eating disorder content. For example, after watching one video by a fitness influencer on TikTok and following her, a teen user named Lauren's feed became dominated by content focused on keeping up a so-called “healthy” lifestyle that pushed her to the dangerous trend of meticulously tracking how many calories they eat.[33] She went from feeling positively about her body, to crying about it every night after watching videos of people saying they hated their body. Four months later, she was diagnosed with an eating disorder. A user's feed can become overwhelmed with this kind of content extremely quickly. When an Australian researcher Dr. Suku Sukunesan started engaging with eating disorder content on TikTok, after a couple hours, the platform suggested 30 different accounts for him to follow, all people living with eating disorder issues.[34]

---

[24] Florian Arendt, Sebastian Scherr, and Daniel Romer, *Effects of Exposure to Self-Harm on Social Media: Evidence From a Two-Wave Study Among Young Adults*, 21 NEW MEDIA & SOCIETY 2422, 2436 (2019). This study sampled participants aged 18 years or older. It is important to consider that teen users are even more likely to use social media, and also more vulnerable to the effects of exposure to harmful content because their brains are still developing.
[25] *Id.* at 2429.
[26] *Id.* at 2430.
[27] Fairplay, Designing for Disorder: Instagram's Pro-eating Disorder Bubble.
[28] *Id.*
[29] *Id.*
[30] *Id.*
[31] Fairplay's report gives a first hand account of how a 17-year-old teen girl named Kelsey's former eating disorder and obsession with weight loss was fueled by social media. *Id.* at 6-7.
[32] *Id.* at 12.
[33] Avani Dias, *The TikTok spiral*, ABC News Australia (Jul. 25, 2021).
[34] *Id.*

In the United States, the number of hospitalizations for eating disorders doubled during the pandemic.[35] Similar trends have been observed in Western Australia,[36] New Zealand,[37] and Canada.[38] While there is less data from Europe, representatives in four out of six European hospitals in a research study described a higher rate of overall admissions for anorexia nervosa during the pandemic.[39] Clinicians also experienced an increase in the severity of symptoms of anorexia compared to before the pandemic.[40] They cited increased exposure to triggering social media as one of the explanations for the rise in eating disorder symptoms, as well as general psychopathology.[41]

Overall, using social media platforms has been associated with body image issues.[42] For example, frequently comparing one's own physical appearance to that of people followed on social media is associated with body dissatisfaction and a drive for thinness.[43]

II. Priority Content

A. Online Abuse, Cyberbullying, and Harassment

The anonymity of being online can expose people, especially children and teens, to online abuse, cyberbullying, and harassment. Common Sense's research has found that children and teens are exposed to these risks in virtual reality (VR) platforms.[44] The anonymity of virtual avatars on VR platforms facilitates the trolling that afflicts social media and enables the mass coordination of harassment campaigns.[45] For example, virtual troll armies are easy to coordinate in private VR rooms, and their threats feel more intimidating because a user's eyes and ears are engulfed with a headset to create a more immersive experience.[46]

Young users also regularly come across sexual content and abuse like virtual strip clubs, sexual grooming, simulated sex acts, and rape threats.[47] Sexual assault, especially in the form of

