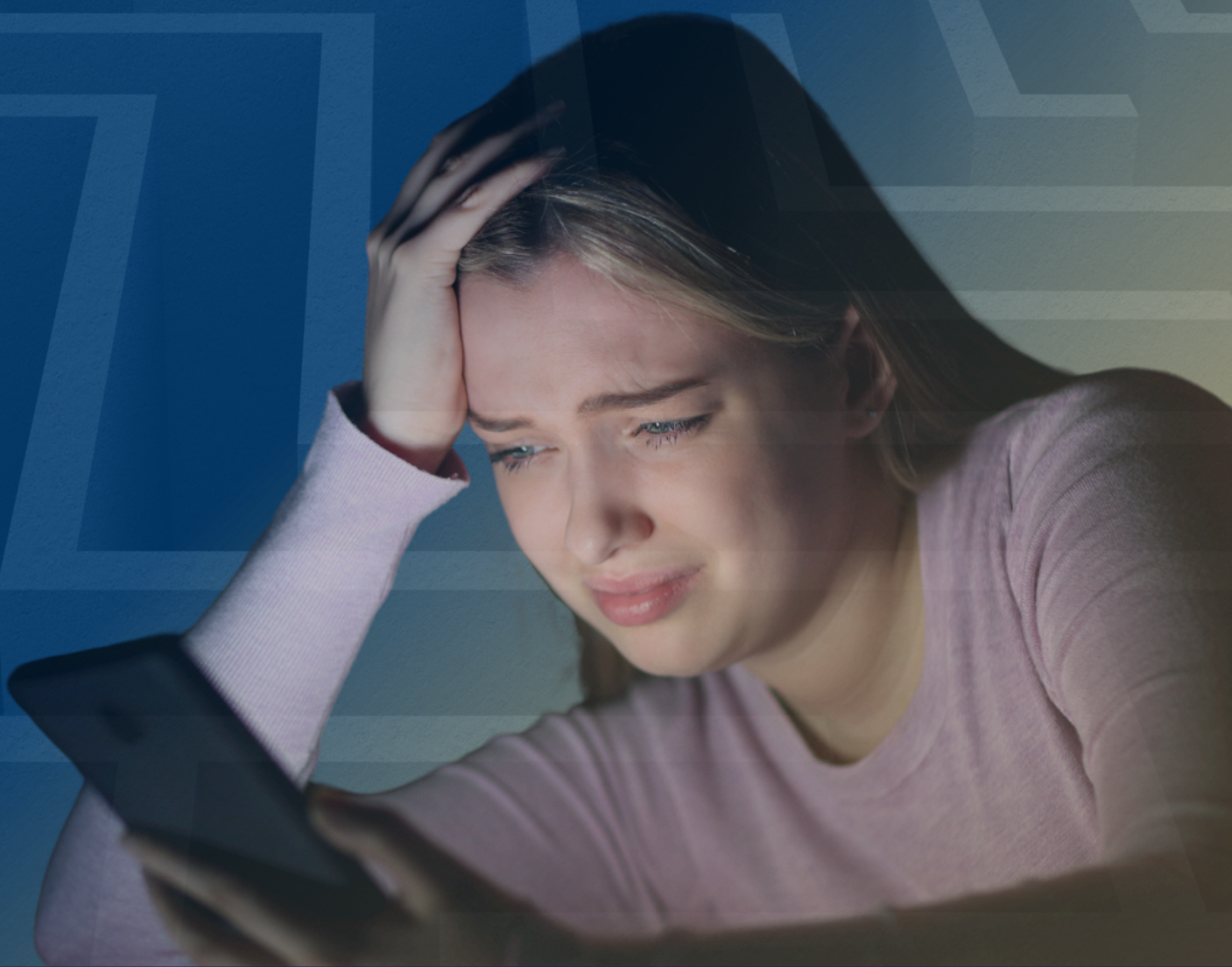


# CHILD CRUELTY BY DESIGN:

*Big Tech's Defiance of California's  
Child Online Safety Mandates*



---

# ABOUT THE CHILDREN'S ADVOCACY INSTITUTE

---

The Children's Advocacy Institute (CAI) was established in 1989 at the University of San Diego School of Law. CAI is an academic, research, and advocacy organization with the mission to strengthen the rights and improve the health, safety, and well-being of all children and youth.

CAI works to achieve its mission by engaging in the academic and clinical training of law students in child advocacy; providing continuing education for attorneys on child- and youth-related areas of practice; conducting research on child-related issues; engaging in public education about the status of children and of the performance of local, state and federal governments to advance their interests; and directly advocating before courts, agencies, and legislatures to seek leveraged results for the benefit of children and youth.

To support CAI's work, please visit [our website](#). Contributions to CAI are tax-deductible to the extent the law allows.

San Diego Office:  
University of San Diego School of Law  
5998 Alcalá Park / San Diego, CA 92110-2429  
(619) 260-4806 / (619) 260-4753 (fax)

Sacramento Office: (916) 844-5646

Washington, D.C. Office: (917) 371-5191

Email: [info@caichildlaw.org](mailto:info@caichildlaw.org) / Website: [www.caichildlaw.org](http://www.caichildlaw.org)



CHILDREN'S ADVOCACY INSTITUTE  
UNIVERSITY OF SAN DIEGO SCHOOL OF LAW

---

# TABLE OF CONTENTS

<b>EXECUTIVE SUMMARY</b>	i
<b>INTRODUCTION: EVALUATING HOW SOCIAL MEDIA PLATFORMS HAVE COMPLIED WITH CALIFORNIA LAWS AIDING SEXUALLY ABUSED AND CYBERBULLIED CHILDREN</b>	01
<b>BACKGROUND</b>	02
Child Sex Abuse Material and Social Media	02
<b>INSTAGRAM (META)</b>	04
<b>FACEBOOK (META)</b>	05
<b>TIKTOK</b>	06
<b>CYBERBULLYING AND SOCIAL MEDIA</b>	07
<b>CALIFORNIA’S LEGISLATURE TAKES ACTION</b>	11
AB 1394 Required Platforms to Create a Reasonably Accessible Mechanism for Reporting CSAM	11
SB 1504 Required Platforms to Create a Prominent Mechanism for Reporting Cyberbullying	12
<b>THE CHILDREN’S ADVOCACY INSTITUTE’S EVALUATION</b>	13
<b>USER EXPERIENCE AB 1394</b>	14
Instagram (Meta)	14
Instagram Summary	16
Facebook (Meta)	17
Facebook Summary	20
SnapChat	21
SnapChat Summary	24
TikTok	25
TikTok Summary	30
<b>FINAL OBSERVATION REGARDING AB 1394</b>	31
<b>USER EXPERIENCE SB 1504</b>	32
Instagram (Meta)	32
Instagram Summary	34
Facebook (Meta)	35
Facebook Summary	38
SnapChat	39
SnapChat Summary	42
TikTok	43
TikTok Summary	45
<b>FINAL OBSERVATION REGARDING SB 1504</b>	46



## EXECUTIVE SUMMARY

This report by the Children’s Advocacy Institute evaluates how major social media platforms have responded to California’s AB 1394 and SB 1504—two landmark laws aimed at protecting children from online sexual exploitation and cyberbullying. AB 1394 requires platforms to establish “reasonably accessible” mechanisms for reporting child sexual abuse material (CSAM), while SB 1504 mandates “prominent” pathways for reporting cyberbullying. These laws reflect California’s bipartisan demand for greater accountability from platforms whose business models have demonstrably harmed youth.

The evaluation covers four major platforms—Instagram, Facebook, Snapchat, and TikTok—and measures their compliance through hands-on testing of their in-app reporting systems on mobile devices. Despite being granted an extra year to prepare for compliance, none of the platforms studied has implemented reporting mechanisms that meet even the most basic standards of accessibility, clarity, or support for survivors.

Key findings include:

- Buried and misleading reporting tools: Reporting pathways are often hidden behind multiple menus, vague terminology, and unclear steps—requiring 5 to 7 clicks or more.
- Confusing language: Platforms use euphemisms like “nudes” instead of more precise and understandable terms, creating confusion and deterring reporting.
- No legal transparency: None of the reporting systems clearly inform users of their rights under California law or the platforms’ obligations to respond.
- Designs that retraumatize: Young survivors must navigate confusing interfaces, contradictory instructions, and character-limited text fields while seeking help—experiences that compound trauma rather than alleviate it.

The report concludes that the failures documented are not due to technological limitations or oversight. These corporations—some of the most profitable and technologically advanced in the world—have deliberately chosen profit over protection. Their actions, or inaction, defy the intent of the law and disrespect both the California Legislature and the vulnerable children these laws were designed to protect.

Immediate regulatory enforcement and stronger legislative action are warranted to ensure that platforms cannot continue to operate in willful noncompliance with laws meant to safeguard children.

# INTRODUCTION: EVALUATING HOW SOCIAL MEDIA PLATFORMS HAVE COMPLIED WITH CALIFORNIA LAWS AIDING SEXUALLY ABUSED AND CYBERBULLIED CHILDREN

*For decades, Big Tobacco lied to the public and targeted children to fuel addiction and profits, but the tactics of today's social media platforms pose a danger that is more insidious and far-reaching. These companies are knowingly exploiting vulnerable children to maximize engagement and revenue, undermining their mental health, development, and sense of self on a generational scale. They are monetizing the erosion of childhood, causing record-shattering numbers of child suicides, depression, and eating disorders, and with the rise of AI, it is only going to get worse.*

*– Jessica Heldman, Executive Director, Children's Advocacy Institute*

In response to widespread, bipartisan outrage over the harms social media platforms are knowingly inflicting on an entire generation of young people, California has enacted a series of landmark laws in recent years.

Three of these laws—Assembly Bills 1394 (Wicks and Flora) and 2481 (Lowenthal) and Senate Bill 1504 (Stern)—require platforms to create mechanisms that allow certain individuals to report content that either violates state law or the platform's own terms of service, particularly when that content poses risks to children. Two of these bills (AB 1394 and SB 1504) have been in effect since January 1, 2025. AB 2481 goes into effect in just a few months, on January 1, 2026. Notably, the effective date for AB 1394, which in part requires platforms to establish mechanisms for platform users to report child sex abuse material (CSAM), was delayed by a year at the request of the platforms, to give them a full year to alter their interfaces to aid sexually abused children.

When it comes to the laws enacted by AB 1394 and SB 1504, this report asks ***whether some of the world's most profitable corporations that are also the world's experts on crafting user interfaces have designed their reporting mechanisms to be easy for victimized children and those adults aiding them to use or whether they have thwarted the intent of these laws by making reporting needlessly obscure and frustrating.***



## BACKGROUND

### CHILD SEX ABUSE MATERIAL AND SOCIAL MEDIA

California Civil Code § 3273.65 defines CSAM on a social media platform as child pornography or obscene matter that depicts a minor personally engaging in, or personally simulating, sexual conduct. CSAM — said more plainly, images and video portraying a child being victimized — on a social media platform is not just a depiction of a past crime. It is an ongoing assault on a child.

Online CSAM has proliferated since the advent of social media. For example, during the time that Instagram went from one million users to one billion, there was “a 9,000% jump in abuse images online, according to the National Center for Missing and Exploited Children (NCMEC).”<sup>[1]</sup> A recent NCMEC report revealed that there were more than 20.5 million reports of online child exploitation received by its Cyber Tipline in 2024, representing 29.2 million separate incidents. These incidents include online enticement of children for sex acts (e.g., sextortion); child sex trafficking; possession, manufacture, and distribution of child pornography; and child molestation. Reports of these incidents contained 62.9 million images, videos, and other files related to the exploitation.<sup>[2]</sup>

It is challenging in a public document to describe how horrifying these videos and images can be:

The images and videos that are reported are not merely sexually suggestive or older teenagers who “look young.” This content depicts crime scene activity. Children — including those who are too young to call for help — are raped, abused, and exploited in this imagery. ... They are revictimized every time a sexually abusive image or video in which they are depicted is traded online and a new predator takes personal gratification in their anguish or uses the imagery to entice another child into sexual abuse.<sup>[3]</sup>

Social media is a primary gateway for child sex trafficking. As noted by FBI Special Agent Michael Syrax, “I think the biggest risk [of child sex trafficking] is through social media ... Some of these children are recruited through Facebook, Instagram, and other social media outlets, much in the same way that a person would interact with them in real life.” Research supports the opinion of FBI Agent Syrax. The Human Trafficking Institute reported that in 2020, “the internet was the most common ... location for recruitment.”<sup>[4]</sup> Furthermore, the

Institute found that 65% of child sex trafficking victims recruited through social media were recruited through Facebook, with Instagram and Snapchat accounting for another 22% combined.<sup>[5]</sup> TikTok has also been documented as a venue where predators use livestreams to exploit minors, often rewarding them with digital gifts or off-platform payments.<sup>[6]</sup>

According to the Human Trafficking Institute, “Traffickers may create, encourage the creation of, upload, or transfer child sexual abuse materials as part of their trafficking scheme.”<sup>[7]</sup> Among the child victims of CSAM, 67% are prepubescent or younger, including infants and toddlers.<sup>[8]</sup>

Appallingly, the trend in recent years appears to be toward more violent and extreme material.<sup>[9]</sup> Although girls are more likely to be depicted in CSAM, research shows that boys are more likely to be depicted in the most violent images. Furthermore, queer youth are uniquely at risk. A survey by the child safety non-profit Thorn revealed that “[n]early 1 in 3 LGBTQ+ minors (32%), reported an online sexual encounter with someone they believed to be over 18, ten percentage points higher than their non-LGBTQ+ peers (22%).”<sup>[10]</sup>

The insidiousness of CSAM lies in its permanence, especially when it is online. As a recent Department of Justice report summarized, “[w]hen these images and videos are posted and disseminated online, the victimization continues in perpetuity. Children often suffer a lifetime of re-victimization, knowing the documentation of their sexual abuse is on the internet, available for others to access forever.”<sup>[11]</sup> For survivors of child sex abuse, the abuse never really ends because, “[o]nce an image [of their abuse] is on the Internet, it is irretrievable and can continue to circulate forever.” The child is re-victimized as the images are viewed again and again.<sup>[12]</sup>

Survivor accounts and research confirm that the psychological toll of CSAM online is not only severe but also enduring. Survivors report feelings of helplessness, violation, and permanent terror at the loss of control over their own bodies and identities. For many, the trauma of knowing the material remains publicly available can be nearly as devastating as the original abuse. As one survivor described in a 2020 report:

I spend hours every day searching for my own content, reporting thousands of accounts and posts sharing CSAM. When platforms don't actively look for or prevent this content from being uploaded, the burden falls on me to have these images removed.<sup>[13]</sup>

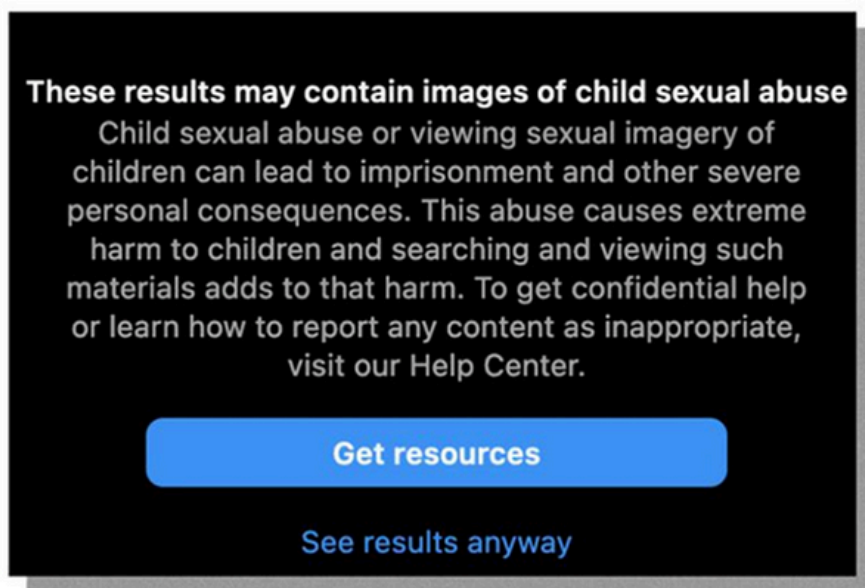


This is made worse because criminals often purposefully produce CSAM where children are seen smiling, leading survivors to worry that others will assume their enjoyment or implicate them in the abuse. Survivors report that perhaps the most difficult part of their re-victimization is the knowledge that their images may be used to groom future victims as a way to normalize the abusive behavior.<sup>[14]</sup>

Unfortunately, not all victims are survivors. Since 2021, NCMEC has received information about more than three dozen teen boys who have committed suicide as a result of being a victim of sextortion, in which a perpetrator extorts a youth after coercing them to produce sexually explicit content and threatening them with sharing the images or video.<sup>[15]</sup>

Platforms unquestionably know how to find and block CSAM content. They elect not to. No better proof exists than from Instagram, which, until shamed by public disclosure, screened for and identified CSAM but allowed pedophile users to “See Results Anyway.”

## INSTAGRAM (META)



This graphic is from a June 7, 2023, article in *The Wall Street Journal* titled “Instagram Connects Vast Pedophile Network.” The story documented the findings of a mere three academics at Stanford University and the University of Massachusetts Amherst working part-time with only the public’s access to Instagram:

*Though out of sight for most on the platform, the sexualized accounts on Instagram are brazen about their interest. The researchers found that Instagram enabled people to search explicit hashtags such as #pedowhore and #preteensex and connected them to accounts that used the terms to advertise child-sex material for sale. Such accounts often claim to be run by the children themselves and use overtly sexual handles incorporating words such as “little slut for you.”*

*Instagram accounts offering to sell illicit sex material generally don't publish it openly, instead posting "menus" of content. Certain accounts invite buyers to commission specific acts. Some menus include prices for videos of children harming themselves and "imagery of the minor performing sexual acts with animals," researchers at the Stanford Internet Observatory found. At the right price, children are available for in-person "meet ups."*<sup>[16]</sup>

*The Journal* concluded: "Instagram doesn't merely host these activities. Its algorithms promote them. Instagram connects pedophiles and guides them to content sellers via recommendation systems that excel at linking those who share niche interests[.]"<sup>[17]</sup>

Compare this "See Results Anyway" corporate culture to trying to upload a copyrighted song to YouTube. Fearing copyright liability, the platform effectively screens for and blocks not only images *but also sounds* from being uploaded. This Instagram "See Results Anyway" window underscores that platforms know how to find and intercept the most horrific content that clearly violates their terms of service. But even a \$5 billion fine assessed by the FTC against Meta was, according to that agency, too puny to motivate the company to comply with child-protecting orders.<sup>[18]</sup>

## FACEBOOK (META)

Facebook has known about child sex traffickers openly using its products since at least 2018, leaked documents show. It was so egregious that in 2019, Apple threatened to pull Facebook and Instagram's access to the App Store, a platform on which the social media giant relies to reach hundreds of millions of users annually.<sup>[19]</sup> A CNN report stated that: "A report distributed internally [within Facebook] in January 2020 found that 'our platform enables all three stages of the human exploitation lifecycle (recruitment, facilitation, exploitation) via complex real-world networks[.]'"<sup>[20]</sup>

In March 2022, a college professor described in *WIRED* magazine how searching for "Facebook groups with names including 10, 11, or 12" concerning "the 10th, 11th, or 12th wards of the city of Pittsburgh" instead revealed to her dozens of "groups targeting children of those ages" with "over 81,000 members" openly soliciting children for sexual exploitation. One 9,000-member group appearing in the search results was named "Buscando novi@ de 9,10,11,12,13 años"—i.e., "[l]ooking for a 9, 10, 11, 12, 13-year-old girlfriend." It gets worse. When the professor "used Facebook's on-platform system" to report this group, an "automated response came back" stating "[t]he group had been reviewed and did not violate any 'specific community



standards.” And, despite (or because of) her reporting this group, along with others, Facebook’s AI algorithms caused “new child sexualization groups” to be “recommended to [her] as ‘Groups You May Like.’”<sup>[21]</sup>

The *WIRED* revelations are not new. In March 2020, the Tech Transparency Project published an analysis that found hundreds of U.S. cases in which suspected pedophiles used Facebook to groom minors and trade images of their sexual abuse.<sup>[22]</sup> The report further concluded that Facebook’s systems are failing to do the bare minimum needed to reduce such abuse.

