FILED UNDER SEAL

EXHIBIT 7

EXPERT REPORT OF DAMON L. McCOY, Ph.D.

State of New Mexico v. Meta Platforms, Inc., et al.

State of New Mexico, First Judicial District Court, Santa Fe County

No. D-101-CV-2023-02838

July 11, 2025

same features after they had already been long released. Efforts to keep users on, or bring them back to, Meta's platforms were often prioritized despite reservations or even objections expressed by Meta's internal integrity and product safety teams tasked with identifying and addressing the serious safety issues that often resulted.

These teams were often given inadequate resources to make meaningful changes before or after these products were unveiled, or, worse, were effectively told to stand down on implementing integrity or safety related changes that could interfere with the key growth and engagement metrics that have powered Meta to the social media and advertising behemoth it has become today, with over $160 billion in revenues in 2024 alone,[4] and a significant fraction of the world's population among the users of its platforms. The company has created a euphemism of "value" for its users, but this metric equates to short-term usage growth that was largely untethered to an understanding of whether its users felt that updates and new products were delivering real value.

Extensive internal documentation and correspondence demonstrates that Meta often stalled in fixing known integrity and safety issues until sufficient public and regulatory pressure had accumulated, and only then deploy watered-down or minimal integrity and safety interventions and, even then, often with what appears to be an eye towards forestalling significant impacts to financial metrics or blunting proposed external regulatory measures.

Once external pressure subsided, Meta sometimes attempted to roll back integrity and safety features if they reduced engagement or otherwise impacted short-term financial metrics. The narrow and oftentimes potentially deceptive metrics in Meta's publicly issued Community Standards Enforcement Reports (themselves only introduced as public scrutiny mounted on the company), as well as collaborations with cherry-picked "friendly" external researchers, fostered deniability and doubt about the scope and extent of the issues. At the same time, independent journalists and researchers' findings, and even findings from Meta own internal employees, were often undermined by account suspensions, PR statements, lobbying efforts, and legal threats.

These conclusions apply to a range of features across Meta's platforms, as explored in greater detail below. Unfortunately, this cycle appears to be starting anew with Meta's latest forays into generative AI and chatbots: Meta continues to publicly tout past product harm reduction reformulations which do not adequately address new integrity/safety issues, even as new problems have emerged in new product offerings and features, while existing problems linger. True to its founding motto, Meta will continue to move fast and break things, leaving behind a wake of product integrity and safety issues that could have been more effectively addressed.

1. Facebook

In February 2004, "TheFacebook" was launched by 19 year-old Mark Zuckerberg (and others) as a social network exclusive to Harvard students; throughout 2004 it expanded access to other universities and, in 2006, to the general public over the age of 13 (or at least those who

[4] *See, e.g.*, Meta Investor Relations, Meta Earnings Presentation Q4 2024, https://s21.q4cdn.com/399680738/files/doc_financials/2024/q4/Earnings-Presentation-Q4-2024.pdf (see page 4, showing $160,663,000,000 in 2024 advertising revenues).

collecting data from consenting users' own feeds and deleting third-party data immediately. Fearing litigation, AlgorithmWatch shut down the project in July 2021.[297]

Meta's CSER Reports

Meta's Community Standards define what types of content is prohibited on Meta's platforms.[298] In May 2018, Meta began issuing Community Standards Enforcement Reports ("CSER"), at first semi-annually, and then quarterly, which detail how the company enforces its content policies across its platforms.[299] Meta's CSERs claim to provide data on the prevalence of violating content, with a focus on minimizing views of harmful content, as measured through prevalence metrics (although they were not applicable for private messaging).[300]

While attempts at providing public transparency into the scope of the dangers present on Meta's platforms is certainly commendable, testimony from former employees indicates that these attempts were, at best, incomplete and potentially misleading. For example, former senior safety researcher Dr. Alison Lee testified that "what they were sharing in their CSER report was an – a deliberately incomplete picture in an underreporting of the rate of harmful content and experiences that were happening on our platform."[301] "I think by underreporting and diminishing the findings that were leaked, and underreporting the prevalence of that, that they prevented the public from understanding the rate and severity of the risks that are associated with being on the platform, particularly for young people."[302] Similar, former employee Arturo Bejar testified:

---

[297] *Id.*

[298] *See* Meta, *Community Standards*, https://transparency.meta.com/policies/community-standards/.