---

[35] David A. Asch et al., *Trends in US Patients Receiving Care for Eating Disorders and Other Common Behavioral Health Conditions Before and During the COVID-19 Pandemic*, 4 JAMA NETWORKING OPEN 11, 2 (2021).
[36] Yasheer V. Haripersad et al., Outbreak of Anorexia Nervosa Admissions During the COVID-19 Pandemic, 106 ARCHIVES OF DISEASE IN CHILDHOOD 1 (2021).
[37] Sara J. Hansen, David B. Menkes, and Alice Stephen, *The Impact of COVID-19 on Eating Disorder Referrals and Admissions in Waikato, New Zealand*, 9 J. EATING DISORDERS 1 (2021).
[38] Holly Agostino et al., *Trends in the Incidence of New-Onset Anorexia Nervosa and Atypical Anorexia Nervosa Among Youth During the COVID-19 Pandemic in Canada*, 4 JAMA NETWORK OPEN 1 (2021).
[39] Josefina Castro-Fornieles et al., *Increase in Admission Rates and Symptom Severity of Childhood and Adolescent Anorexia Nervosa in Europe During the COVID-19 Pandemic: Data From Specialized Eating Disorder Units in Different European Countries*, 16 CHILD & ADOLESCENT PSYCHIATRY MENTAL HEALTH 46, 3 (2022).
[40] *Id.* at 5.
[41] *Id.* Other possible explanations included the interruption in regular sports activities, too much spare time, lack of social contacts, and a reduction in treatment offers.
[42] *See e.g.* Melanie Duval et al., *Social Media Use and Body Image Disorders: Association Between the Frequency of Comparing One's Own Physical Appearance to That of People Being Followed on Social Media and Body Dissatisfaction and Drive for Thinness*, 18 INT'L J. ENV'T RSCH. PUB. HEALTH 6, 10 (2021).
[43] *Id.*
[44] Katie Joseff and Nelson Reed, What are Kids Doing in the Metaverse?, Common Sense (Mar. 23, 2022), https://www.commonsensemedia.org/kids-action/articles/what-are-kids-doing-in-the-metaverse.
[45] *Id.* at 8.
[46] *Id.*
[47] *Id.* at 10.

groping, is already seen in VR.[48] The immersive nature of VR gives this abuse the potential to be more traumatic than when it occurs in other online formats.[49]

In addition to the specifically outlined categories of priority content, our report also details other harms children and teens experience on VR platforms. Primarily, children and teens face major privacy violations.[50] Spending just 20 minutes in a VR enables the platform to collect just under two million unique recordings of a user's body language.[51] VR devices contain cameras and sensors that constantly record user movements, tracking body movements 90 times per second.[52] They collect sensitive biometric information like facial and eye movements, which can be used to reveal their behaviour and desires.[53] For example, eye tracking and pupil dilation can potentially signal personality traits, cultural affiliation, skills, preferences and aversions.[54] This gives companies even more invasive personal information they can use or sell to other advertisers to use to target users with ads or content, which furthers a person's potential exposure to harmful content.[55] Additionally, overuse and overexposure to VR can lead to addiction and problematic internet use, aggression, and dissociation.[56]

B. Content Depicting or Encouraging Violence

Many dangerous challenges like the "blackout challenge," where people choke themselves until they pass out on camera, have become viral online, which has resulted in grave physical harm or death to the children and teens who try them.[57] The blackout challenge, which went viral on TikTok, has been linked to the deaths of at least 15 kids aged 12 or younger since it reemerged in 2021.[58] This challenge, also known as the "choking game," dates back to at least 2008, when 82 children died trying to record themselves doing the challenge.[59] Most of the kids that died that year were between the ages of 11 and 16, from across 31 states.[60]

Although platforms do not allow content that encourages dangerous or illegal activities, when you search for terms like "blackout challenge," you can still easily find examples of them online. Numerous other dangerous challenges have become viral, such as ones in which people "dry scoop" pre-workout powder without water, climb tall stacks of milk creates, and eat massive

---

[48] *Id.*
[49] *Id.*
[50] *See id.* at 6–8.
[51] *Id.* at 6.
[52] *Id.*
[53] *Id.* at 7.
[54] *Id.*
[55] *Id.*
[56] *Id.* at 11.
[57] Fairplay, Dared by the Algorithm: Dangerous Challenges Are Just a Click Away (Sept. 29, 2022), https://fairplayforkids.org/wp-content/uploads/2022/09/Dangerous-Challenges.pdf.
[58] Olivia Carville, *TikTok's Viral Challenges Keep Luring Young Kids to Their Deaths*, Bloomberg Businessweek (Nov. 30, 2022), https://www.bloomberg.com/news/features/2022-11-30/is-tiktok-responsible-if-kids-die-doing-dangerous-viral-challenges.
[59] Press Release, Center for Disease Control, CDC Study Warns of Deaths Due to the "Choking Game" (Feb. 14, 2008), https://www.cdc.gov/media/pressrel/2008/r080214.htm.
[60] *Id.*

amounts of frozen honey and corn syrup.[61] These challenges have resulted in physical harm and death as well. For example, dry scooping led a 20-year-old woman to be rushed to the hospital for heart attack symptoms.[62]