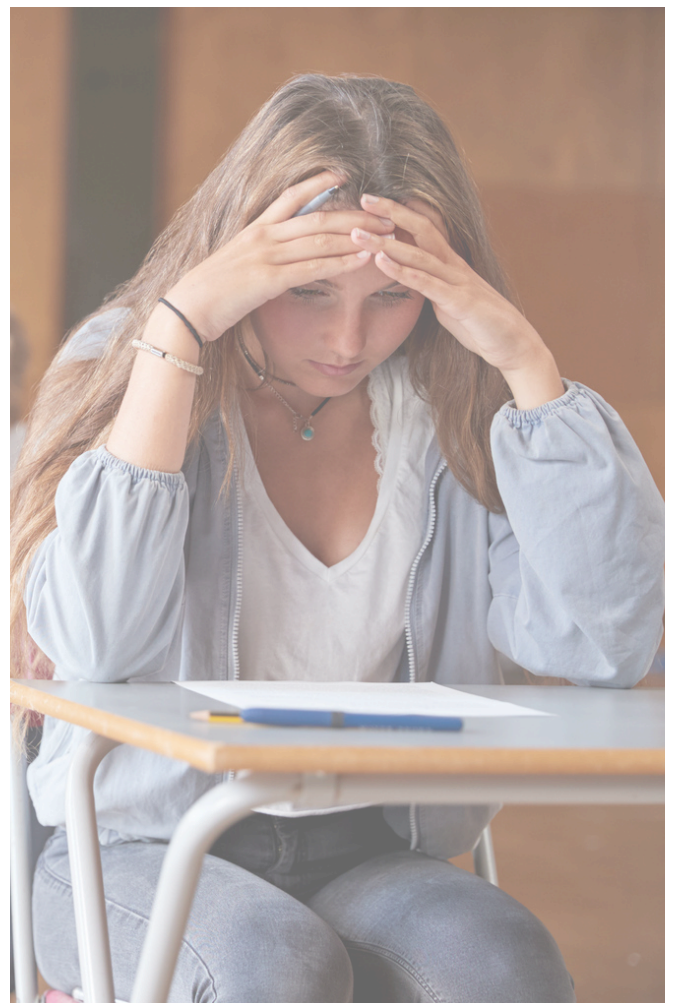
Furthermore, in the vast majority of the criminal cases of child sex trafficking and exploitation involving Facebook, Facebook was not the source of the tip-off to law enforcement about such heinous crimes occurring on its own platform, on its own supposed watch. Tips instead came from the public or other investigations.<sup>[23]</sup> The Tech Transparency Project reported that “[o]nly 9% of the cases [involving Facebook] were initiated because Facebook or the National Center for Missing and Exploited Children (which receives cyber tips from Facebook) reported them to authorities, raising questions about the effectiveness of Facebook’s monitoring of criminal activity targeting children.”<sup>[24]</sup>

## TIKTOK

Multiple investigative reports have documented how TikTok permits users to urge children to commit sexual or sexualized acts. For example, in 2022, a *Forbes* review of hundreds of recent TikTok livestreams reveals how viewers regularly use the comments to urge young girls to perform acts that appear to toe the line of child pornography — rewarding those who oblige with TikTok gifts, which can be redeemed for money, or off-platform

payments to Venmo, PayPal or Cash App accounts that users list in their TikTok profiles.

“It’s the digital equivalent of going down the street to a strip club filled with 15-year-olds,” says Leah Plunkett, an assistant dean at Harvard Law School and faculty associate at Harvard’s Berkman Klein Center for Internet & Society, focused on youth and media. Imagine a local joint putting a bunch of minors on a stage before a live adult audience that is actively giving them money to perform whatever G, PG or PG-13 activities they request, she said. “That is sexual exploitation. But that’s exactly what TikTok is doing here.”<sup>[25]</sup>



## CYBERBULLYING AND SOCIAL MEDIA

As documented by The Wall Street Journal,<sup>[26]</sup> Arturo Bejar, a consultant formerly on Instagram's Well-Being team "with unusual access to top executives" testified before Congress about the knowing, stubborn failure of Instagram to take basic steps to protect its child users from cyberbullying, including sexual harassment cyberbullying. His whistleblowing included many internal documents like the email excerpt below, affirming, based on Instagram's own secret research, the alarming breadth and consistency of bullying and both (i) the impenetrability of Instagram's system for receiving complaints and (ii) the unresponsiveness of the giant company when complaints are filed. For example, this email affirms that more than 21% of users 13 to 15 years of age reported they "were subject to bullying":

From: Arturo [REDACTED]  
**Subject: Gap in our understanding of harm and bad experiences**  
 Date: October 5, 2021 at 9:37:59 PM PDT  
 To: Mark Zuckerberg [REDACTED]  
 Cc: Sheryl Sandberg [REDACTED], Chris Cox [REDACTED], Adam Mosseri [REDACTED], Mark Zuckerberg [REDACTED]

Dear Mark,

I saw the note you shared today after the testimony, and I wanted to bring to your attention what I believe is a critical gap in how we as a company approach harm, and how the people we serve experience it. I've raised this to Chris, Sheryl, and Adam in the last couple of weeks.

I want to start by saying that my personal experience, and what I believe, is that you and m-team care deeply about everyone we serve, and my goal in sending this is to be of service to that. It's been 2 years since I've been back part-time.

**51% of Instagram users say 'yes' to having had a bad or harmful experience in the last 7 days. Out of those 1% of report and of those 2% have the content taken down (i.e. 0.02%). The numbers are probably similar on Facebook.**

Two weeks ago my daughter [REDACTED], 16, and an experimenting creator on Instagram, made a post about cars, and someone commented 'Get back to the kitchen.' It was deeply upsetting to her. At the same time the comment is far from being policy violating, and our tools of blocking or deleting mean that this person will go to other profiles and continue to spread misogyny. I don't think policy/reporting or having more content review are the solutions.

There is detailed data about what people experience in TRIPS, a statistically significant survey. We ran a more detailed survey, I've attached the full age breakdown below, **but here are some key numbers (these questions are in the last 7 days):**

**21.8% of 13-15 year olds said they were the target of bullying.**  
 39.4% of 13-15 year olds said they experienced negative comparison.  
 24.4% of 13-15 year old responded said they received unwanted advances.



It is worth noting how the rampant cyberbullying generally and sexual harassment cyberbullying specifically of children came to the attention of the whistleblower. According to *The Journal*:

Having joined the company long before its initial public offering, [Bejar] had the resources to spend the next few years on hobbies—including restoring vintage cars with his 14-year-old daughter, who documented her new pastime on Instagram.

That’s when the trouble began. A girl restoring old cars drew plenty of good attention on the platform—and some real creeps, such as the guy who told her that the only reason people watched her videos was “because you’ve got tits.”

“Please don’t talk about my underage tits,” Bejar’s daughter shot back before reporting his comment to Instagram. A few days later, the platform got back to her: The insult didn’t violate its community guidelines.

Bejar was floored—all the more so when he learned that virtually all of his daughter’s friends had been subjected to similar harassment.

Another leaked document — unsurprisingly — reveals the yearning of users for a safer experience online:

Hate speech, divisive civic content, and graphic violence are frequently and intensely experienced, and have been shown to have a negative effect on sentiment and legitimacy, particularly with repeated exposures over time.

\*\*\*

Users want Facebook to act. They hold us responsible for negative experience, and most think Facebook should automatically remove severe integrity-related content and hide less severe content. They perceive exposure to integrity harms as worse than false positive actions on benign posts.

\*\*\*

. . . lack of transparency & understanding of ranking & enforcement.<sup>[27]</sup>

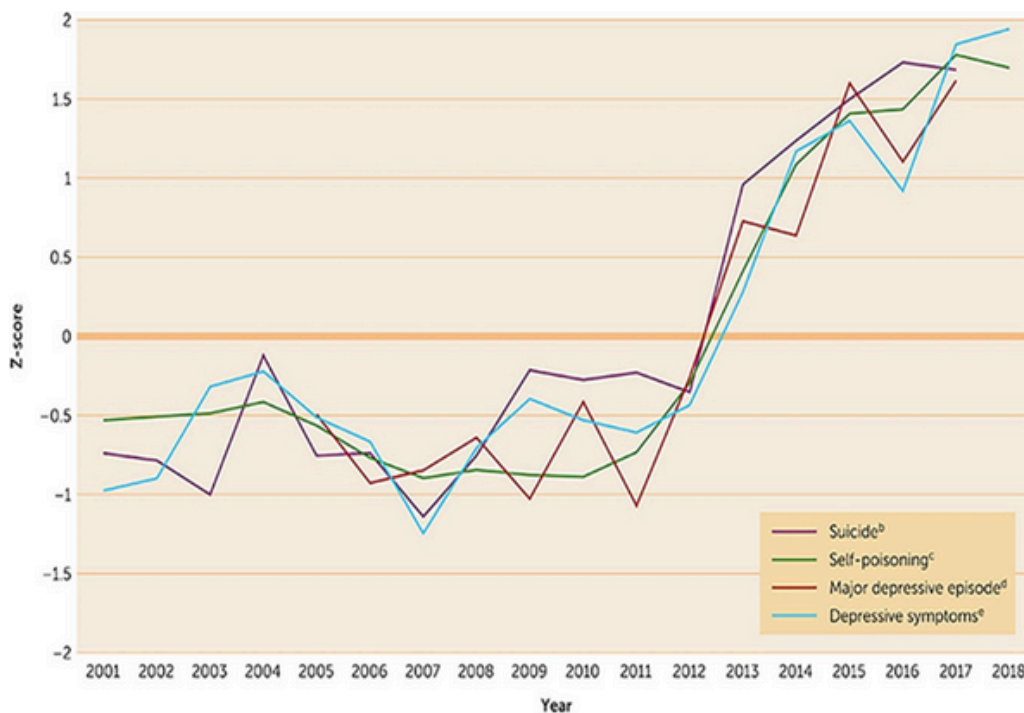
The harm associated with cyberbullying can be severe. As California Attorney General Rob Bonta has correctly observed, unlawful “cyberbullying can destroy a young life. It takes the worst of youthful cruelty and puts it on that most public of forums – the Internet. Too many American young people keep quiet about online abuse. And too many kill themselves over it.”<sup>[28]</sup> He adds:

As many as 56 percent of teens report being cyberbullied, and certain groups, such as lesbian, gay, bisexual and transgender teens, are targeted more than others. Teenagers who are cyberbullied are more likely to struggle with depression and substance abuse. They are at a higher risk offline to be victims of sexual harassment and physical assault.<sup>[29]</sup>

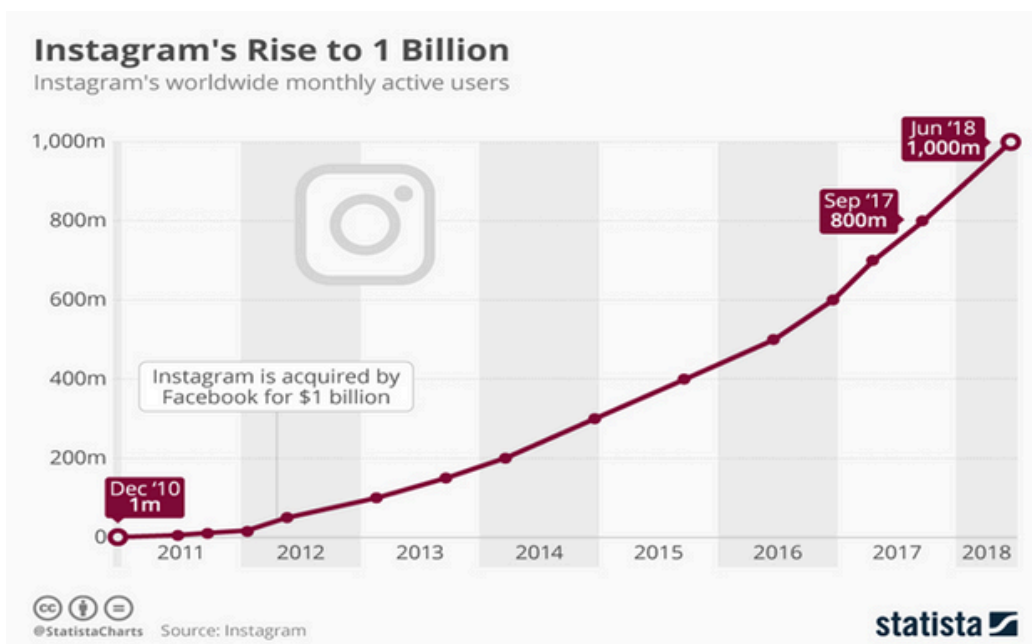
Black or Hispanic teens are more likely than White teens to say cyberbullying is a major problem for people their age.<sup>[30]</sup>

Relatedly, suicides, self-harm, and major depression are spiking in ways never before seen, especially among teen girls, and two graphs show why.

**Increases in Depression, Self-Harm, and Suicide Among U.S. Adolescents**  
**FIGURE 1. Indicators of poor mental health among U.S. girls and young women, 2001–2018 (note, before COVID)**



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



Research affirms the association between these charts.<sup>[32]</sup>

So do the facts, in individual cases like that of Molly Russell's. A coroner's inquest is like a trial in the United Kingdom. An inquest there investigated the alleged suicide of a fourteen-year-old, Molly. Voluminous evidence was taken and many witnesses called, including executives from Facebook. In a ruling that made headlines throughout Europe, the Coroner ruled the algorithms that curate a social media user's experience had pushed harmful content to Molly that she had not requested.<sup>[33]</sup>

Thousands of images, videos and other social media material from Molly's accounts were revealed during the investigation, one of the largest public releases of its kind. That provided the sort of detail that researchers studying the mental health effects of social media have long complained that platforms like Meta, which owns Facebook and Instagram, withhold on privacy and ethical grounds.

Molly's social media use included material so upsetting that one courtroom worker stepped out of the room to avoid viewing a series of Instagram videos depicting suicide. A child psychologist who was called as an expert witness said the material was so "disturbing" and "distressing" that it caused him to lose sleep for weeks.<sup>[34]</sup>

Molly is far from alone in her suffering. Consider these findings from a 2021 U.S. Surgeon General Advisory:

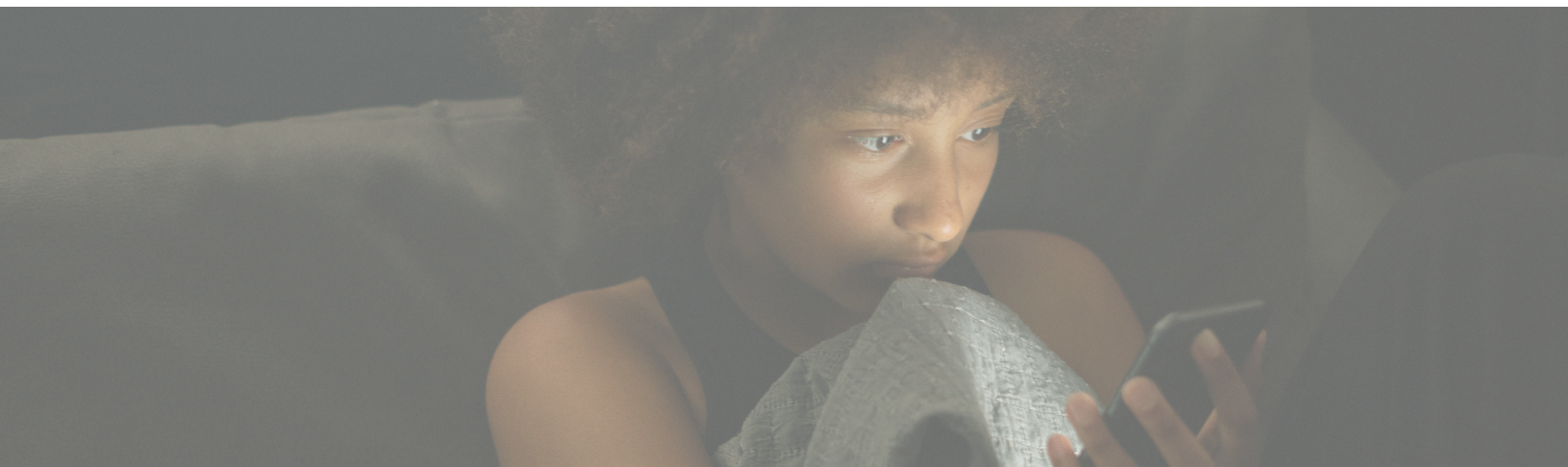
From 2009 to 2019, the proportion of high school students reporting persistent feelings of sadness or hopelessness increased by 40%; the share seriously considering attempting suicide increased by 36%; and the share creating a suicide plan increased by 44%. Between 2011 and 2015, youth psychiatric visits to emergency departments for depression, anxiety, and behavioral challenges increased by 28%. Between 2007 and 2018, suicide rates among youth ages 10-24 in the US increased by 57%.<sup>[35]</sup>

As Facebook's first President, Sean Parker, has admitted:

God only knows what it's doing to our children's brains.

The thought process that went into building these applications, Facebook being the first of them ...was all about: "How do we consume as much of your time and conscious attention as possible?" And that means that we need to sort of give you a little dopamine hit every once in a while ...you're exploiting a vulnerability in human psychology. The inventors, creators . . . understood this consciously.

And we did it anyway.<sup>[36]</sup>



The response of social media platforms to bi-partisan fury caused by their child-endangering practices, hundreds of lawsuits (which they are increasingly failing to have dismissed), a constant drumbeat of the worst publicity imaginable, and repeatedly shocking revelations about the lethal and damaging consequences of their products, has been, unbelievably, to cut back on safety even as their profits soar: “Meta Platforms' gross profit for the twelve months ending December 31, 2023, was \$108.943B, a 19.25% increase year-over-year. Meta Platforms' annual gross profit for 2023 was \$108.943B, a 19.25% increase from 2022.”<sup>[37]</sup>

It is apparent that these unimaginably wealthy corporations and the executives that lead them — Mark Zuckerberg earned \$29 billion in a single day<sup>[38]</sup> — are making so much money from the status quo of offering unsafe, addictive products to children that bad publicity and lawsuits filed under current law can be and are being easily absorbed as just the cost of doing business that is still otherwise fantastically profitable. Vigorous laws are, therefore, needed.