[299] Rosen Dep. Tr. 398:20-22, 101:1-3.

[300] Rosen Dep. Tr. 415:3-10, 528:11-529:10.

[301] Lee Dep. Tr. 457:22 – 458:1.

[302] Lee Depo Tr. 458:17-22; *see also, e.g.*, Bejar Dep. Tr. 228:23-229:16 ("I mean, when I looked at these community standards and other documentation that Instagram produces, to me it says this is an environment where your kid is not going to get unwanted sexual content. Your kid is not going to experience inappropriate contact. Your kid is not going to be exposed to graphically violent content or self-harm content. And I have spoken to the parents of many kids who, again, have lived harm experience. And one thing that was in common is they all told me I wasn't worrying about what my kid was getting recommended because they thought that part was going to be safe. They were worried about who their kid was talking to and what their kid was posting. So I believe it's absolutely critical for parents to know the substance of what Meta is recommending."); Bejar Dep. Tr. 321:10-18 ("What Meta has in its Transparency Center, it says a fraction of a percent, like point zero-something or other. Like, it's tiny, tiny numbers. And for something like eating disorder content, you would look at that and believe there is no way they are going to be recommending that to my kid. They seem to be really on top of it. And that is not the case in my experience or my testing."); Bejar Dep. Tr. 344:3-347:1.

> [T]here was a lot of emphasis put on what Meta calls its Transparency Center where they listed out different harms, the kind we have been talking about today. But if you talked about, for example, some nudity and sexual content, the number that was quoted is something called prevalence. And the prevalence of nudity was a fraction of a percent. While internal surveys, like the negative experiences [BEEF] survey, was saying one in five people experienced it in the last seven days. And so I believe that this Transparency Center was deeply misleading as to what is the likelihood that your kid is going to experience a certain harm. But the numbers -- I mean, I think to me the best example of this is on, like, self-suicidal ideation content, SSI.[303]

> And if you look at the Transparency Center around these dates around the topic of suicide and self-harm, what it says is the number is so small we don't even put it up here. And so as a parent, as a professional, as a person, I'm not sure how they can say that the number of suicidal -- SSI content on the platform is so small they can't report it when 12 percent of people are telling you -- of the people who responded to the [BEEF] survey, are telling you that they saw it in the last seven days.[304]

My review of Meta's internal documents and features, and my own experience as a researcher indicates that Meta does not disclose most of the limitations and incompleteness of its CSER reports. Some of the most significant gaps in Meta's CSER reports include the failure to better disclose internal data and awareness of the rate and prevalence of harmful content and behavior, particularly in Meta's messaging surfaces. Based on my expertise and experience, Meta's CSER reports misrepresent and downplay the likely true prevalence of harmful experiences, particularly for teen Instagram users, by both what they included and what they left out, especially regarding exposure to unwanted content and behavior.

For example, internal documents show that, as of November 15, 2021, Meta knew (and had measured internally) that "~1 in 100 IG teen users in the US are exposed to IIC [inappropriate interactions with children] convos (child grooming) with adults ***everyday***."[305] However, with regard to "Child Endangerment: Nudity and Physical Abuse and Sexual Exploitation," its CSER

---

[303] Bejar Dep. Tr. 204:9-25.

[304] Bejar Dep. Tr. 205:9-19.