C. Differences in the Impact of Priority Content on Children in Different Age Groups

Although we are not aware of research on the impact of categories of priority content on children in different age groups, overall, the younger a child is, the more vulnerable they are to online harm because their brain is still developing. The brain structures of adults and adolescents are very different, which causes children and teens to respond to stimuli differently from adults.[63] The limbic system and the prefrontal cortex of our brains grow synchronously, but at different speeds.[64] The limbic system is associated with survival and contains the part of our brain that controls certain emotional responses such as our "fight or flight" response.[65] Meanwhile, the prefrontal cortex is associated with higher-level functions such as planning, problem solving, reasoning, and impulse control, and will not mature until closer to adulthood.[66] Before the prefrontal cortex is fully matured and able to counterbalance the limbic system, children and teens are less equipped than adults to make rational decisions, consider long term consequences, and control impulses.[67] Their critical thinking skills are developing, making them more vulnerable and largely defenceless against advanced and personalised techniques like targeted advertising.[68] Consequently, while all children and teens as well as adults can face harm to their physical and mental health from being exposed to priority content such as online harassment, younger children are more subject to harmful impact.

Question 11: What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements for children (including children of different ages)?

---

[61] Addison Aloian, Sarah Felbin, and Sabrina Talbert, *The "Blackout Challenge" Has Resurfaced on TikTok, And It's Still Just as Dangerous As it Was 17 Years Ago*, Women's Health (Jan. 17, 2023), https://www.womenshealthmag.com/health/a38603617/blackout-challenge-tiktok-2021/.

[62] Kate Dwyer, *What is Dry Scooping? Why Experts Warn Against Trying This Viral TikTok Trend*, Women's Health (June 8, 2021), https://www.womenshealthmag.com/health/a36632539/what-is-tiktok-dry-scooping-trend/.

[63] *See* Brief for Common Sense Media and Frances Haugen as Amici Curiae Supporting Petitioners 6–9, Gonzales v. Google, __ U.S. __ (2022) (No. 21-1333) (discussing the structural disparities between the adult and adolescent brain and how they lead adults and adolescents to respond to stimuli differently).

[64] B.J. Casey et al., *The Adolescent Brain*, 28 Developmental Rev. 62, 63 (2008).

[65] *See* Velayudhan Rajmohan and Eladath Mohandas, *The Limbic System*, 49 Indian J. of Psychiatry 132–39 (2007) (providing an overview of the components and functions of the limbic system).

[66] Edward E. Smith and John Jonides, *Storage and Executive Processes in the Frontal Lobes*, 283 Science 1657, 1659–60 (1999).

[67] *See* Angela Griffin, *Adolescent Neurological Development and Implications for Health and Well-Being*, 5 Healthcare 62, 63 (2017) (describing how the prefrontal cortex is late-evolving and enables individuals to learn how to manage long term planning, monitor what is going on, and adjusting smoothly to surroundings while keeping emotions and behaviors context-appropriate).

[68] Common Sense Media, AdTech and Kids: Behavioral Ads Need a Time-Out (May 13, 2021).

We do not believe current privacy policies or terms of service for applications or services used by children or teens enhance clarity and accessibility without accountability.

Just as we regulate our food system to protect the public's health and safety of food products in the grocery store, we also need to urgently protect the privacy of kids and families. The Common Sense Privacy Program reads and evaluates the privacy practices of popular applications and services used by children in order to provide evidence of a company's privacy practices to help parents and children make better informed decisions.[69]

Enhanced clarity and accessibility of the privacy practices of an application or service can help children, teens, and their parents make better informed decisions about the apps they use everyday and hopefully choose better privacy protecting products. If children under the age of 13, or teens younger than 18, use an application or service, the company should disclose how they better protect the privacy of children and teens in their privacy policy. In addition, if an application or service has worse privacy practices for adult users or consumers, such as selling their data to third parties for profit, or displaying targeted advertising that could inadvertently harm children, then the company should disclose additional protections are in place to protect children or teens from these practices by default. Moreover, even if an application or service is not intended to be used by children or teens, the product may still appeal to children and therefore companies should still disclose in their privacy policy how they will handle information inadvertently collected from children or teens with stronger privacy protections.