## CALIFORNIA'S LEGISLATURE TAKES ACTION

### AB 1394 REQUIRED PLATFORMS TO CREATE A REASONABLY ACCESSIBLE MECHANISM FOR REPORTING CSAM

Section 230 of the Communications Decency Act, as currently construed by courts, prevents California from passing a law directly compelling platforms to remove CSAM. However, in 2023, in response to overwhelming evidence that social media platforms play a pivotal role in promoting child sex trafficking and facilitating the efficient delivery of images and videos of CSAM to pedophiles, California's Legislature passed and Governor Newsom signed the bi-partisan AB 1394 (Wicks and Flora). In pertinent part, this landmark law requires social media platforms to create a “reasonably accessible” “mechanism” for survivors of child sexual abuse to report CSAM. It also enacted requirements governing how platforms must respond to such reports, including what they must say in response to those who report. Prior to enactment of AB1394, platforms were “not required to engage in efforts to combat revictimization, and [there was] no civil recourse for survivors when [platforms] refuse to engage in these efforts.”<sup>[39]</sup>

During the 2023 legislative process, the platforms asked for and were granted a year's delay in the effective date of the law to provide them with time to change their systems to comply with it. Thus, the law became effective January 1, 2025.

Those who work for social media platforms are the world's experts in designing platform user interfaces. Yet, as this report documents, screenshot-by-screenshot, the major social media platforms have created a confusing and daunting maze for survivors to report CSAM. After 14 months, children continue to suffer as:



- Reporting pathways remain buried behind multiple menus
- Language remains confusing, technical, and not user-friendly
- Survivors must guess, scroll, and click excessively to find the correct reporting path.

As this report shows, when it comes to child sex abuse survivors simply seeking to have images and videos of their abuse blocked — images and videos that violate the platforms' terms of service, that show crimes, and that aren't supposed to be on the platform anyway — the platforms have deployed their expertise with cruelty rather than compassion, demonstrating contempt for California's law and lawmakers in the process.

## **SB 1504 REQUIRED PLATFORMS TO CREATE A PROMINENT MECHANISM FOR REPORTING CYBERBULLYING**

Building on the precedent of AB 1394, in 2024, California's Legislature passed and the Governor signed the bi-partisan SB 1504 (Stern), which became binding law January 1, 2025.<sup>[40]</sup> SB 1504 addresses cyberbullying specifically. California defines "Cyberbullying" as:

...any severe or pervasive conduct made by an electronic act ... committed by a person or group of people directed toward one or more minors that has, or can be reasonably predicted to have, the effect of any of the following:

- (1) Placing a reasonable minor in fear of harm to that minor's person or property.
- (2) Causing a reasonable minor to experience a substantially detrimental effect on the minor's physical or mental health.
- (3) Causing a reasonable minor to experience substantial interference with the minor's academic performance.
- (4) Causing a reasonable minor to experience substantial interference with the minor's ability to participate in, or benefit from, the services, activities, or privileges provided by a school.<sup>[41]</sup>

SB 1504, like AB 1394, requires a reporting mechanism (a "prominent" mechanism, actually, in the case of 1504) and certain responses on set timetables.<sup>[42]</sup>



## THE CHILDREN'S ADVOCACY INSTITUTE'S EVALUATION

The following evaluation examines how four major platforms — Instagram, Facebook, Snapchat, and TikTok — have implemented reporting mechanisms on their mobile apps after AB1394 and SB1504. Regarding AB1394, our analysis focused specifically on whether their reporting systems are "reasonably accessible" as required by Civil Code Section 3273.66(a).

For our purposes, "reasonably accessible" means reporting tools should be:

- Visible within one or two clicks from the main interface
- Labeled with clear, direct language (e.g., "Report Child Sexual Abuse Material")
- Designed with consideration for traumatized young users with limited capacity to navigate complex systems
- Designed for smartphone use. Youth almost always access these platforms through smartphones, and so, what is "reasonably accessible," must be viewed through smartphone interfaces.<sup>[43]</sup>

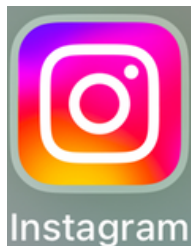
We documented the number of steps required to reach reporting tools, assessed language clarity, evaluated visibility from home screens, and identified confusing elements in each platform's reporting process.

Regarding SB1504, our analysis focused on whether reporting mechanisms were 'prominent' and easy for young survivors to access and navigate.

As our findings detailed in the following pages show, the platforms have created reporting pathways that force users through multiple confusing steps, bury relevant options behind vague terminology, and create unnecessary barriers for vulnerable users — particularly the young adolescents who are most frequently victimized.



## USER EXPERIENCE AB 1394



### Instagram (Meta)

Step-by-step User Experience

To report CSAM through Instagram, survivors—many of whom are children—are forced to scramble through an obscure and confusing process.

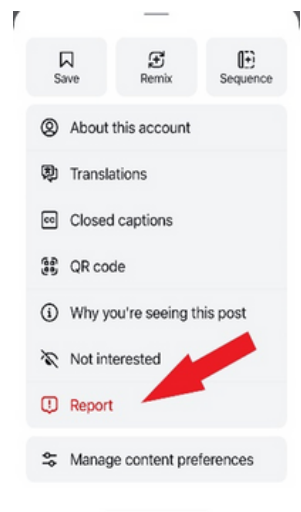
#### Click 1:

The process begins by clicking the three-dot menu on the top right of the post, but only if the user already knows to click the three-dot menu in the upper right corner of the post. (click1).



#### Click 2:

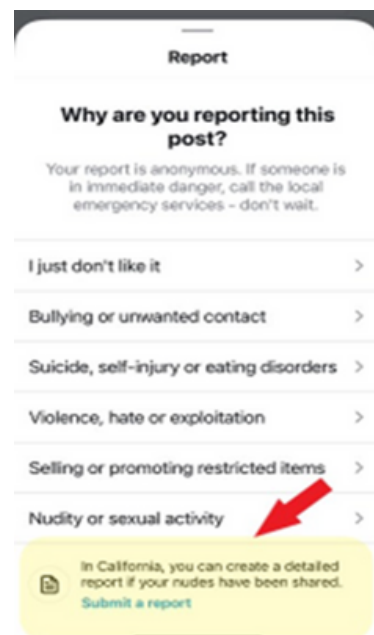
The “Report” option appears in red but at the very bottom of the pop-up. (click 2). The word “Report” is ambiguous. It is not immediately clear that this is where a user would go to request content to be blocked. Labels such as “Complain about an image or video” or “Report abuse” would more clearly communicate the purpose. As opaque as this early step is, this is the most intuitive part of the process. It gets worse from this point.



#### Click 3:

The next screen after clicking the ambiguous “Report” offers multiple reporting categories, but the California-specific reporting mechanism is referenced only in tiny print at the very bottom: *“In California, you can create a detailed report if your nudes have been shared.”* (Remember, this will likely be viewed on a cell phone.) The accompanying “submit a report” link is blue (click 3).<sup>[44]</sup>

While Instagram’s mention of California law might seem commendable, the execution is fundamentally flawed—perhaps intentionally so. The font size is significantly smaller than all other items on the page, making it easy to miss. More importantly, the language is deeply misleading. The term “nudes” fails to capture the nature of the harm in many cases. Survivors are often reporting images or videos depicting rape, abuse, or other sexual exploitation—not just nudity. Framing these as “nudes” diminishes the gravity of the abuse and may deter victims from coming forward.



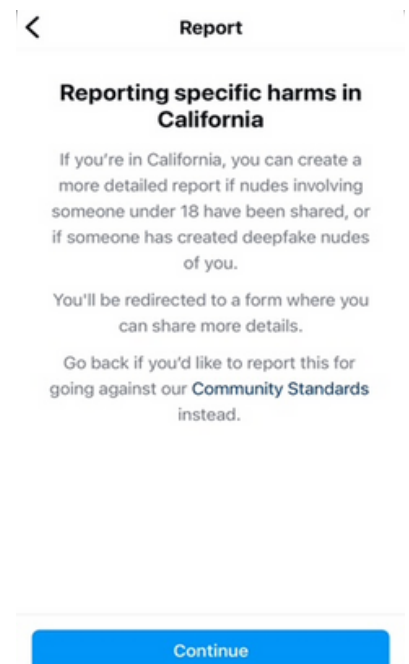
Additionally, the phrase “*your nudes*” introduces confusion. It implies that the survivor must have taken the image or held it in their possession, which is rarely the case in instances of CSAM. Just above this link is another category labeled “sexual activity,” which is far more accurate and aligned with the nature of CSAM and abuse, yet is not linked to the California-specific reporting process. This creates a confusing and counter-productive contrast.

Finally, at a minimum, the California reporting link should be displayed in equal font size and labeled with more precise, survivor-centered language—something like: “*Nudity or sexual activity involving minors (California residents).*” This would more accurately reflect the legal standard and the nature of the abuse being reported.

#### Click 4:

Clicking the blue “submit a report” link at the bottom of the page opens a new explanation screen describing California’s reporting rights, with a prompt to continue to a separate form.

The screen reads: “**Reporting specific harms in California.** If you’re in California, you can create a more detailed report if nudes involving someone under 18 have been shared, or if someone has created deepfake nudes of you. You’ll be redirected to a form where you can share more details. Go back if you’d like to report this for going against Community Standards instead.” (Note that there is a link to Community Standards that will take the user to a page describing community standards and not to a reporting page.) To continue to the reporting screen, the user must click the blue “continue” button at the bottom of the screen.



Instagram  
help.instagram.com

Instagram

How can we help?

Help Center

**Reporting specific harms in California**

If you're in California, you can create a more detailed report if:

nudes involving someone under 18 have been shared  
someone has created deepfake nudes of you

What are you reporting?  
Include the URL for what you're reporting. Copy the link from your browser and paste it below.

<https://www.instagram.com/reel/DJDgeHmp3Ym/>

What is the reason for this report?

**Nudes involving someone under 18**

Someone is threatening to share or have shared your online nudes.

**Deepfake nudes**  
AI-generated nudity content made to look like you.

Send

**Click 5:**

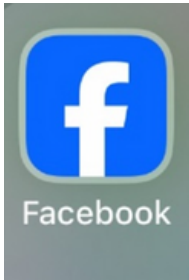
The form first asks the user to paste the URL of the post they are reporting, but the URL field is prefilled with a link that might be a link to the post the user is attempting to report, or it might not. The form instructs the user to “paste” the link, which implies that user action is required, even though a link is already populated. This mixed messaging is confusing. A user who did not copy the post URL before beginning the reporting process—after all, the user had no reason to expect they would need to do this—may now feel compelled to abandon the form, return to the post they are attempting to report, and start the process from scratch.

Assuming the user has been able to figure out how to post the URL, the form then asks the user to select one of two categories: (1) “Nudes involving someone under 18,” or (2) “Deepfake nudes.” The user then clicks “Send” to complete the report. (click 5). Note, again, the reference to “nudes” is too limiting to capture the requirements of California’s reporting law. At this point, the process ends—though no confirmation screen or next steps were reviewed, as the test did not involve submitting a false report. California law spells out with specificity what is required to happen next. But no information about what is legally required to happen next is provided to the reporting user, at any point prior to clicking “submit.”

We did not test what happens after submission to avoid generating false reports, so it is unknown whether Instagram provides a confirmation screen or complies with the follow-up communication requirements of Cal. Civ. Code § 3273.66(e). What is clear, however, is that the platform provides no information during the reporting process about what is legally required to happen next. Knowing that this process triggers legal obligations under California law for a platform to respond on certain timetables could influence a survivor’s decision to proceed and give them confidence that the platform is required to act. Yet none of that information is offered in the user interface prior to the user’s submission of the form, five clicks in.

**INSTAGRAM SUMMARY**

In sum, the difficulties, omissions, and ambiguities in Instagram’s CSAM, cyberbullying, and other harms reporting process are present from the very first step and persist throughout the entire experience, culminating in a confusing and unsupported final submission. The platform repeatedly and at every stage obscures access to the correct reporting path, and burdens survivors with navigating unclear instructions and misleading prompts. Given a full year to comply and Instagram’s vast resources and unparalleled expertise in user interfaces, it is implausible that these obstacles are accidental.<sup>[45]</sup>



## Facebook (Meta)

Step-by-step User Experience

Facebook, like Instagram, directs users through a generic process focused on “nudes” and “deepfake nudes,” which confusingly narrows the scope and risks excluding victims of other forms of child sexual exploitation. Like Instagram’s mechanism, Facebook’s process is vague, confusing, and likely to deter the very users it is intended to serve.

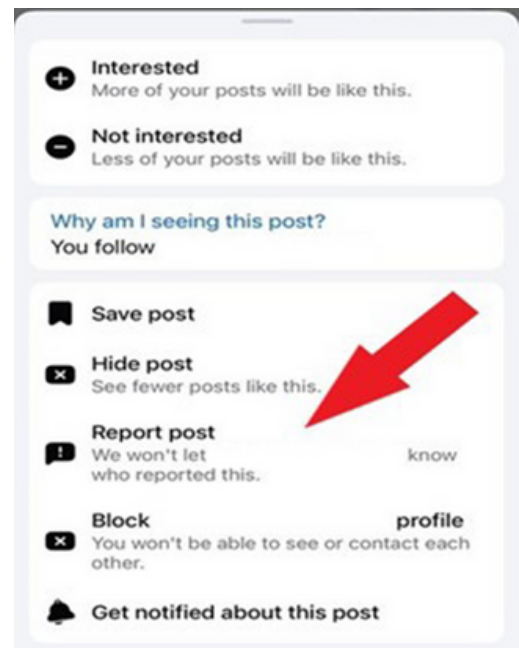
### Click 1:

The reporting process begins, again, only if the user already knows to click the three-dot icon in the upper right corner of the post.



### Click 2:

The “Report Post” option that appears after clicking on the three dots is about halfway down the resulting pop-up menu. It appears in plain black text, the same font and size as all other options, and lacks any visual emphasis. For a survivor in distress, there is absolutely nothing to signal that this is the place to request removal of CSAM.

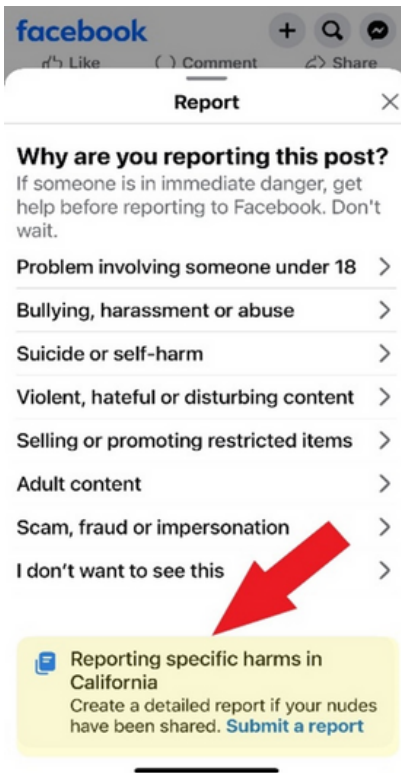


**Click 3:**

After Selecting “Report Post,” the user is directed to a screen titled “Why are you reporting this post?” At the very bottom of the screen—buried beneath numerous other reporting options such as “Problem involving someone under 18,” “Bullying, harassment or abuse,” “Violent, hateful or disturbing content,” and “Adult content”—Facebook includes a small section labeled “Reporting Specific Harms in California,” accompanied by a small blue icon.

While many of the options listed above could plausibly apply to CSAM, Facebook does not provide any clear direction to help users identify the legally required pathway for reporting CSAM under California law. Nor does it distinguish the California-specific option from Facebook’s general community standards reporting process. The lack of clear labeling or guidance significantly increases the risk that survivors will select an incorrect reporting category and never access the form intended to comply with state requirements.

Further, the design of the page gives no visual prominence to the California-specific reporting option. While the language may differ slightly from the general community standards categories, the text is the same size as all other entries, and there are no prominent visual cues to signal its legal significance or importance under California law. This makes it difficult for users to recognize that this is a distinct and legally significant pathway for reporting CSAM.

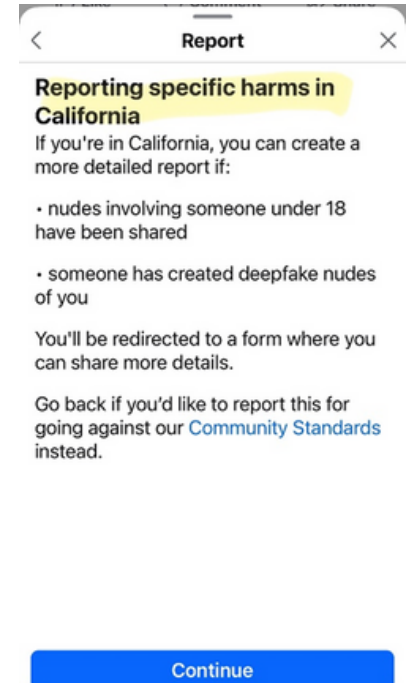


It reads: “Create a detailed report if your nudes have been shared. Submit a report.” (click 3) This language presents the same problems seen on Instagram. The word “nudes” is misleadingly narrow and fails to account for many forms of abuse that may not involve nudity. The possessive “your” may also imply that only content created or owned by the user can be reported, which could discourage survivors from proceeding.



**Click 4:**

After the user clicks the California-specific “Submit a report,” Facebook brings them to another screen titled “Report: Reporting specific harms in California.” The page explains that they can report either nudes involving someone under 18 or AI-generated deepfake nudes. Again, this is the same problem as above. “Nudes” deflects from the core problem – child sex abuse material. The text tells users they will be redirected to a form where they can share more details and offers a link to the platform's Community Standards for those who prefer to use that reporting route. Clicking the link opens yet another page without removing them from the reporting process. The user then clicks “Continue.”

**< Reporting specific harms in Calif...**
**Click 5:**

The next screen—“Reporting specific harms in California”—is identical to the one found on Instagram and presents the same challenges:

- **Prefilled URL field:**

- The form displays a URL that appears to correspond to the post being reported.
- However, there is no confirmation that the link is correct or an explanation of what it leads to.
- Instructions on the form tell the user to paste a link, despite one already being prepopulated, creating a confusing contradiction.

- **User confusion:**

- No previous screen informs the user that they will need a URL.
- A survivor may assume they need to return to the original post to retrieve the correct link, even though one is already present.
- The mixed messages increase uncertainty in a process that should be seamless and supportive.



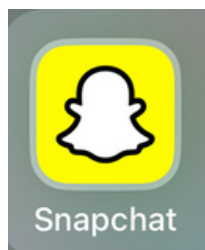
- **Ambiguity as a design flaw:**
  - This lack of clarity adds stress and doubt to what is likely already a traumatic experience.
- **Problematic category language:**
  - The form asks the user to choose one of two categories:
    - "Nudes involving someone under 18"
    - "Deepfake nudes"
  - As discussed above, the term "nudes" is too narrow to reflect the full range of CSAM.
  - Many CSAM images do not involve nudity, or nudity may not be central to what is being reported.
  - The phrase "your online nudes" is also problematic—it may suggest that only content personally created or owned by the victim qualifies for reporting.
- **Risk of deterrence:**
  - This language may discourage or confuse survivors, especially minors, from completing the report.

After selecting a category, the user must click "Send" to submit the report (click 5).

Again, we did not test what happens after submission to avoid generating false reports, so it is unknown whether Facebook provides a confirmation screen or complies with the follow-up communication requirements of Cal. Civ. Code § 3273.66(e). What is clear, however, is that the platform provides no information during the reporting process about what is legally required to happen next. Knowing that this process triggers legal obligations under California law for a platform to respond on certain timetables could influence a survivor's decision to proceed and give them confidence that the platform is required to act. Yet none of that information is offered in the user interface prior to the user's submission of the form, five clicks in. While it is possible that Facebook supplies this information after the user submits the form, that comes far too late in the process. The user should understand early on that this pathway is distinct from reporting content under the platform's Community Standards.

## FACEBOOK SUMMARY

Facebook's reporting mechanism suffers from the same flaws seen on Instagram: too many clicks, legally insufficient language, small and buried text, confusing and vague instructions, and a final form that risks losing the user entirely. These are not minor oversights. Facebook is a platform with vast resources and granular insight into user behavior. Facebook should be more than capable of designing a reporting system that is not only compliant with the law but also intuitive and supportive for traumatized users who are child victims of exploitation. That it has not done so speaks volumes.



## Snapchat

### Step-by-step User Experience

Snapchat's process for reporting CSAM lacks both legal clarity and accessibility, particularly for child survivors of sexual abuse. Unlike Instagram and Facebook, Snapchat does not provide a California-specific reporting option tied to the obligations outlined under AB 1394. Instead, the platform presents a series of generalized options that could apply to CSAM but fail to clearly identify or guide users toward a legally protected pathway for blocking and redress.

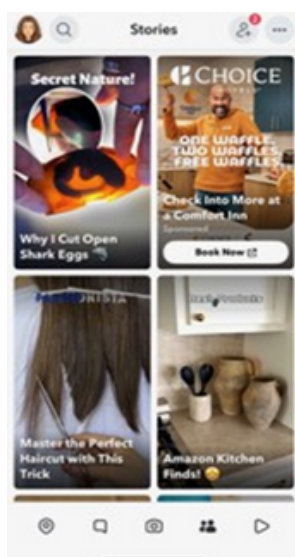
### Click 1:

The process begins when a user attempts to report a “Snap” or a “Tile,” but how to do so is far from intuitive.

On Snapchat, “**Tiles**” are thumbnail previews for Stories that appear in the Discover feed—a section featuring curated and personalized content from publishers, influencers, and creators. To report a Tile, the user must already know to press and hold it—there is no on-screen label or visual cue that indicates this action will open a reporting menu. If a user taps the Tile instead, it opens the associated Snap.

A “**Snap**” is a photo or short video that a user sends to friends, groups, or shares publicly. Once a Snap is opened—either directly from a friend or by tapping on a Tile—the user can access the reporting menu in one of two ways: by pressing and holding the Snap or by clicking on the three dots in the lower right corner of the screen. However, neither of these pathways is labeled or explained anywhere in the interface. The platform provides absolutely no visible guidance to indicate that pressing and holding or using the three-dot menu are the necessary steps to begin a report.

For survivors, especially young users experiencing distress, this lack of instruction creates a serious barrier to reporting abuse and accessing legal protections. Snapchat’s reliance on hidden functions and user familiarity, rather than clear and accessible reporting tools, undermines the platform’s responsibility to support survivors.



### Tiles

*Press and Hold*



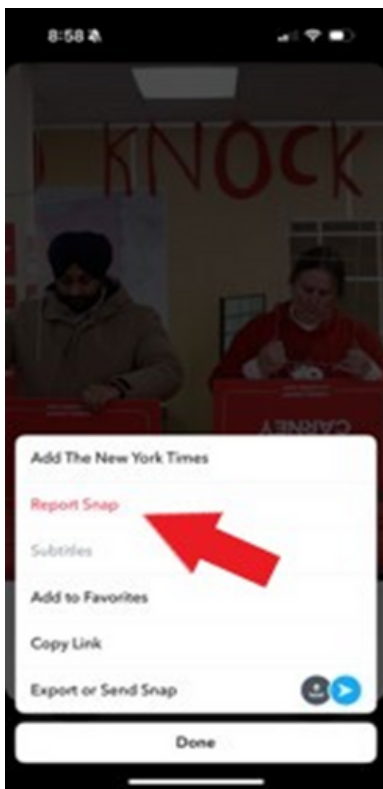
### Snap

*Press and Hold  
OR  
Click on 3 dots*

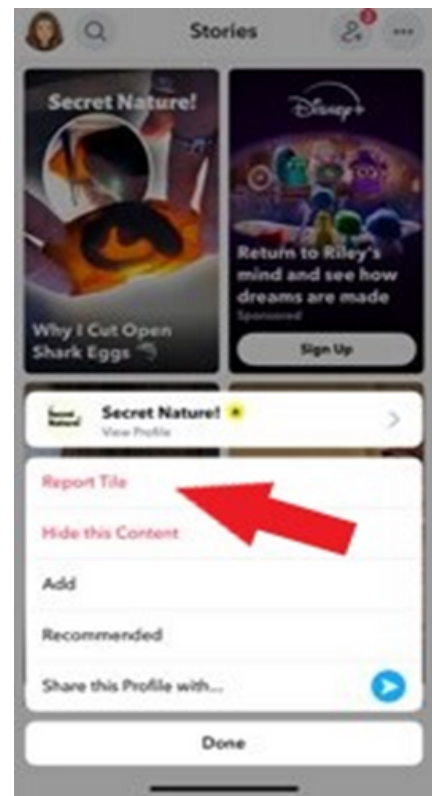
**Click 2:**

When the user initiates a report, the next screen varies depending on whether the report starts from a Tile or a Snap. If the user presses and holds a Tile, the reporting menu appears with the username of the account that posted the Tile at the top. Below that, the first clickable item, in red text, is “Report Tile.”

If the user reports from a Snap—either by pressing and holding or clicking on the three dots in the lower right corner—a similar menu appears, but the red “Report Snap” option is listed second. This inconsistency between Tile and Snap interfaces can create additional confusion for users unfamiliar with Snapchat's layout. Although the red font adds some visual emphasis, neither label clearly indicates that this is the appropriate option to report CSAM and to request removal under California law.



**Snap**  
Screen 2



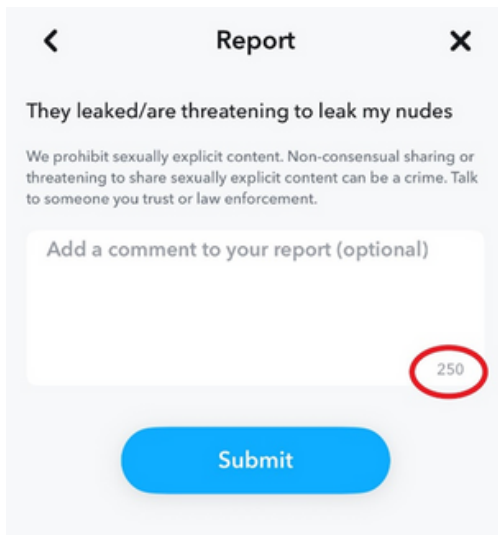
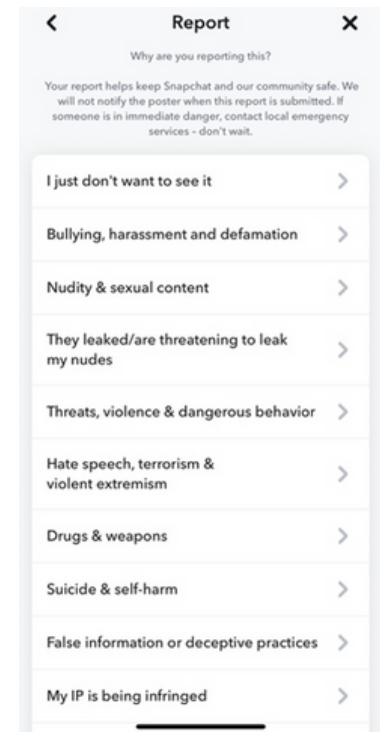
**Tiles**  
Screen 2



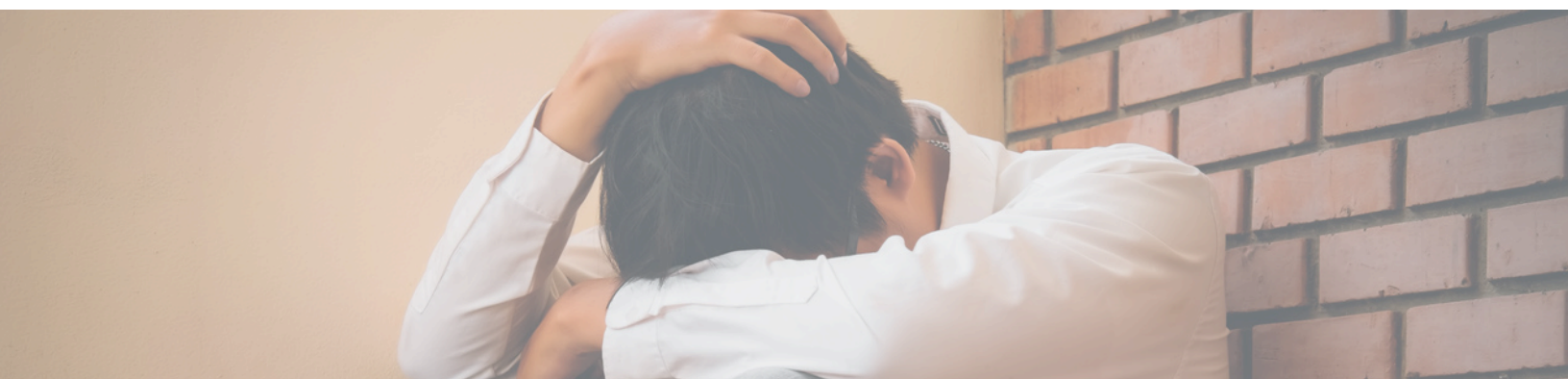
**Click 3:**

After clicking “Report,” a screen titled “Report” appears. From this point forward, the reporting process is the same whether the user started from a Tile or a Snap. The user must choose from one of eleven options, none of which clearly indicate a path for reporting CSAM in compliance with California law. Several of the options may seem relevant, but the platform offers no guidance on which to choose. For example, the user might select “Nudity & Sexual Content” or “They leaked/are threatening to leak my nudes,” both of which could plausibly relate to CSAM. But the lack of any explanation leaves the user to guess which is the most appropriate path or if their reporting will fall under “other” (the 11<sup>th</sup> option). This ambiguity can lead to confusion, incomplete reporting, and ultimately, failure to access the protections the law is meant to guarantee.

The following two examples demonstrate how this lack of clarity plays out in practice, revealing how Snapchat’s reporting options force users to navigate vague and ambiguous language with no clear guidance for reporting CSAM in compliance with California law.

**Click 4:**

If the user selects “They leaked / are threatening to leak my nudes,” a new screen appears with that heading. In very small print, it states: “We prohibit sexually explicit content. Non-consensual sharing or threatening to share sexually explicit content can be a crime. Talk to someone you trust or law enforcement.” The screen includes a text box allowing up to 250 characters for the user to add a comment to the report. To put this into perspective, 250 characters is roughly 40 to 50 words—barely enough for a short paragraph. That is all the space a survivor has to describe the horrific content they’re trying to report. And in many cases, the person making the report is a victimized and vulnerable child, already under immense emotional stress. Expecting them to compress their report into a handful of sentences is not just unreasonable, it is cruel, retraumatizing, and unjust.



After writing their comment (which is optional), the user can click the blue “Submit” bar to send the report (click 4). We did not submit a report, but up to this point in the process, the platform has provided no information about how a CSAM report, once submitted, is handled under California law. There is no explanation of what to expect next, what obligations the platform must fulfill, or what rights the user has.

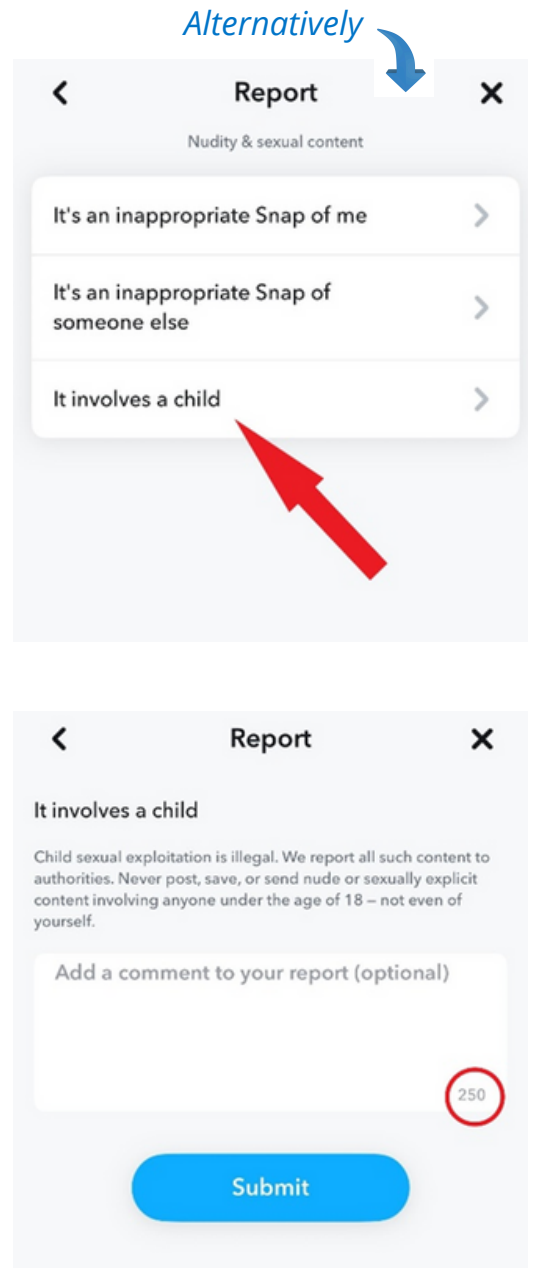
Alternatively, the user might select “Nudity & Sexual Content,” which brings up another set of options: “It’s an inappropriate Snap of me,” “It’s an inappropriate Snap of someone else,” and “It involves a child.” If the user selects “It involves a child” (click 4), a final screen appears.

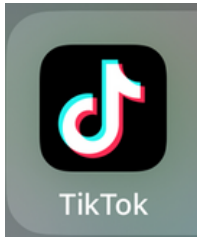
The screen that appears contains a brief explanation: “Child sexual exploitation is illegal. We report all such content to authorities. Never post, save, or send nude or sexually explicit content involving anyone under the age of 18—not even of yourself.” Again, 250-character text box follows, allowing the user to submit further details. (click 5)

While this step is the closest Snapchat comes to acknowledging CSAM, it still does not inform the user of their rights under California law. The platform provides no assurance that it will follow mandated timelines, send confirmation of receipt, or update the user on the outcome of their report. Nor does it differentiate the legal obligations of reporting CSAM from a standard violation of community guidelines.

## SNAPCHAT SUMMARY

Snapchat offers a fragmented and inadequate reporting mechanism for survivors of child sexual abuse. The platform provides no clear legal pathway for reporting CSAM under California law and places the burden on users to guess which reporting path to choose. Survivors are left to decipher vague options, scroll through buried menus, and rely on incomplete language, all in the hope of accessing protections the law is supposed to guarantee. These omissions are not excusable oversights; they are the result of a system that prioritizes profitable ambiguity over accountability.





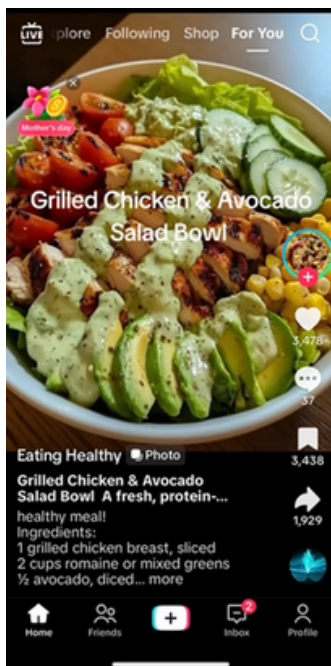
## TikTok

### Step-by-step User Experience

TikTok's process for reporting CSAM is highly fragmented, confusing, and inadequate. Although TikTok includes some references to child sexual abuse in its menu structure and support materials, it fails to provide a direct, clearly labeled, or easily accessible pathway for survivors to report CSAM in compliance with California law. Instead, the platform presents the user with multiple overlapping menus, vague and ambiguous labels, and inconsistent instructions that leave even experienced users confused, and vulnerable users, particularly children, completely unsupported.

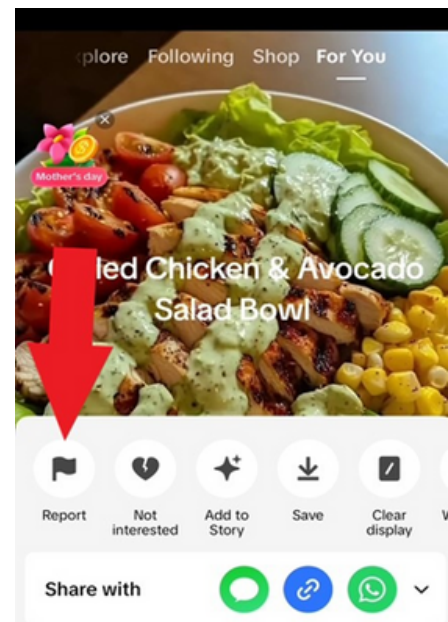
#### Click 1:

To begin the reporting process, the user must know to press and hold on the TikTok post they want to report. This action is not labeled or explained.



#### Click 2:

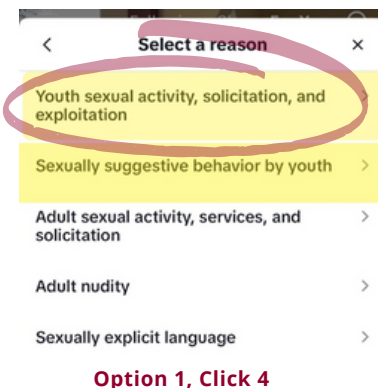
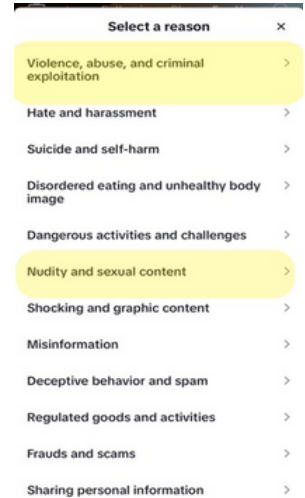
Once the user presses and holds, a menu appears at the bottom half of the screen. To proceed, the user must click a black flag icon on a white circle—the Report icon—located in the upper left side of the menu.