I. 2021 State of Kids' Privacy Research

Our research into the privacy practices of the most popular applications and services used by children and teens found that the majority of applications and services do not transparently disclose clarity or accessibility of stronger privacy protections for children.[70] In addition, our evidence shows among companies with mixed audience products intended for both children and adult users, we see increased transparently with updates to privacy policies that carve out exceptions to prohibit selling children's data, not displaying targeted ads to children, and not tracking child users of the product when the company has actual knowledge the user is a child. However, approximately half of all companies in 2021 likely avoid obtaining actual knowledge of whether a user is a child under 13 years of age through the product's experience with an age-gate or required birth date, which can lead to inadvertently exposing children using these products to data monetization practices that are intended to only apply to teen and adult users.[71]

Rather, companies likely have constructive knowledge that children under 13 are using their products — information that a company is presumed to have, regardless of whether or not they actually do. If a product has features such as child profiles, content directed to children such as cartoons, or interactions clearly intended for children or that would likely appeal to children under 13 years of age, then companies should know children are using the respective product and should be required in place stronger privacy protections.

Our research into the state of kids' privacy also examined whether given a company's policies disclose the product is intended for children, does the company also disclose the qualitatively "worse" practice that they display targeted advertisements. The evidence indicates approximately 95% transparency for products that explicitly disclose whether or not children are

[69] Common Sense Media, Privacy Program, https://privacy.commonsense.org.
[70] Kelly, G., Graham, J., Bronfman, J., & Garton, S. (2021). 2021 State of Kids' Privacy. San Francisco, CA: Common Sense, https://www.commonsensemedia.org/research/state-of-kids-privacy-report-2021.
[71] Id. at 51-58.

intended users. For the products intended for children, this transparency results in a split between "better" 52% and "worse" 37% practices, which indicate that data collected from users of the product may be used to target advertising to kids.[72] In addition, for the 62 products evaluated in our report that disclose they are not intended for children, but are still used by children everyday, the data indicates that nearly 60% of products have "worse" practices that use personal information to display targeted advertising to other users of the product, who could include consumers, parents, and educators.[73]

However, for products where the policies disclose children are the intended audience and also display targeted advertising, it may be the case that these companies are limiting display of targeted ads to only adult users of the product or anonymous users which often include children. Mixed-audience products often allow children to create accounts without indicating their age with age-gates or other birth date verification systems – which would inadvertently expose children to targeted advertising practices unless the company has actual knowledge the user is a child under 13 years of age and prevents displaying targeted advertising to child users of their product.

II. Privacy of Streaming Apps and Devices Research

Additional research into the privacy practices of the most popular streaming media services used by children, such as Netflix, Disney, Hulu and others has found evidence that stronger privacy protections for children are not disclosed in privacy policies, even when child profiles or age-appropriate content moderation features are available.[74] Streaming apps and devices with kid and family directed content should minimally include child profiles or child accounts to provide a safer experience with age-appropriate content recommendations and stronger privacy practices that protect children's data when they are using the streaming app or device. Additional privacy protections that apply to children's data when using separate child profiles also need to be clearly communicated to parents with a separate child privacy policy that explains what stronger privacy protecting practices are actually in place when children are using the streaming app or device.[75]

To enhance clarity and accessibility of privacy practices for children, adequate privacy protections for children require a separate child profile and child privacy policy that clarifies different data collection and use practices are in place for child accounts. Children should be prevented from viewing or experiencing child-directed content through an adult account or device without mechanisms in place to provide notice and direct parents to create a safer child account or profile with stronger privacy protections. However, none of the most popular streaming apps and devices we evaluated in our research report provided a separate child profile with stronger privacy practices for children across all evaluation criteria.[76] If the most popular media streaming applications with children do not disclose how they protect children with stronger privacy protections, or engage in worse privacy practices for adults and with child-directed content in adult accounts, but do not disclose how they protect kids, then there can be no meaningful clarity and accessibility of privacy for children.

---

[72] Id. at 55.
[73] Id. at 55-56.
[74] Kelly, G., Graham, J., Bronfman, J., & Garton, S. (2021). Privacy of Streaming Apps and Devices: Watching TV that Watches Us. San Francisco, CA: Common Sense Media, https://www.commonsensemedia.org/research/privacy-of-streaming-apps-and-devices-watching-tv-that-watches-us.
[75] Id. at 18.
[76] Id. at 20.