### Click 3:

The user is then taken to a “Select a reason” menu with no fewer than 14 different but sometimes overlapping choices. On a mobile phone, the user must scroll to view all the options. None of these choices is labeled as pertaining to CSAM or to California's legal reporting process. Several options could apply to CSAM, including “Nudity and sexual content,” “Violence, abuse, and criminal exploitation,” and “Human exploitation.” The user is left to guess which is most appropriate. There is no explanation of what CSAM is, no information about the user's rights under California law, and no mention of the platform's obligations when a user reports CSAM.

Below, we explore three of the several unclear options for the user.



### Option 1 - “Nudity and sexual content” (click 4 and 5)

From the list of 14 choices, the user might reasonably select “Nudity and sexual content” as a path to report CSAM, given the lack of any explicit CSAM category. This option appears to be the most intuitive match for content that is sexually exploitative, especially involving minors. If the user selects “Nudity and sexual content,” they are presented with five subcategories:

- Youth sexual activity, solicitation, and exploitation
- Sexually suggestive behavior by youth
- Adult sexual activity, services and solicitation
- Adult nudity
- Sexually explicit language

A CSAM post could plausibly contain content that falls into one or more of these categories. If the user selects “Youth sexual activity, solicitation, and exploitation” (click 4), they are directed to a “Report” screen. The top of the screen reads in bold: “Youth sexual activity, solicitation and exploitation,” followed by: “We don't allow the following,” and five bullet points. The first bullet states: “Showing or promoting child sexual abuse material (CSAM) or youth sexual activity.”

Thus, it is not until this point, well into the in-app reporting process and deep into fine print, that CSAM is explicitly mentioned, and it appears without any explanation of user rights, next steps, or the platform's obligations under California law and is only found after selecting a prompt which could cover consensual activity. The screen features a bright red “Submit” button, but does not offer the user the opportunity to add context or further details prior to submitting the report. (click 5)

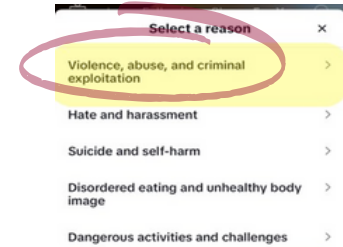


### Option 1, Click 5 (Submit)

Adding to the confusion, the report screen for “Youth sexual activity, solicitation, and exploitation” also states in a gray box at the bottom of the page, just above the submit button, “If this is cyberbullying, submit a separate report.” There is no guidance on how or where to do this, although pressing the bold print in the box directs the user to another page related to cyberbullying (note that it is not immediately obvious that the gray box is a link).

## Option 2 - Violence, abuse, and criminal exploitation” - Exploitation...under 18 (starting at click 3).

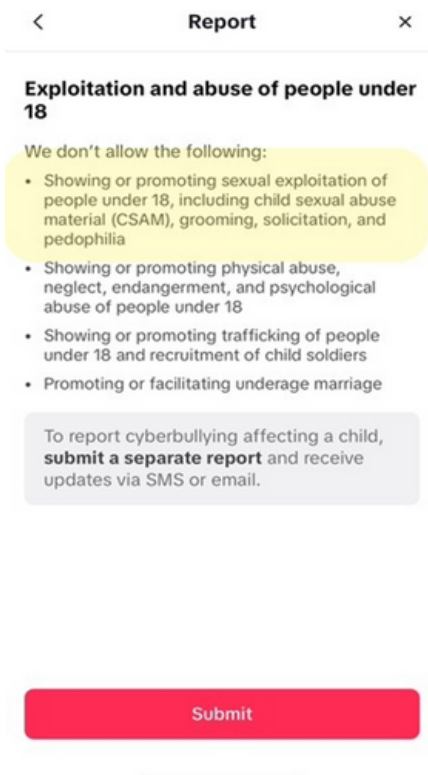
If the user selects “Violence, abuse, and criminal exploitation” (click 3), they are taken to another menu with six options. One option —“Exploitation and abuse of people under 18”— seems relevant to CSAM (click 4). If selected, it leads to a report page titled “Exploitation and abuse of people under 18.” The page states: “We don't allow the following,” and includes four bullet points. The first bullet reads: “Showing or promoting sexual exploitation of people under 18, including child sexual abuse material (CSAM), grooming, solicitation, and pedophilia.” This is again the first time in this reporting path that CSAM is mentioned, many steps into the process. This exact language could be elevated much earlier, of course. Moreover, at no point is the user informed of their rights under California law or the platform’s obligations related to CSAM reports.



Option 2, Click 3



Option 2, Click 4



Option 2, Click 5 (Submit)

At the bottom of the page, again, a gray box reads: “To report cyberbullying affecting a child, submit a separate report and receive updates via SMS or email.”

Following the gray box is a bright red “submit” button for the user to submit a report without any option to provide further context or detail.

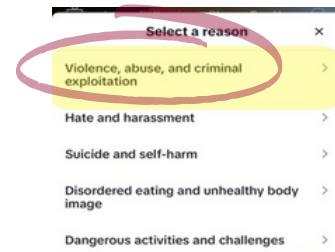


### Option 3 - Starting from "Violence, abuse, and criminal exploitation" option (click 3)

If the user selects "Sexual exploitation and abuse," from the "violence, abuse, and criminal exploitation" menu, (click 4) they are brought to a similar page with the same structure: a title in bold, followed by four bullet points. However, in this case, none of the bullet points explicitly mention CSAM, which is bizarre as the title specifically mentions "sexual exploitation and abuse."

The user can click the bright red "submit" button at the bottom of the page without providing more information.

- or -



Option 3, Click 3



Option 3, Click 4

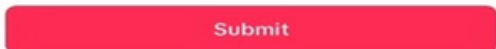


#### Sexual exploitation and abuse

We don't allow the following:

- Showing or promoting non-consensual sexual acts that are real or fictional, including rape and molestation
- Showing or promoting non-consensual sharing of intimate content or threats to share such content
- Editing content to sexualize someone or create the appearance of them engaging in sexual activity
- Unwanted or degrading statements, such as statements about someone's private body parts, sexual activity, or private sex life

To report child sexual abuse content or sexually explicit digital identity theft affecting you or a child, **submit a separate report** and receive updates via SMS or email.



Option 3, Click 5 (Submit or Grey Box)

Alternatively, at the bottom of the page, another gray box appears with the message: "To report child sexual abuse content or sexually explicit digital identity theft affecting you or a child, submit a separate report and receive updates via SMS or email."

This message is vague and confusing. It does not explain where or how to submit the separate report, and it is unclear whether the user should click the red "Submit" button or follow an alternative reporting path. Moreover, grouping identity theft under "sexual exploitation" is confusing. The gray box itself appears passive and non-interactive; it contains no clear link or button. However, if the user presses the bold print in the gray



box, it will actually take the user to a form titled "Submit a request." (Note, "request," not "report") This page includes a lengthy explanation citing the federal definition of child pornography and explaining "sexually explicit digital identity theft."

[Submit a request - Sexual Exploitation and A...](#)

## Submit a request

At TikTok, our goal is to create a safe and welcoming community. We do not allow content, real or AI-generated, containing any form of sexual exploitation or abuse. If you or your child are the person depicted, fill out this form to report:

- Child sexual exploitation / abuse** (i.e. child pornography as defined by 18 U.S.C § 2256 or other obscene matter that depicts you or your child engaging in or simulating sexual conduct)
- Sexually explicit digital identity theft** (i.e. a non-consensual digitally created or altered image or video of your intimate body part or you engaging in a sexual act)

We report incidents of child sexual abuse and exploitation to the [National Center for Missing and Exploited Children \(NCMEC\)](#), as required by law. We also report to relevant law enforcement authorities when there is a specific, credible, and imminent threat to human life or serious physical injury. **If you have been the victim of a crime or are in imminent risk of harm, please contact law enforcement directly for immediate assistance.**

If you have a question about your TikTok account, your content, or our [Community Guidelines](#), or for reporting other issues, visit our [Help Center](#).

If you wish to report cyberbullying of a child, please use this [form](#).

Fields marked with an asterisk (\*) are required.

Your email address\*

How do you want TikTok to communicate with you?\*

Email

State\*

Select your state or territory of residence in the US. For residents outside the United States, please review TikTok's Safety Center for reporting options: <https://www.tiktok.com/safety/en/reporting>

What app is your report about? \*

What issue are you reporting?\*

What type of content are you reporting? \*

**More details (optional)**  
If you don't have the link to the content, provide as many details as possible for us to locate the reported content.

**Declaration**

I am the person depicted in the material reported (or their parent/guardian/representative).\*

I ensure, to the best of my ability and knowledge, that all the information disclosed above is accurate and true.\*

I acknowledge TikTok's Terms of Service and Privacy Policy. TikTok may share the information submitted with the National Center for Missing and Exploited Children or law enforcement as required by law.\*

**Submit**

**Option 3, Click 6/Submit**

Scroll...

The page notes that TikTok reports incidents to NCMEC and includes links to:

- NCMEC
- TikTok Community Guidelines
- Help Center
- Cyberbullying report form

Below this information, the user must scroll to complete the form. Required fields include:

- Email address
- Preferred communication method (Email or SMS)
- State of residence
- App the report is about (TikTok or Lemon8)
- Issue being reported (“Child sexual exploitation and explicit digital identity theft”)
- Type of content being reported (Video, Photo, Comments, Direct Messages, Other)

An optional text field allows the user to add more details. Finally, the user must check three boxes under a legal declaration affirming they are the person depicted (or their guardian), that the information is accurate, and that they acknowledge TikTok’s privacy policy and understand the report may be shared with law enforcement or NCMEC.

After completing this separate form and including all required information, the user can click the red “Submit” button at the bottom of the page. (click 6)

- Violence, abuse and criminal exploitation
  - Sexual exploitation and abuse (this pathway)
  - Human exploitation
- Nudity and Sexual Content
  - Sexually suggestive behavior by a youth
  - Adult sexual activity, services, and solicitation
  - Adult nudity
  - Sexually explicit language
- *NOTE that the most intuitive place – “youth sexual activity, solicitation, and exploitation does NOT include a pathway to this reporting mechanism. Instead, it includes a pathway to report cyberbullying (see above).*

## TIKTOK SUMMARY

TikTok’s CSAM reporting system is overly complex, poorly labeled, and inconsistently executed. While the platform appears to be aware of its legal obligations, it has failed to design a clear, direct path for survivors to report abuse and access the protections required by California law. Instead, it requires users to guess their way through confusing menus, misleading categories, and vague language.

The result is a process that is not only opaque and inadequate but is so convoluted that it borders on obstruction. For survivors, particularly children, this is likely to be more than a frustrating user experience; it is retraumatizing. Given TikTok’s vast financial and technological resources and its ability to optimize user behavior with pinpoint precision, it is difficult to believe this confusion is accidental. If TikTok can algorithmically serve personalized content in milliseconds, it can certainly build a clear, accessible reporting path for its most vulnerable users. That it has not done so reflects a failure not just of compliance, but of conscience.

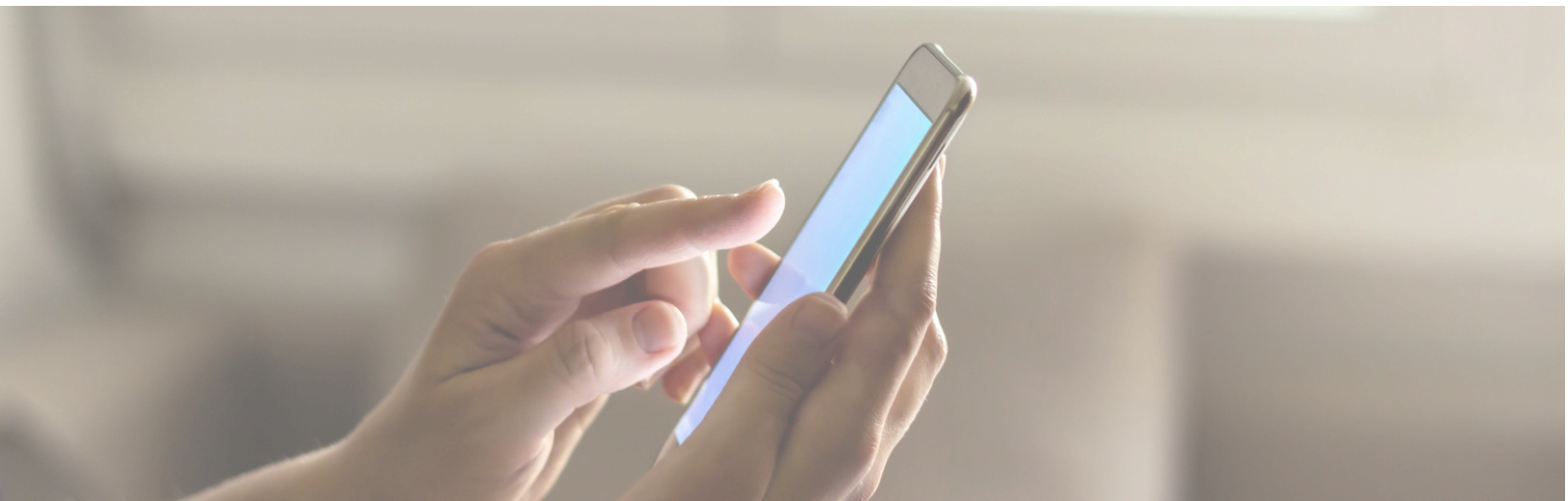
**FINAL OBSERVATIONS REGARDING AB 1394: INSTAGRAM, FACEBOOK, SNAPCHAT, AND TIKTOK FAIL TO ACT WITH THE COMPASSION EVERY REGULAR PERSON WOULD OFFER WITHOUT HESITATION, THUMBING ITS NOSE AT CALIFORNIA'S GOVERNOR AND LEGISLATURE**

The context:

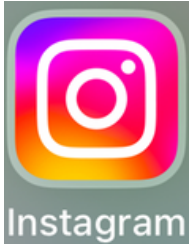
- These corporations are the world's experts in how to facilitate user engagement on social media platforms.
- The images and videos show images of child rape and sexual abuse. The victims are the most sympathetic imaginable; the material is among the most disturbing imaginable.

Granted by the California Legislature a full year to comply with a modest and compassionate law offering hope to survivors of child rape and abuse, each of the four major platforms have purposefully made it needlessly complicated and frustrating to report images and videos of sex crimes against children; abuse that shock the conscience, abuse they know how to screen for and detect, abuse that should never be offered for viewing on their products in the first place.

Given this experiment in offering these corporations the benefit of the doubt, both in enacting flexible standards and a year to comply, it is hard to come to any conclusion but this: these vastly wealthy corporations have elected to thumb their nose at both common decency and this Legislature's leniency. A legislative inch was given. They took not just one but all the miles.



## USER EXPERIENCE SB 1504



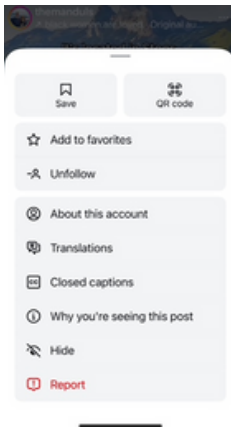
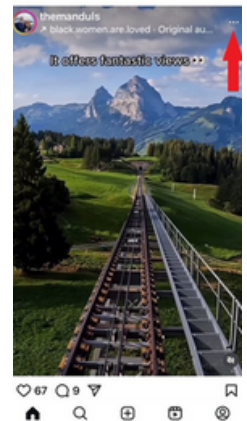
### Instagram (Meta)

#### Step-by-Step User Experience

Instagram's cyberbullying reporting mechanism is not prominent, as required by California law. Key steps are hidden behind unclear menus, and the process takes numerous clicks. For youth in distress, the lack of clarity and visibility creates an intimidating and ineffective experience.

#### Click 1

To report cyberbullying from an Instagram post, the user must first identify how to access the reporting function. This step is not intuitive. There is no visible "report" button on the post itself. Instead, the user must already know—or discover through trial and error—that tapping the three horizontal dots in the upper right-hand corner of the post opens a menu of options. The reporting mechanism lacks prominence. There is no visual or textual cue indicating that this menu contains reporting options, which could make the process difficult to navigate for users unfamiliar with the platform or in distress, especially minors, parents, or school personnel trying to act quickly.

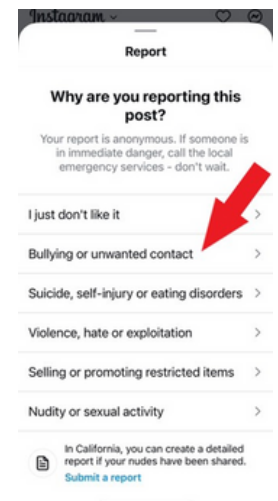


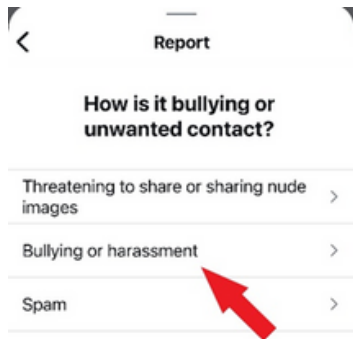
#### Click 2

Clicking the three dots brings up a short menu. Near the bottom of this menu—below options like "Hide" and "Why am I seeing this post"—is a red "Report" option. The color draws attention to the "Report" option once the menu is opened. The user selects "Report."

#### Click 3

After selecting "Report," a new screen opens with the title "Why are you reporting this post?" followed by a note that reports are anonymous and that if someone is in immediate danger, emergency services should be contacted. The screen displays nine categories the user must choose from, including "spam," "violence, hate or exploitation," and "nudity of sexual activity". The user must scroll through to see all of the options. The second option listed is "Bullying or unwanted contact," which the user selects to report bullying.





**Click 4**

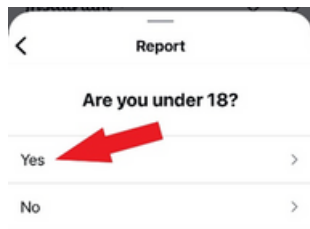
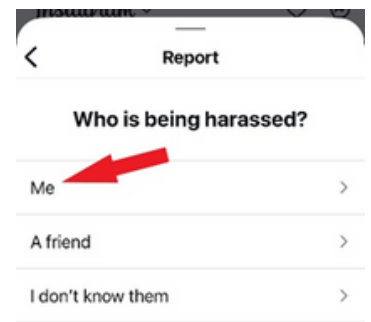
At this point, a new screen appears titled “How is it bullying or unwanted contact?” The user is given three choices:

- Threatening to share or sharing nude images
- Bullying or harassment
- Spam

To continue the cyberbullying report, the user taps “Bullying or harassment.”

**Click 5**

On the next screen, the platform asks the user “Who is being harassed?” with three response options: “Me,” “A friend,” or “I don't know them.” Notably, the term “harassed” is used exclusively here, dropping the word “bullying” which had previously been part of the labeling in earlier steps. This shift in terminology may confuse some users, particularly youth or caregivers, who may not equate bullying with harassment or may expect consistent language throughout the reporting process. Despite this, the screen itself is generally clear. The user selects “Me.”