III. Privacy Misinformation

To enhance clarity and accessibility of a product's privacy practices, some companies have published separate "privacy principles" or "privacy centre" webpages to help summarise their privacy practices, but often these privacy summary websites do not actually disclose any of the "worse" privacy practices they engage in with children, teens, or adults' data.[77] This can create a false sense of safety for children and parents who believe these products are more privacy protecting then they actually are. In addition, Apple's recent introduction of its App Store "privacy nutrition privacy label" and Google's new "Data Safety section" in the Play Store for app developers have attempted to redefine what "privacy" means to consumers in the App Stores for hundreds of millions of users to improve clarity and accessibility.

Common Sense reads and rates the privacy policies of apps in both the Apple and Google App Stores and is therefore able to also validate a company's privacy information displayed to consumers in App Store labels and product pages. However, privacy and safety labels are self-reported by app developers without validation by Apple or Google. Recent research has found evidence of apps in App Stores, and elsewhere, with worse privacy practices for children such as selling data to third parties and targeted ads that use misinformation in their privacy labels claiming they have safer privacy practices then what their privacy policies state.[78] This misinformation practice is spreading with upwards of 60% of popular apps by children and teens in both App Stores currently displaying false privacy information to parents and children which completely negates any clarity or accessibility of a product's privacy practices.[79] Privacy misinformation is inherently unfair and deceptive because it is misleading parents and consumers to download and purchase apps for themselves and their children that they believe to be healthier and more privacy-protecting and therefore are more likely to use.

IV. Holding Companies Accountable

Parents desperately want to know which apps are better for privacy, but now they have an additional challenge of identifying which apps also have false privacy information. In order to help enhance clarity and accessibility of privacy practices, the Common Sense privacy ratings help parents and children navigate the complexity of an app's privacy practices and our additional verification of privacy labels can help provide more confidence to consumers.[80] However, companies need to be held accountable when their "privacy principles" or "privacy centre" webpages do not reflect both their better *and* worse practices. Companies also need to be held accountable when they market themselves as privacy protecting in App Stores to increase the likelihood of more downloads, but their privacy policies transparently disclose inconsistent worse privacy practices, or disclose no stronger protections at all for children. The problems of how to enhance clarity and accessibility of privacy practices for children are complex and evolving, but the solutions require companies to be transparent and honest about

---

[77] *See* Microsoft Privacy Principles, https://privacy.microsoft.com/en-US ("Microsoft's privacy center does not disclose its"worse " privacy practices of selling data, displaying targeted advertisements, or tracking users across the internet and over time for commercial purposes.")

[78] Anne Stopper and Jen Caltrider, Mozilla, No Evil: Loopholes in Google's Data Safety Labels Keep Companies in the Clear and Consumers in the Dark, https://foundation.mozilla.org/en/campaigns/googles-data-safety-labels.

[79] Li, Y., Chen, D., Li, Tianshi, Agarwal, Y., Cranor, L., & Hong, J.I. (2022, April 28). Understanding iOS privacy nutrition labels: An exploratory large-scale analysis of app store data. CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts). https://doi.org/10.1145/3491101.3519739.

[80] Common Sense Privacy Program, Privacy Ratings, https://privacy.commonsense.org/resource/privacy-ratings.

their privacy practices. Companies also need to be held accountable for privacy misinformation so parents and children can make better informed decisions about the apps and services they use everyday.

Question 15: What actions do or should services take in response to reports or complaints about online content harmful to children (including complaints from children)?

Online services should provide a meaningful response to each report or complaint. However, Ofcom should put in place a process by which firms are held accountable.

The US Consumer Financial Protection Bureau (CFPB), the regulator responsible for the safety of consumer financial products and services, has put in place a complaint submission and response process to ensure that firms are held accountable. To be sure, many of these firms have their own department that handles complaints, but it is far more effective when the regulator is also involved in the complaint process.