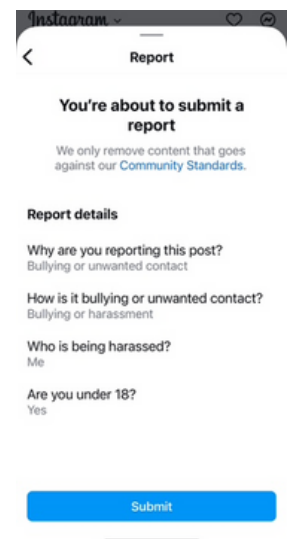
**Click 6**

The following screen asks, “Are you under 18?” with simple “Yes” or “No” options. The user selects “Yes” (click 6). Notably, this is the first point at which the platform acknowledges the user’s age, which is highly relevant to SB 1504’s focus on protecting minors from cyberbullying.

**Click 7**

The final screen in the sequence states “You’re about to submit a report.” It explains that Instagram only removes content that violates its Community Standards, and provides a link to those standards. The user is shown a summary of their report, confirming the options selected in the previous steps. To submit, the user must tap a large blue “Submit” button. (click 7)

We did not test what happens after submission to avoid generating false reports, so it is unknown whether Instagram provides a confirmation screen or complies with the follow-up communication requirements of California Business and Professions Code § 22589.1(b).



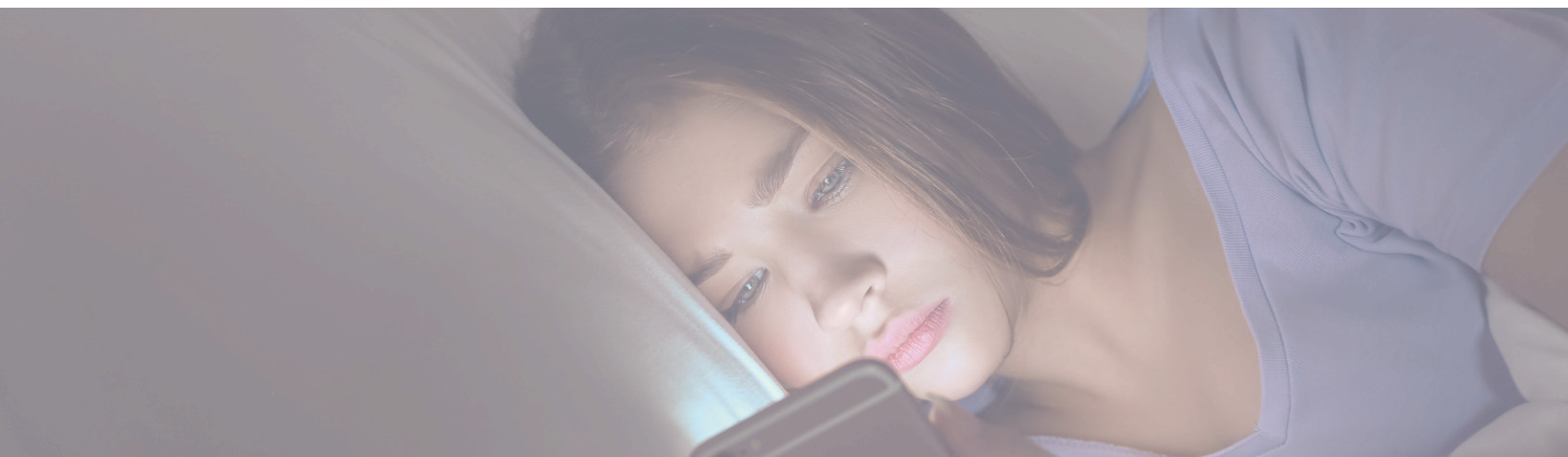
As mentioned, users are informed that Instagram only removes content that violates its Community Standards. While this message may manage expectations, it could also discourage users from completing the report, especially if they are uncertain whether the conduct they are reporting will be considered a violation. Crucially, the platform has failed to provide the user with information about their rights or the platform's responsibilities under California law. Therefore, while the platform provides a message to manage expectations, the user is left without a clear understanding of their rights and the platform's responsibilities following a report.

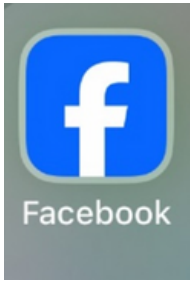
## INSTAGRAM SUMMARY

Instagram's reporting process for cyberbullying, while functional, requires seven clicks and ultimately fails to provide essential, prominent, easy-to-access protections for its most vulnerable users. It is important to note that, in this evaluation, no report was submitted in order to avoid creating a false report. As such, it is possible that some elements required by SB 1504 (as codified in Cal. Bus. & Prof. Code § 22589.1) that would make the process more user friendly and effective, such as the ability to upload evidence, provide contact information, or receive follow-up communication, could appear only after the user completes the final step. However, based solely on the process visible up to the point of submission, these features are entirely absent.

There is no opportunity to upload screenshots or provide identifying details, no option to enter a preferred method of contact, and no explanation of what to expect after submission. Users are told only that Instagram removes content that violates its Community Standards. This may have a chilling effect, discouraging users from completing a report if they are unsure whether the content constitutes a violation.

Crucially, there is no indication—before submission—that the user will receive a confirmation within 36 hours, periodic updates, or a final determination. Instagram's current design lacks the clarity and transparency necessary to encourage users to complete the reporting process and ensure meaningful access to protections under California's Cyberbullying Protection Act.





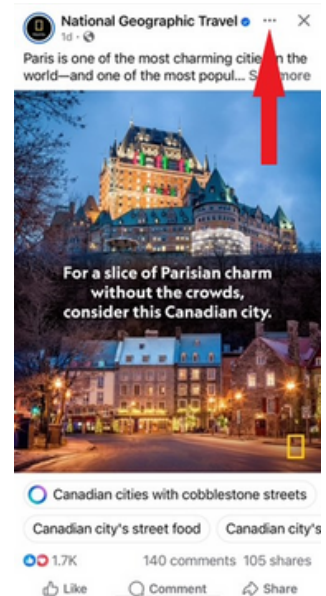
## Facebook (Meta)

### Step-by-Step User Experience

Facebook offers two paths to report cyberbullying, but both are confusing, repetitive, and lack clear labeling. These design flaws undermine the law's intent to provide a clear and accessible process for youth.

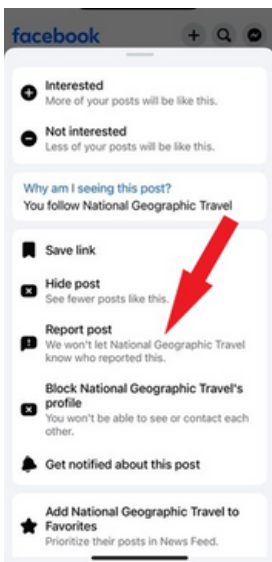
#### Click 1

To report cyberbullying from a post on Facebook, a user must first know to click on the three small horizontal dots in the upper right-hand corner of a post. This is not visually highlighted or labeled in a way that signals its purpose as a reporting tool. The three small dots do not stand out at all and, in fact, appear to be designed not to stand out, making it very easy to overlook this important tool.



#### Click 2

Upon clicking the three dots, a menu appears. "Report Post" is the fifth option down, presented in black font on a white background, again, without any visual prominence (such as color, size, or bolding) to make it stand out. This placement and formatting make the reporting option neither prominent nor intuitive, especially to youth in distress or individuals unfamiliar with platform functions.



#### Click 3

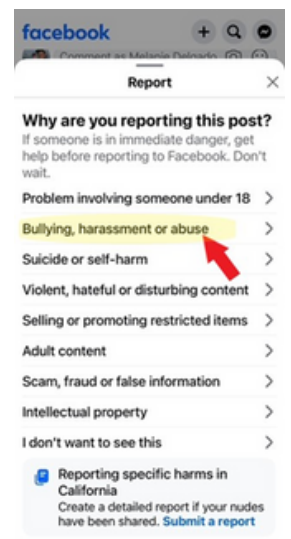
The user is then taken to a screen titled "Why are you reporting this post?" with a reminder that if someone is in immediate danger, users should get help before reporting to Facebook.

At this stage, the user is presented with several reporting categories. Two separate pathways exist that a user might reasonably use to report cyberbullying.

In the first pathway (Path 1), the user clicks the top option: "Problem involving someone under 18." (click 3) In the second pathway (Path 2), the user clicks the option "Bullying, harassment or abuse." (click 3)



Path 1

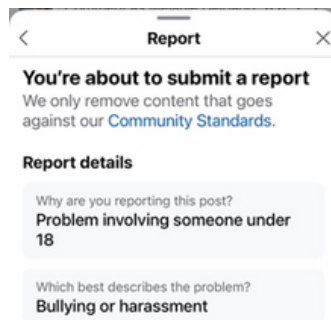
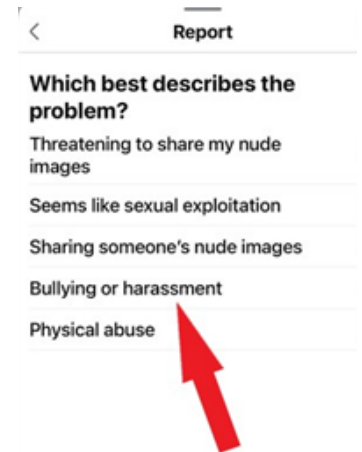


or

Path 2

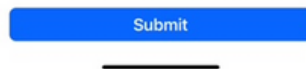
**Path 1****Click 4**

If the user chooses the first option (Problem involving someone under 18), they are directed to a screen asking, “Which best describes the problem?” with five choices. The fourth choice in the list, near the bottom, is “Bullying or harassment.” The user selects this option to report bullying. (click 4)

**Path 1****Click 5**

Clicking “Bullying or harassment” leads to a submission screen that recaps the selected report details - the issue involves someone under 18, and the problem is bullying or harassment.

Note that at the top of the page, Facebook includes a statement that it only removes content that violates its Community Standards (with a link to those standards). No information is provided at this point regarding what the user can expect after submission; however, the platform clearly states that only content violating its standards will be removed. This framing, absent any assurance of follow-up or explanation of user rights, may discourage users from completing a report—especially if they are uncertain whether the content qualifies.

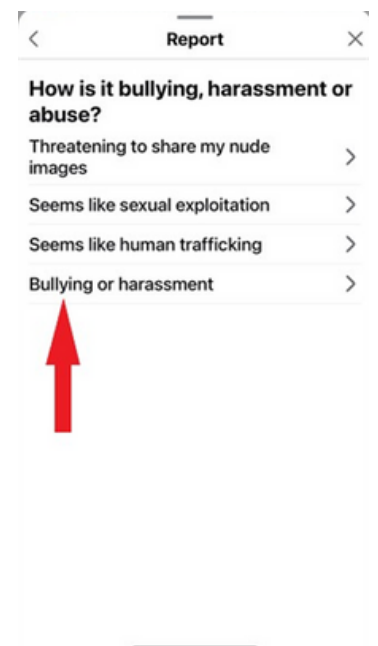


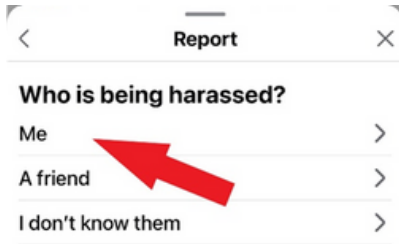
A blue “Submit” button appears at the bottom of the screen. The user clicks the button to submit. (click 5) We did not submit a false report; therefore, this is where our evaluation for Path 1 ended.

**Path 2****Click 4**

If the user clicks the second option, “Bullying, harassment or abuse,” they are directed to a screen with a follow-up question: “How is it bullying, harassment or abuse?” The fourth and final item at the bottom of this list is again “Bullying or harassment.” (click 4)

This leads to a follow-up question: “How is it bullying, harassment, or abuse?” The fourth and final item in the provided list of options is again “Bullying or harassment,” essentially requiring the user to repeat the same answer they selected previously.





### Path 2 Click 5

The user is taken to a screen that asks, "Who is being harassed?" There are three choices: "Me," "A friend," or "I don't know them." The user chooses "Me."

### Path 2 Click 6

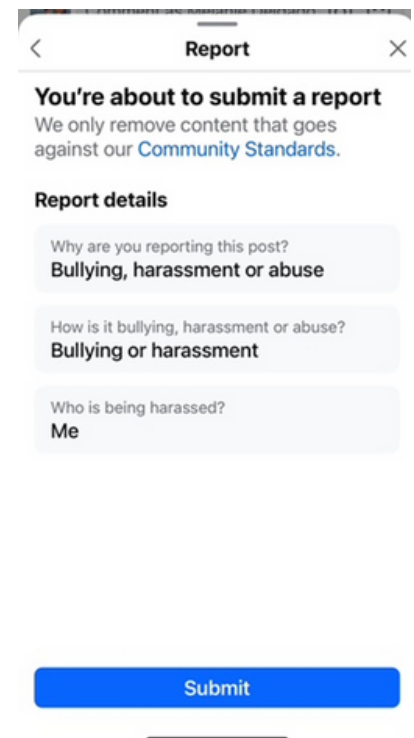
The user is taken to the final report screen, which, as in path 1, includes a link to Facebook's Community Standards at the top of the page and summarizes the report based on prior responses: Reporting bullying, harassment, or abuse, and the user is being harassed (me).

As in the first pathway, Facebook clearly states that only content violating its standards will be removed and provides a link to community standards. There is no information provided about follow-up, user rights, or what to expect after submission. Again, this language, absent context or any assurance of follow-up, may cause confusion or hesitation, particularly among users uncertain whether the content meets the criteria.

Notably, the report here does not indicate whether or not the user is under age 18. This omission introduces confusion about how reports are classified. Under California law (Cal. Bus. & Prof. Code § 22589), cyberbullying is defined as severe or pervasive conduct, committed by electronic means, and directed toward one or more minors. It is not clear whether a report submitted through the "bullying, harassment or abuse" path without also identifying the victim as under 18 is categorized or logged as cyberbullying. This is problematic because a victim of cyberbullying could quite reasonably select 'bullying' as the most relevant description, without realizing that doing so may prevent the report from being handled as cyberbullying under California law.

A blue "Submit" button appears at the bottom of the screen. The user clicks this button to submit. (click 6)

We did not submit a false report; therefore, this is where our evaluation ended for path 2.



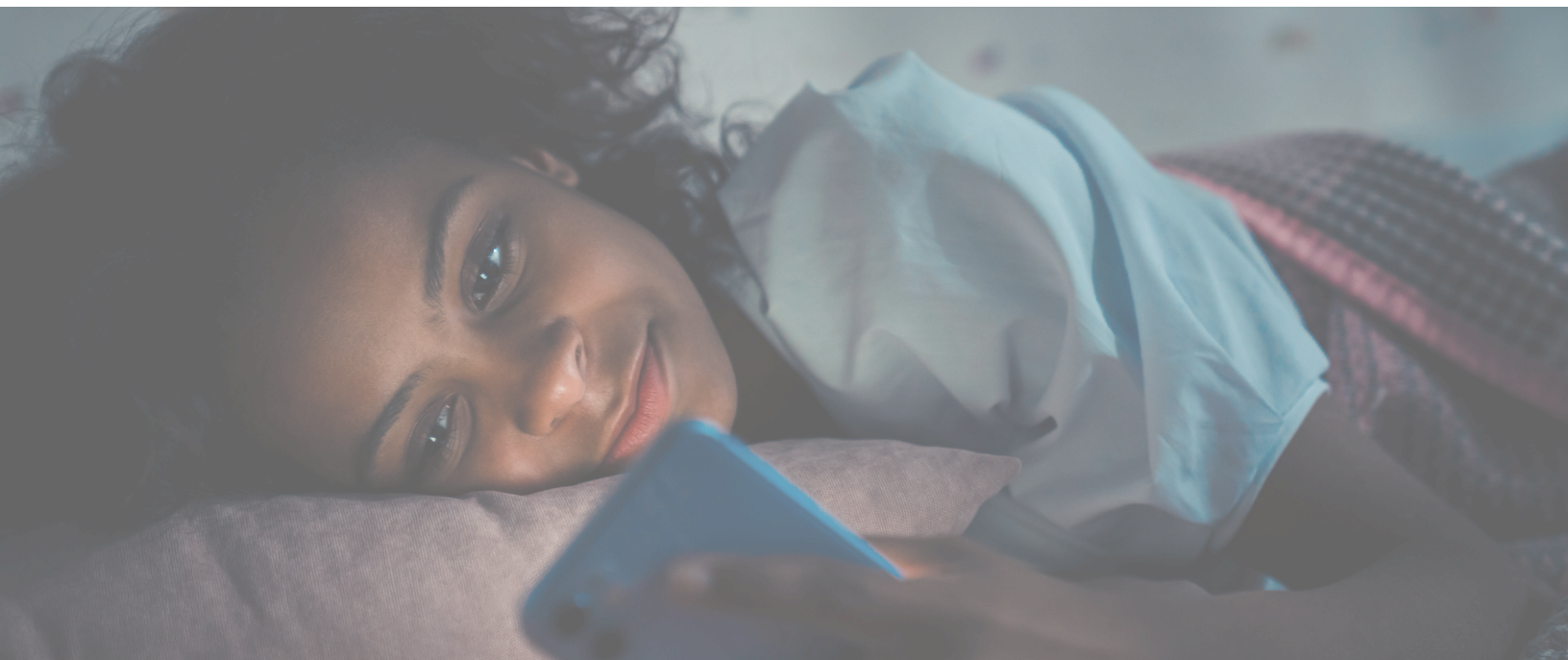
## FACEBOOK SUMMARY

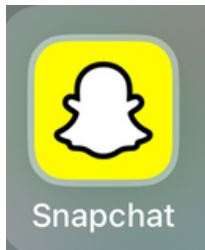
While Facebook offers a pathway to report cyberbullying, its design does not prioritize ease of use for young, anxious, or traumatized users. The reporting option is hidden behind a small three-dot menu, with labels and paths that are functional but not intuitive. Users must choose between two similar reporting routes — one beginning with "problem involving someone under 18," the other with "bullying, harassment or abuse" — and are prompted to select essentially the same option multiple times. This redundancy adds unnecessary frustration.

Although "bullying or harassment" is a reasonably clear term, the lack of an explicit "cyberbullying" label, combined with the absence of guidance on how Facebook categorizes reports, raises concern. A victim may choose the more intuitive "bullying" path without realizing that doing so — without separately indicating the victim is under 18 — could mean the report is not flagged as cyberbullying.

No information is provided about what happens after submission. There is no guidance on user rights, no indication of follow-up, and the consistent reminder that only content violating Community Standards will be removed may discourage uncertain users from completing a report. For someone already navigating fear or trauma, this vague and bureaucratic system feels more like a barrier than a support.

Facebook has the capability to design seamless, highly responsive user experiences — it does so routinely for profit-generating features. But that same level of design care is not evident here. If the reporting mechanism were treated with the urgency and attention of a revenue tool, it could be far more visible, user-friendly, and effective. As it stands, it misses the opportunity to truly serve the needs of vulnerable users.