The CFPB's Consumer Response (CR) division is responsible for handling complaints submitted by consumers about a firm and ensuring that consumers receive a meaningful response.[81] Once a consumer submits a complaint through the regulator's website, CR routes that complaint to the appropriate entity. The company then has a given period of time to respond depending on certain factors. Once the company responds, the CFPB publishes the complaint[82] along with the firm's response in the agency's Consumer Complaint Database, which is publicly available.

This process helps to address those instances where firms fail to respond to direct complaints. Further, the CFPB can monitor whether firms are in fact providing meaningful responses to consumer complaints and identify those firms that are routinely providing inadequate responses. The CFPB can use its supervisory authority to initiate an examination or its enforcement authority to investigate any wrongdoing. Finally, this process serves as a public shaming tool because anyone can access it, from reporters in search of a story to potential consumers who place a high value on customer service.

Question 17: To what extent does or can a service adopt functionalities or features, designed to mitigate the risk or impact of content that is harmful to children on that service?

Currently, most online services are not developed or designed in a way that is safe for children, and instead prioritise company profits and engagement first. Online services demonstrate this

[81] *See* 12 U.S.C. 5534(b) Pursuant to Title X of the Dodd-Frank Wall Street Reform and Consumer Protection Act, the CFPB has the authority to require covered firms to explain:

- The steps that were taken to respond to the consumer complaint;
- Any response that consumer provided; and
- Any follow-up actions or planned follow-up actions that will be taken to further respond to the complaint.

[82] CFPB publishes only those complaints that consumers have affirmatively provided their consent to publish.

through their mass data collection, the use of targeted advertising and algorithmic recommendations, and the use of design features that keep users glued onto their platform. However, there are many functionalities and features an online service can adopt to mitigate the risk or impact of content that is harmful to children.

Social media platforms have taken some steps to keep children safe, but they have largely relied on offering parents and teens themselves more control rather than changing their own practices. These controls have focused on giving parents the ability to see who their children are interacting with and how much time they are spending on platforms.[83] Platforms have also rolled out features young users can use, such as Instagram's "take a break" and "nudge" features. Young users under a certain age (typically 16 years old) who sign up for platforms like TikTok and Instagram are also now defaulted to private accounts.[84] While these can be helpful tools, this approach is insufficient and ineffective because the parent or teen must choose to utilise these controls, and they do not address the root cause of online harms like invasive data privacy practices or algorithmic amplification.

To address invasive data privacy practices, online services should have the strongest, most privacy-protective settings be the default for all minor accounts. This ensures every minor user has a baseline level of privacy protection without requiring them to try to change the often difficult-to-navigate privacy settings of all the online services and apps they use. Companies would still be able to collect the information they strictly need for the online service to function. Coupled with data minimization principles, as we propose in the response to question 4, companies would have less data about individual users to recommend harmful content and target them with ads.

Online services should also stop utilising design features like endless scroll and autoplay that keep users on their platform and increase their chances of seeing harmful content and becoming addicted. Former employees at tech companies have spoken out on how social media companies are deliberately addicting users to their products to make profit.[85] For example, Aza Raskin, the former Mozilla and Jawbone engineer who designed endless scroll, has said it does not give users' brains time to catch up with their impulses, leading them to keep scrolling.[86] Raskin did not set out to addict people and now feels guilty about the creation, but said many designers were driven to create addictive app features because of the business models of large companies that try to get more funding and keep their stock prices up.[87] Similarly, autoplay, where videos start playing automatically as soon as they appear on screen, can keep users on a platform longer than they intend to by making them start to watch a video,

---

[83] Julie Jargon, *How to Use Parental Controls on YouTube, TikTok, Instagram, and Snapchat*, Wall Street Journal (Apr. 16, 2022), https://www.wsj.com/articles/how-to-use-parental-controls-on-youtube-tiktok-instagram-and-snapchat-11650065233 (detailing the parental controls YouTube, TikTok, Instagram, and Snapchat offer, and how to use them).

[84] Eric Han, *Strengthening Privacy and Safety for Youth on TikTok*, TikTok (Jan. 13, 2021), https://newsroom.tiktok.com/en-us/strengthening-privacy-and-safety-for-youth; Sarah Perez, *Instagram Now Defaults New Users Under 16 to Most Restrictive Content Setting, Adds Prompts for Existing Teens*, Tech Crunch (Aug. 25, 2022), https://techcrunch.com/2022/08/25/instagram-now-defaults-new-users-under-16-to-most-restrictive-content-setting-adds-prompts-for-existing-teens/.