## Snapchat

### Step-by-step User Experience

Snapchat's reporting system for cyberbullying is difficult to discover and filled with vague, legalistic language. It does not confirm whether the victim is a minor, and offers no explanation of what happens after submission. The experience is opaque and disempowering, especially for young users trying to seek help.

### Snap

#### Click 1

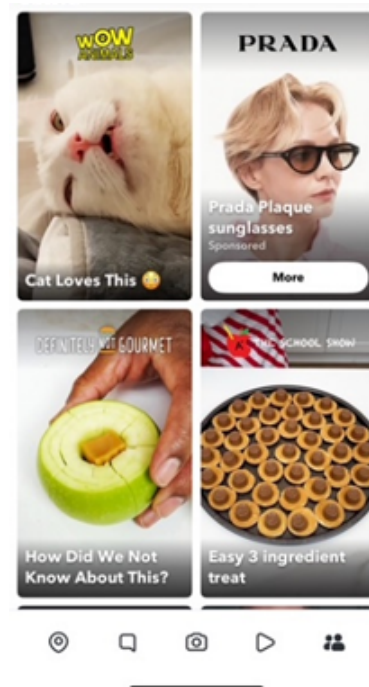
To report cyberbullying from a Snap on Snapchat, the user must first discover that they need to press and hold on the Snap itself to bring up the reporting menu. This is not intuitive, especially for users in distress who may be seeking a clearly visible, user-friendly path to report harmful content. The method is not labeled or guided in any way on the screen. User presses and holds on the Snap.



### Tile

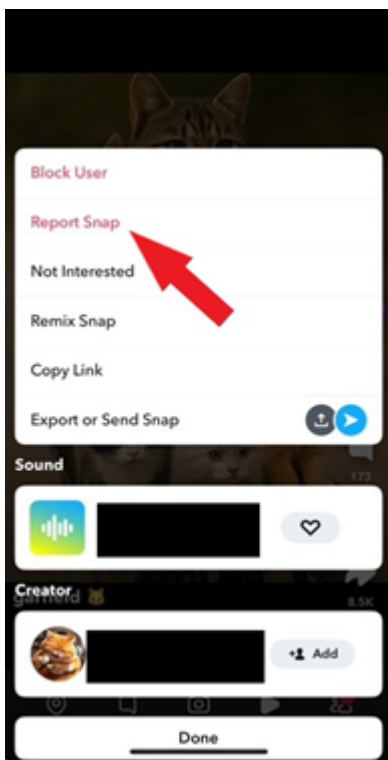
#### Click 1

To report cyberbullying from a Tile on Snapchat, the user must first discover that they need to press and hold on the Tile they want to report to bring up the reporting menu. Again, this is not at all intuitive. User presses and holds on the Snap.



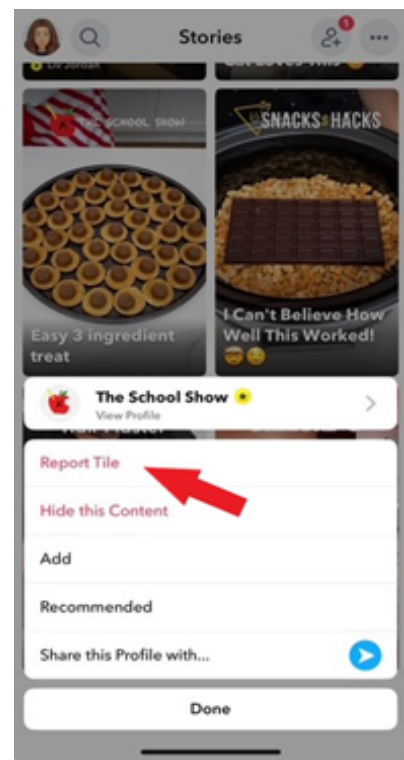
**Snap**  
**Click 2**

Pressing and holding the Snap or clicking on the three horizontal dots in the lower right-hand corner of the Snap brings up a menu of options. The top two, "Block User" and "Report Snap," are presented in red, making them stand out against the rest of the list, which appears in black text on a white background. The reporting function is the second option on the list. User chooses "Report Snap."



**Tile**  
**Click 2**

Pressing and holding the Tile brings up a menu of options. The top option, below the name of the profile that posted the Tile is "Report Snap" presented in red, making it (and "Hide this Contact" beneath it) stand out against the rest of the list, which appears in black text on a white background. The reporting function is the first option on the list. User chooses "Report Tile."

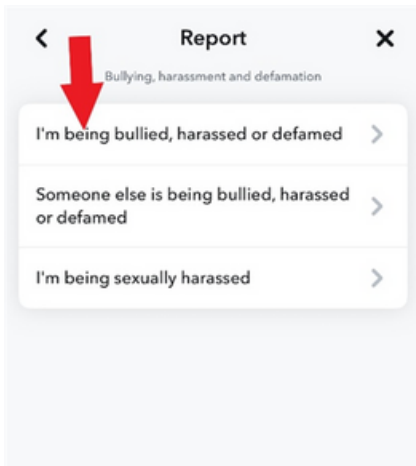
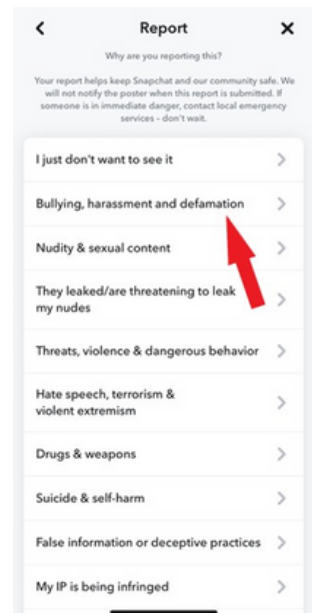


### From here on, reporting is the same for Snaps and Tiles:

#### Click 3

Clicking "Report Snap" or "Report Tile" brings up a screen titled "Report," with the subheading "Why are you reporting this?" Below that, in small dark grey text on a light grey background, Snapchat offers a message: "Your report helps keep Snapchat and our community safe. We will not notify the poster when this report is submitted. If someone is in immediate danger, contact emergency services - don't wait." While this content is thoughtful and potentially reassuring to users experiencing harm, the color scheme and small font size visually deemphasize the critical message.

After the message at the top of the page, the user is presented with a list of 11 reporting options. The second on the list is "Bullying, harassment and defamation." This is the closest match to a cyberbullying category, though there is no specific "cyberbullying" label. That said, the term "bullying" is a clear and logical option—it describes the conduct in question, even if it does not precisely mirror the legal terminology. For many users, particularly youth, "bullying" would be an intuitive choice.



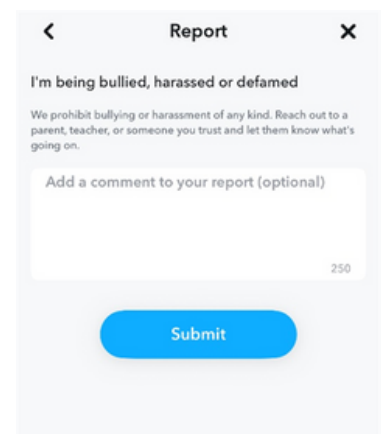
#### Click 4

Next, the user is taken to another report screen labeled "Bullying, harassment and defamation," which offers three choices. The user selects "I'm being bullied, harassed, or defamed." (click 4)

Notably, the language groups three distinct concepts under one umbrella, and terms like "harassed" or "defamed" may be unfamiliar or intimidating to youth. A person being bullied may hesitate to choose this if they aren't sure the broader language applies.

#### Click 5

The final screen is labeled "I'm being bullied, harassed, or defamed." It includes a short message in the same small dark grey text on a light grey background: "We prohibit bullying or harassment of any kind. Reach out to a parent, teacher, or someone you trust and let them know what's going on." Again, this is good advice, but visually minimized with small grey font on a light grey background. The screen includes a space to add a comment (up to 250 characters), which equates to roughly 1-2 sentences, not nearly enough room to describe context, impact, or details that may help Snapchat understand the severity of the situation.



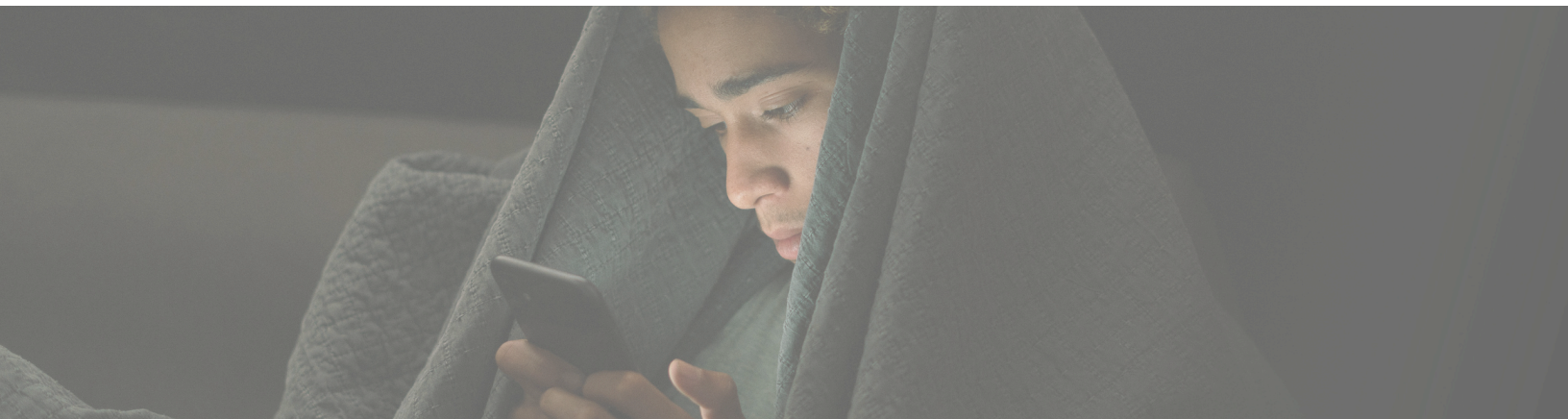
We did not submit a false report, therefore, this is where our evaluation ended.

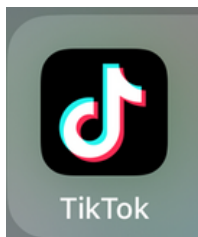
At no point in the process does Snapchat inform the user of what will happen once the report is submitted. There is no description of user rights, follow-up procedures, or how the platform handles these reports. There is also no way to indicate whether the person being targeted is under 18, despite California law (Cal. Bus. & Prof. Code § 22589) defining cyberbullying specifically as behavior directed toward minors. This raises questions about how Snapchat categorizes such reports and whether they are flagged and addressed under the legal definition of cyberbullying.

## SNAPCHAT SUMMARY

The process of reporting cyberbullying on Snapchat is functional but not intuitive. Users must know to press and hold on a Snap to begin, a step that is not signposted or explained. This lack of prominent cues in the Snapchat reporting process creates a barrier for distressed or inexperienced users. While the interface becomes more direct after this initial hurdle, important guidance is often deemphasized through small grey text on a grey background, including assurances about anonymity and community safety.

The reporting category most relevant to cyberbullying is labeled “Bullying, harassment and defamation.” Although the term “bullying” is recognizable and appropriate, its pairing with legalistic terms like “defamation” may cause confusion, especially for younger users. The platform does not ask whether the victim is under 18, nor does it inform the user about what happens and the platform's responsibilities after submission. The absence of transparency around follow-up procedures and categorization leaves unclear whether reports are processed as cyberbullying under California law. Overall, the process lacks clarity, visibility, and user-centered design that could reasonably be expected from a platform of Snapchat's sophistication and resources.





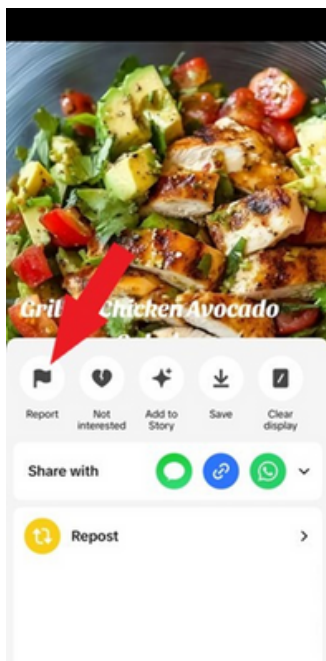
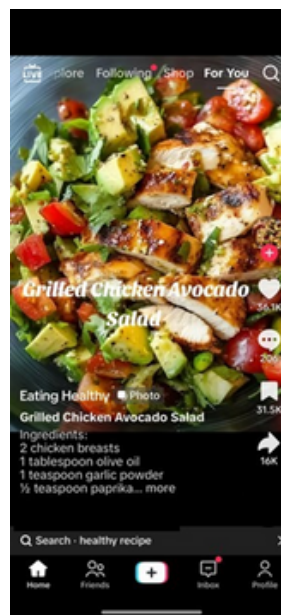
## TikTok

### Step-by-step User Experience

TikTok's process for reporting cyberbullying is buried behind unclear categories and inconsistent terminology. The reporting path varies depending on user guesses. The result is a confusing system that fails to meet the needs of vulnerable users or the legal standard of prominence.

#### Click 1

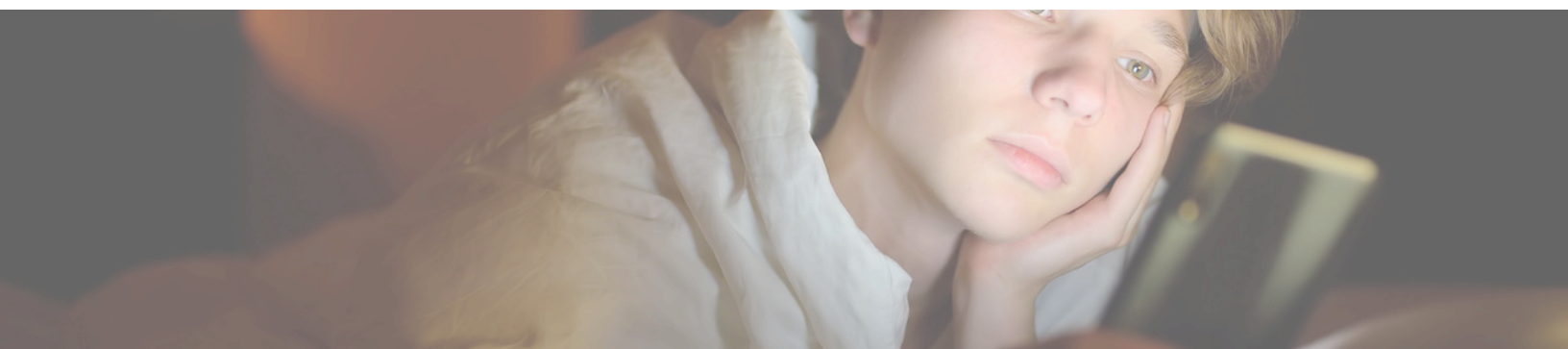
To report cyberbullying from a post on TikTok, a user must first know to press and hold on a post to bring up the reporting menu. This step is not clearly labeled or explained, making it difficult to access for users in distress or unfamiliar with platform norms. Again, the lack of prominent cues in the TikTok reporting process creates a barrier for distressed or inexperienced users. To report, the user presses and holds the post.



#### Click 2

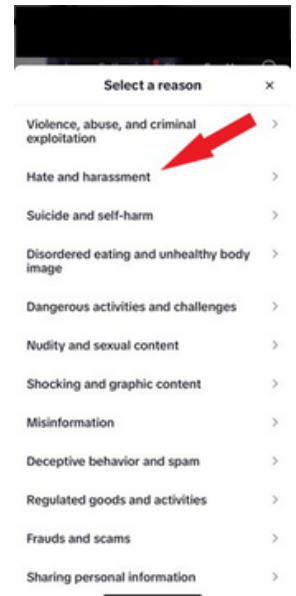
The user is taken to a menu. In the upper left corner, a small black flag icon appears above the word "Report" on a white background. This is the only indication that reporting is possible from this screen, the icon and the word "report" are not set apart visually with a different color or larger font.

The user clicks on "Report."



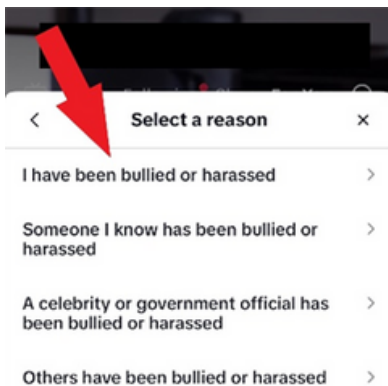
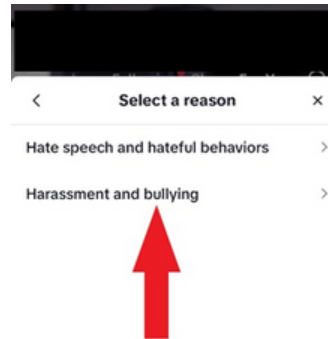
**Click 3**

Clicking the report icon brings the user to a screen titled "Select a reason." Fifteen options are displayed for the user to scroll through, but notably and incredibly, not one mentions "bullying." The user, who may be young and distressed, must correctly guess that "Hate and harassment" is the relevant category. If the user guesses correctly, they click "Hate and harassment."



**Click 4**

The user is brought to another "Select a reason" screen with two options. Finally, here, on the 4<sup>th</sup> screen, the term "bullying" appears for the first time — in the second option "Harassment and bullying." The user clicks on Harassment and bullying.



**Click 5**

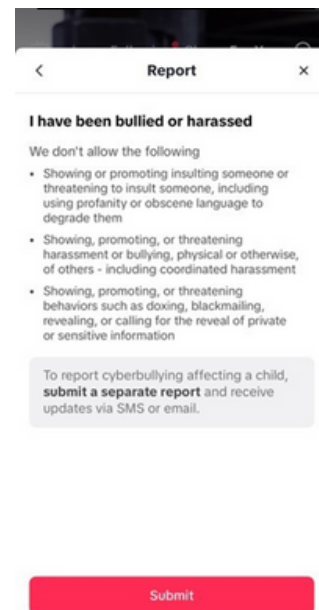
The user is taken to a third "Select a reason" screen, this time with four options. The user selects "I have been bullied or harassed," the first option, to report bullying. Note this extra step includes "bullied or harassed," which is redundant of "Harassment and bullying" in the prior step,

**Click 6**

This leads to the final Report screen. The screen states, "We don't allow the following," followed by a list of prohibited behaviors: showing, promoting, or threatening to insult, harass, or bully others (including physical harassment), and behaviors such as doxing.

Beneath this list, in a light grey box, the platform includes a message: "To report cyberbullying affecting a child, submit a separate report and receive updates via SMS or email." The phrase "submit a separate report" is bolded but not hyperlinked or explained, and there is no indication of how to do so.

We did not submit a false report; therefore, this is where our evaluation ended.