[85] *See* Hilary Andersson, *Social Media Apps are "Deliberately" Addictive to Users*, BBC News (July 4, 2018), https://www.bbc.com/news/technology-44640959.

[86] *Id.*

[87] *Id.*

even if they did not want to. At a very minimum, minors' accounts should have these features turned off by default. While this is not as effective as a direct prohibition, it better ensures less young users have these features enabled.

Companies are in a position to adopt functionalities and features that will mitigate the risk or impact of content that is harmful to children, demonstrated by the intentional design choices they currently make to increase their engagement on the platform. It is past time for companies to start prioritising young users' health and well-being before their profits.

Question 22: How are human moderators used to identify and assess content that is harmful to children?

I. The Role of Content Moderators

The typical experience of a human content moderator who identifies and assesses whether content is harmful to children and teens has transformed over recent years. Moderators work in modern day sweatshops where tens of thousands of outsourced contractors across the globe work tirelessly to police the worst content on the Internet. Moderators are expected to evaluate user generated content in only a few seconds against the rules of a company's content policy. In many respects, social media companies would not be able to function without human content moderators making hundreds of individual content decisions each day about whether to keep or remove user generated content that is published and shared to millions of users of online social media services such as Facebook.[88] The problem and scale of moderating hundreds of millions of unique individual content types each day in dozens of languages is mind bogglingly complex. Content moderation requires contractors to consistently make objective decisions about what content should remain published or be deleted with arbitrary and often changing rules that serve only as a floor, not a ceiling, or even best practices to protect children.[89] Content moderators serve as a type of online first responder who work to protect users–especially the most vulnerable users such as children--from the spread of harmful content shared by others. In their role, moderators are exposed daily to a wide range of images, text, and video content including harmful content to children such as hate speech, violent imagery, sexual abuse, rape, murder, and porngraphy.

Moderators are required to quickly decide whether each content type uploaded each day should be removed for violating the company's community standards or policies that are complex and require language fluency, cultural nuance, context of speech, and can vary from company to company. Companies use Artificial Intelligence content-based algorithms to detect and delete most of the obvious and most extreme terrorist or violent content. However, these filters can’t stop all harmful content and therefore social media companies need human content moderation to fill the gap and manually view and evaluate everything else Artificial Intelligence cannot determine is harmful or not. Social media companies could simply block or reject all other content that AI filters miss by erring on the side of blocking more content to protect their users, but the current content moderation business model intentionally chooses to only remove the minimum amount of potential harmful content necessary to create the illusion of a positive social media experience for the majority of its users. By design, the content moderation business

[88] Jason Koebler & Joseph Cox, The Impossible Job: Inside Facebook’s Struggle to Moderate Two Billion People, Motherboard, (Aug. 23, 2018) https://www.vice.com/en/article/xwk9zd/how-facebook-content-moderation-works.
[89] Facebook Community Standards, https://transparency.fb.com/policies/community-standards.

model serves to keep published the maximum amount of content possible for social media companies to monetize user engagement for profit, while erring on the side of showing harmful content to kids and teens. This business model needs to change.

II. The Risks and Harms of Content Moderation

Investigations into the workplace of content moderators in the United States have found evidence that many contractors earn close to minimum wage and are required to work in a high-stakes environment that demands near-perfect accuracy of determining whether content should remain published or be removed while being subjected continuously to violent imagery and hate speech.[90] Content moderators subjected to harmful content on a daily basis have stated they suffer from mental illness, started to believe many of the conspiracy theories they are exposed to everyday, and even experience PTSD after they leave the job.[91] Since investigations and evidence was found into the poor working conditions and risks to the mental health of moderators, a prominent content moderating service company Cognizant, has since left the industry and other content moderation companies have limited the scope of their contracts with social media companies citing liability risks.[92] To minimise potential liability operating in the United States and high turnover, several content moderation outsourcing companies have moved to other countries to exploit weaker labour laws, lower minimum wages, and high unemployment to find more contractors. A recent investigation found evidence of poor working conditions with a content moderation company in Kenya that paid contractors less than 2 dollars per hour to moderate content for several social media companies.[93] Facebook and other social media companies' content filtering Artificial Intelligence algorithms are decades away from achieving the necessary level of sophistication to moderate all potential harmful content automatically. Therefore human content moderators are still critically important to allowing social media companies to function, but poor working conditions and low wages have resulted in contractors submitting open letters to Facebook demanding equality and change.[94]