## TIKTOK SUMMARY

The pathway to report cyberbullying on TikTok is far from intuitive and lacks prominence: it requires multiple screens, none of which provide a clear or consistent visual cue that cyberbullying reports are being categorized appropriately. The fact that the term "bullying" only appears after three screens—and only when navigating through a path labeled "hate and harassment"—makes the mechanism needlessly difficult to use, particularly for young or distressed users. It also remains unclear how, or whether, TikTok classifies such reports as cyberbullying under California law (Cal. Bus. & Prof. Code § 22589), which applies to conduct directed at minors. There is no place for the user to indicate the age of the person being bullied or harassed.

Although the final screen makes some effort to define what the platform prohibits, the absence of visible follow-up procedures, contact method selection, or categorization guidance suggests that the reporting system is not designed with the urgency or clarity this issue deserves. Given TikTok's vast resources and technological capabilities, it is reasonable to expect a more compassionate and user-centered design—particularly for a platform that knows its users well enough to generate billions in revenue. TikTok should apply the same level of design sophistication to its reporting mechanisms as it does to its profit-generating features.



**FINAL OBSERVATIONS REGARDING SB1504: INSTAGRAM, FACEBOOK, SNAPCHAT, AND TIKTOK FAIL TO ACT WITH THE COMPASSION EVERY REGULAR PERSON WOULD OFFER WITHOUT HESITATION, THUMBING ITS NOSE AT CALIFORNIA'S GOVERNOR AND LEGISLATURE**

The context:

- These corporations are the world's experts in how to facilitate user engagement on social media platforms.
- The children here are victims of cyberbullying, which often leads to lifelong trauma and struggles at school, which can destroy a child's educational and professional potential. At worst, cyberbullying can lead to a child taking their own young life.

Yet, despite their expertise, each of the four major platforms have made it needlessly complicated and frustrating to report cyberbullying. Again, this leads to the conclusion that these vastly wealthy corporations are thumbing their nose at both common decency and the California Legislature's and Governor's policymaking. Here, too, a legislative inch was given. Once more, they took not just one but all the miles.

## END NOTES

- <sup>[1]</sup> Joanna Gill, *Can an EU Law Save Children from Harmful Content Online?*, REUTERS (July 12, 2022, 1:35 PM), <https://www.reuters.com/legal/litigation/can-an-eu-law-save-children-harmful-content-online-2022-07-12/>.
- <sup>[2]</sup> *2024 CyberTipline Report*, Nat. Ctr. for Missing and Exploited Children, (2025), <https://www.missingkids.org/content/dam/missingkids/pdfs/cybertiplinedata2024/2024-CyberTipline-Report.pdf>.
- <sup>[3]</sup> *United States Senate Committee on the Judiciary “Protecting Our Children Online”*, Nat. Ctr. for Missing and Exploited Children, 2-3 (2023) (emphasis added), <https://www.judiciary.senate.gov/imo/media/doc/2023-02-14%20-%20Testimony%20-%20DeLaune.pdf>.
- <sup>[4]</sup> Kyleigh Feehs & Alyssa Currier Wheeler, *2020 Federal Human Trafficking Report*, HUM. TRAFFICKING INST., 44 (2021), <https://traffickinginstitute.org/wp-content/uploads/2022/01/2020-Federal-Human-Trafficking-Report-Low-Res.pdf>.
- <sup>[5]</sup> *Id.*
- <sup>[6]</sup> Alpert, L. I., *Tiktok Knew Live Video Feature ‘Groomed’ Minors*, Utah AG Claims, BLOOMBERG (January 3, 2025), <https://www.bloomberg.com/news/articles/2025-01-03/tiktok-knew-live-video-feature-groomed-minors-utah-ag-claims>.
- <sup>[7]</sup> Alicen Rodolph, Angela Gray, & Lindsey Lane, *2023 Federal Human Trafficking Report*, HUM. TRAFFICKING INST., 72 (2021), <https://traffickinginstitute.org/wp-content/uploads/2024/06/2023-Federal-Human-Trafficking-Report-WEB-Spreads-LR.pdf>.
- <sup>[8]</sup> *Child Sexual Abuse Material (CSAM)*, THORN, <https://www.thorn.org/research/child-sexual-abuse-material-csam/>.
- <sup>[9]</sup> *Id.*
- <sup>[10]</sup> *Responding to Online Threats: Minors’ Perspectives on Disclosing, Reporting, and Blocking*, THORN, 11 (2021), [https://info.thorn.org/hubfs/Research/Responding%20to%20Online%20Threats\\_2021-Full-Report.pdf](https://info.thorn.org/hubfs/Research/Responding%20to%20Online%20Threats_2021-Full-Report.pdf).
- <sup>[11]</sup> *Child Sexual Abuse Material*, U.S. Dept. Of Just., 1 (2023), [https://www.justice.gov/d9/2023-06/child\\_sexual\\_abuse\\_material\\_2.pdf](https://www.justice.gov/d9/2023-06/child_sexual_abuse_material_2.pdf).
- <sup>[12]</sup> *Child Pornography*, U.S. Dept. Of Just., (August 11, 2023), <https://www.justice.gov/criminal-ceos/child-pornography>. Sadly, these feelings usually persist and even intensify over time. U.S. Dept. Of Just., *The National Strategy For Child Exploitation And Prevention And Interdiction*, 11 at D-12 (2010), available at <http://www.justice.gov/psc/docs/natstrategyreport.pdf> (finding that almost ninety-five percent of CSAM victims suffer lifelong psychological damage and may never overcome the harm, even after lifelong therapy). The problem has taken on a new dimension as CSAM involves increasingly younger victims and is becoming more violent and graphic over time.
- <sup>[13]</sup> *Reviewing Child Sexual Abuse Material Reporting Functions on Popular Platforms*, CAN. CTR. FOR CHILD PROT., 6 (2020), [https://protectchildren.ca/pdfs/C3P\\_ReviewingCSAMMaterialReporting\\_en.pdf](https://protectchildren.ca/pdfs/C3P_ReviewingCSAMMaterialReporting_en.pdf).
- <sup>[14]</sup> Palmer, T. & Stacey, L., *Just One Click: Sexual Abuse of Children and Young People Through the Internet And Mobile Phone Technology* (Barkingside, UK: Barnardo’s, 2013); *See also United States v. Williams* (11th Cir.2006) 444 F.3d 1286, 1290 (“Our concern is not confined to the immediate abuse of the children depicted in these images but is also to enlargement of the market and the universe of this deviant conduct that, in turn, results in more exploitation and abuse of children.”).
- <sup>[15]</sup> *2024 Cyber Tipline Report*, Nat. Ctr. for Missing and Exploited Children (2024), <https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata>.
- <sup>[16]</sup> Jeff Horwitz & Katherine Blunt, *Instagram Connects Vast Pedophile Network*, WSJ (June 7, 2023, 7:05 AM), <https://www.wsj.com/articles/instagram-vast-pedophile-network-4ab7189>.
- <sup>[17]</sup> *Id.*
- <sup>[18]</sup> *FTC Proposes Blanket Prohibition Preventing Facebook from Monetizing Youth Data*, FED. TRADE COMM’N (May 3, 2023), <https://www.ftc.gov/news-events/news/press-releases/2023/05/ftc-proposes-blanket-prohibition-preventing-facebook-monetizing-youth-data>.

- <sup>[19]</sup> Clare Duffy, *Facebook Has Known It Has a Human Trafficking Problem for Years. It Still Hasn't Fully Fixed It*, CNN (Oct. 25, 2021), <https://www.cnn.com/2021/10/25/tech/facebook-instagram-app-store-ban-human-trafficking#:~:text=A%20report%20distributed%20internally%20in,accounts%20to%20help%20with%20detection.>
- <sup>[20]</sup> *Id.* (emphasis added).
- <sup>[21]</sup> Lara Putnam, *Facebook Has a Child Predation Problem*, WIRED (Mar. 13, 2022, 8:00 AM), <https://www.wired.com/story/facebook-has-a-child-predation-problem/>.
- <sup>[22]</sup> *Broken Promises: Sexual Exploitation of Children on Facebook*, TECH TRANSPARENCY PROJECT (March 3, 2020), <https://www.techtransparencyproject.org/articles/broken-promises-sexual-exploitation-children-facebook.>
- <sup>[23]</sup> *Id.* at 3.
- <sup>[24]</sup> *Id.*
- <sup>[25]</sup> Alexandra S. Levine, *How TikTok Live Became 'A Strip Club Filled With 15-Year-Olds'*, FORBES (Apr. 27, 2022), <https://www.forbes.com/sites/alexandralevine/2022/04/27/how-tiktok-live-became-a-strip-club-filled-with-15-year-olds/?sh=75efad7b62d7.>
- <sup>[26]</sup> Jeff Horwitz, *His Job Was to Make Instagram Safe for Teens. His 14-Year-Old Showed Him What the App Was Really Like*, WSJ (Nov. 2, 2023 9:00 PM), <https://www.wsj.com/tech/instagram-facebook-teens-harassment-safety-5d991be1.>
- <sup>[27]</sup> *Bejar Whistleblower Docs*, SOC. MEDIA VICTIMS, 4 (2023), <https://socialmediavictims.org/wp-content/uploads/2023/11/Pages-from-Bejar-Whistleblower-Docs.pdf>.
- <sup>[28]</sup> *Cyberbullying*, CA Office of the Attorney General, <https://oag.ca.gov/cybersafety/children/cyberbullying.>
- <sup>[29]</sup> Kamala D. Harris, *Attorney General Harris on Cyberbullying*, CA Office of the Attorney General (2011), <https://oag.ca.gov/cybersafety/online/opinion.>
- <sup>[30]</sup> Emily A. Vogels, *Teens and Cyberbullying 2022*, PEW RESEARCH CTR., (December 15, 2022), <https://www.pewresearch.org/internet/2022/12/15/teens-and-cyberbullying-2022/>.
- <sup>[31]</sup> Jean M. Twenge, *Increases in Depression, Self-Harm, and Suicide Among U.S. Adolescents After 2012 and Links to Technology Use: Possible Mechanisms*, 2 PSYCHIATRIC RSCH. AND CLINICAL PRAC. 1, (2020), <https://prcp.psychiatryonline.org/doi/full/10.1176/appi.prcp.20190015.>
- <sup>[32]</sup> Elizabeth J. Ivie et al., *A Meta-Analysis of the Association Between Adolescent Social Media Use and Depressive Symptoms*, 275 J. of Affective Disorders 165, 165–174 (2020), <https://tinyurl.com/bdzu6h8h>; Alan Mozes, *As Social Media Time Rises, So Does Teen Girls' Suicide Risk*, U.S. News (Feb. 16, 2021), <https://tinyurl.com/49hzmm9v>.
- <sup>[33]</sup> *Id.*
- <sup>[34]</sup> Angus Crawford & Bethan Bell, *Molly Russell Inquest: Father Makes Social Media Plea*, BBC (2022), <https://www.bbc.com/news/uk-england-london-63073489.>
- <sup>[35]</sup> *U.S. Surgeon General's Advisory: Protecting Youth Mental Health*, HHS, 8 (2021), <https://www.hhs.gov/sites/default/files/surgeon-general-youth-mental-health-advisory.pdf>.
- <sup>[36]</sup> Mike Allen, *Sean Parker unloads on Facebook: "God only knows what it's doing to our children's brains,"* AXIOS (Nov 9, 2017), <https://www.axios.com/2017/12/15/sean-parker-unloads-on-facebook-god-only-knows-what-its-doing-to-our-childrens-brains-1513306792.>
- <sup>[37]</sup> *Meta Platforms Gross Profit 2010-2025*, MACROTRENDS, <https://www.macrotrends.net/stocks/charts/META/meta-platforms/gross-profit#:~:text=Meta%20Platforms%20gross%20profit%20for%20the%20twelve%20months%20ending%20December,a%2019.25%25%20increase%20from%202022.> Snap, too, saw record growth: Stacy Jo Dixon, *Global annual revenue of Snap from 2015 to 2024*, STATISTICA (May 21, 2025), <https://www.statista.com/statistics/552702/snapchat-annual-revenue/#:~:text=In%202023%2C%20Snap%20Inc%20generated,annual%20revenue%20throughout%20recent%20years.>

<sup>[38]</sup> Eva Rothenburg, *Mark Zuckerberg Made \$29 Billion This Morning After Meta Stock Makes Record Surge*, CNN (February 2, 2024, 10:31 AM), <https://www.mercurynews.com/2024/02/02/mark-zuckerberg-made-29-billion-this-morning-after-meta-stock-makes-record-surge/>.

<sup>[39]</sup> *United States Senate Committee on the Judiciary "Protecting Our Children Online,"* Nat. Ctr. for Missing and Exploited Children, 3 (2023), <https://www.judiciary.senate.gov/imo/media/doc/2023-02-14%20-%20Testimony%20-%20DeLaune.pdf>.

<sup>[40]</sup> As mentioned, AB 2481 is set to take effect this upcoming January. It is similar to but in some sense broader than SB 1504. It requires platforms earning more than \$1 billion a year or has more than 100 million active monthly users to establish a reporting mechanism where certain "verified" users (school principals and mental health professionals) can report violations of terms and conditions and "social media-related threats" such as content that "promotes, incites, facilitates, or perpetrates any of the following: (1) Suicide. (2) Disordered eating. (3) Drug trafficking. (4) Substance abuse. (5) Fraud. (6) Human trafficking punishable pursuant to Section 236.1 of the Penal Code. (7) Sexual abuse. (8) Cyberbullying. (9) Harassment. (10) Distribution of harmful matter, as defined by Section 313 of the Penal Code [obscene matter]." The law additionally addresses what are labeled "severe risks" to children; namely, content that is "a social media-related threat that more likely than not will cause serious bodily or mental harm to a child.

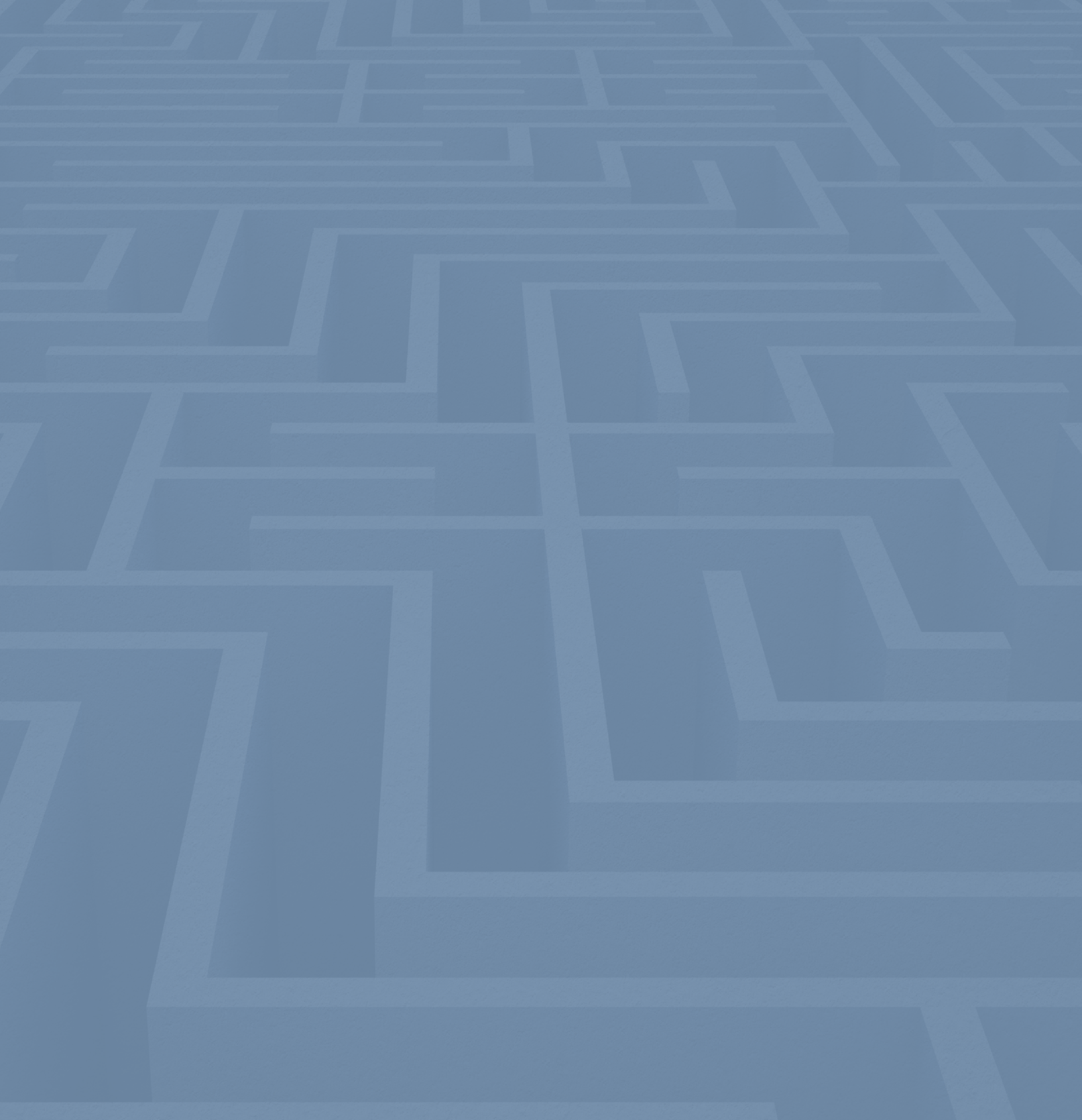
<sup>[41]</sup> California Business & Professions Code § 22589(b).

<sup>[42]</sup> Key language: (b) (1) A social media platform shall establish a prominent [note the requirement of "prominence"] mechanism within its internet-based service that allows any individual, whether or not that individual has a profile on the internet-based service, to report cyberbullying or any content that violates the existing terms of service related to cyberbullying.

<sup>[43]</sup> According to Pew Research Center, 95% of teens have access to a smartphone, more than access to laptops or desktops (90%) and gaming consoles (80%). Vogels, E. A., Gelles-Watnick, R., & Massarat, N., *Teens, Social Media And Technology 2022*, PEW RSCH. CTR., (2022), [https://www.pewresearch.org/wp-content/uploads/sites/20/2022/08/PI\\_2022.08.10\\_Teens-and-Tech\\_FINAL.pdf](https://www.pewresearch.org/wp-content/uploads/sites/20/2022/08/PI_2022.08.10_Teens-and-Tech_FINAL.pdf).

<sup>[44]</sup> In the "User Experience" sections, graphics and highlights have been added to the screenshots by CAI.

<sup>[45]</sup> **A note about Instagram's Teen Accounts.** Meta has claimed that Instagram's new Teen Accounts include safety features designed to protect minors from online exploitation. These features include tools like default privacy settings, restrictions on adult-to-teen direct messaging, and a combined "block and report" option within direct messages. Independent reviews have found that harmful, sexualized, and age-inappropriate content continues to appear on feeds on Instagram Teen Accounts, despite Meta's assurances. This highlights the importance of clear and conspicuous reporting tools on all accounts. Our review demonstrated that Teen Accounts, like adult accounts, lack clear and conspicuous reporting pathways for content appearing in Instagram feeds, leaving children and those trying to help them to navigate confusing, opaque systems that fail to meet the standards established by California law.



University of San Diego®  
**SCHOOL OF LAW**



CHILDREN'S ADVOCACY INSTITUTE