III. Content Moderation Includes Artificial Intelligence (AI)

Content moderation has continued to evolve in recent years with new innovations with natural language Machine Learning models that need moderators to identify and assess harmful content. Artificial Intelligence (AI) services have become popular recently with consumers and families that include helpful answers to spoken or written questions from Alexa, Siri, Google,

---

[90] Casey Newton, The Trama Floor, The secret lives of Facebook moderators in America, The Verge, (Feb 25, 2019),
https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.

[91] Alex Castro, Facebook will pay $52 million in settlement with moderators who developed PTSD on the job, The Verge, (May 12, 2020)
https://www.theverge.com/2020/5/12/21255870/facebook-content-moderator-settlement-scola-ptsd-mental-health.

[92] Alex Castro, A Facebook content moderation vendor is quitting the business after two Verge investigations, The Verge, (Oct. 30, 2019),
https://www.theverge.com/2019/10/30/20940956/cognizant-facebook-content-moderation-exit-business-conditions-investigation; See Adam Satariano and Mike Isaac, The Silent Partner Cleaning Up Facebook for $500 Million a Year, New York Times, (Aug. 31, 2021),
https://www.nytimes.com/2021/08/31/technology/facebook-accenture-content-moderation.html.

[93] Billy Perrigo, Inside Facebook's African Sweatshop,Time, (Feb. 17, 2022),
https://time.com/6147458/facebook-africa-content-moderation-employee-treatment.

[94] Open letter from content moderators re: pandemic, (Nov. 18, 2020)
https://www.foxglove.org.uk/2020/11/18/open-letter-from-content-moderators-re-pandemic.

and now Dall-E[95] and ChatGPT.[96] These AI services refer to themselves as Machine Learning or "ethical AI" services but similar to social media companies they still require human content moderators to clean-up the content used to train the AI systems, which include hate speech and violent imagery. An investigation into content moderation of the popular AI service ChatGPT found evidence of similar poor working conditions as social media companies, with contractors paid 2 dollars per hour in other countries around the world to moderate examples of violence, hate speech, and sexual abuse to teach AI examples of toxic langauge.[97]

IV. The Content Moderation Model Needs to Change

The bottom line is more content moderation by humans does not mean children will be more protected online from harmful content, but rather more humans will need to be exploited in order for social media companies and now AI services to make more profit and claim their products are safe to use by children.[98] The content moderation business model needs to be reexamined in the context of protecting the emotional and mental health of content moderators and paying them actual living wages, or even hazard-related compensation, rather than exploiting their individual vulnerabilities and country's poor economy. Ultimately this is a complex social, ethical, and legal issue we need to confront as a society. We need to decide whether the role of content moderation is too hazardous to contractors' mental health compared to its marginal benefits of more user engagement to justify its unregulated expansion and exploitation of human beings for profit.

In conclusion, less content moderation by humans will mean less potentially harmful content will be approved for publication and sharing on social networks which means harmful content will be accessible to children. Without more content moderation by more humans, less overall content will be published and shared on social media and used by AI services by default. This will likely result in less user engagement, less data monetization, and less profit for social media and content sharing companies. These are difficult issues that require complex solutions which need to be balanced against the beneficial use of technology for children to communicate, collaborate, grow, and share content safely on the Internet without exposure to harmful content.

[95] Open AI, Dall-e-2, https://openai.com/dall-e-2.
[96] Open AI, ChatGPT, https://openai.com/blog/chatgpt.
[97] Billy Perrigo, Exclusive: OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic,Time, (Jan. 18, 2023), https://time.com/6247678/openai-chatgpt-kenya-workers.
[98] Vittoria Elliot, Meta Lurches Toward Another Moderation Crisis, Wired, (Jan. 23, 2023), https://www.wired.com/story/metas-new-moderation-contractor-may-be-worse-than-its-last-one.