[305] METANMAG-011-01633788 at -93 (emphasis added); *see also, e.g.*, Bejar Dep. Tr. 45:12-46:5 (discussing grooming); Bejar Dep. Tr. 227:3-20 ("If you create an environment where if you are a pedophile looking for young girls dancing in revealing ways, right, and you reward that by making it easy to find that content and recommending it, and you create an environment that takes young girls, including girls under 13, and you reward them for creating that kind of content, you are creating an environment that facilitates grooming, sexual endangerment, child pornography, account selling child pornography, which I found in the course of my testing.... What I found is were accounts of kids that appeared to be under 13, like 7, 8, 9, with a Cash App link on their profile. Or they would use the words selling. And so that's an environment, right, in which the worst kind of person on the internet are rewarded by the way the application is designed.")

report for approximately the same period (Q3 2021),[306] failed to disclose anywhere that a minor on Instagram would face approximately a 1% chance of exposure to grooming behavior each and every day they were on the platform, according to Meta's own internal research. Instead, the report blandly states that:

> We do not allow content that endangers children, such as content that contains nudity or physical abuse or content that sexually exploits children on Facebook and Instagram. When we find this type of violating content, we remove it, regardless of the context or the person's motivation for sharing it. We may also disable the account of the person who shared it, unless it appears the intent was not malicious (for example, to spread awareness of child exploitation).... "[w]e are working on estimating prevalence for child endangerment violations [for Instagram]. We will continue to expand prevalence measurement to more areas as we confirm accuracy and meaningful data."[307]

Nor does Meta CSER report explain how a child's measurable daily chance of being subjected to a grooming conversation does not constitute "child endangerment" on its platform. Meta's CSER report continues:

> We report apparent child exploitation to the National Center for Missing and Exploited Children (NCMEC), a nonprofit that refers cases to law enforcement globally, in compliance with US law. We choose to remove content depicting non-sexualized child nudity to reduce the potential for abuse of the content by others.... [and] "[w]e are working on estimating prevalence for child endangerment violations [for Instagram]. We will continue to expand prevalence measurement to more areas as we confirm accuracy and meaningful data."[308]

Despite touting the raw volumetric data of how many NCMEC CyberTip reports involving child endangerment were sent from Facebook, Instagram and Threads (a figure in the tens of millions) with graphs and charts to emphasize its efforts, Meta did not include the added context that many of these reports were simply inactionable or useless to law enforcement,[309] and again,

---

[306] METANMAG-005-00841862 at -874.

[307] METANMAG-005-00841862 at -875.

[308] *Id.*

[309] *See, e.g.*, METANMAG-007-00020093 (2019 NCMEC CyberTipline Report on Facebook and Instagram noting the fact that "submitted reports do not contain enough context or pertinent information to be actioned by law enforcement"; that inactionable reports caused NCMEC "CyberTipline analysts [to] spend a disproportionate amount of time identifying and manually processing these reports" which "takes time away from processing actionable, high valuable reports" and affects law enforcement's "triaging and takes time away from processing actionable reports."); METANMAG-002-00461193 (letter from NCMEC in early 2022 regarding "[f]ollow-up to Discussions Regarding Widespread Flaws in Meta's Reporting to the CyberTipline," including longstanding issues including law enforcement's receiving mass numbers of erroneous reports, NCMEC's concerns about the "severe impact" those errors are having its mission,

without the added context that teen and underage Instagram users regularly faced the chance of being subject to grooming conversations. The contemporaneous internal document noted above also indicated that, "Over the last 1 year, we detect, delete and report 36.37M pieces of content containing CEI to NCMEC, where 8 out of 10 come from Messenger,"[310] but did not disclose in their CSER report that the majority of detected CEI was from Messenger, again depriving parents, teens and the public of this knowledge of critical risks known to Meta.[311] The result of Meta's failure to publicly communicate this knowledge is that parents and teens were not warned about important internally known risks about using Meta's Messenger app.[312] Meta's failure to include this type of information in their CSER reports meant U.S. parents and teens were less able to evaluate the risks of using Instagram, while Meta was able to take credit for its purported increase in transparency and responsiveness to widespread concerns.

As I previously mentioned in my report, Meta intentionally reduced users' ability to report harmful content and conduct, including on its direct message surfaces Messenger and Instagram Direct.[313] Specifically, an April 2021 presentation regarding "Messenger, Instagram Direct, and

---

incomplete reports missing crucial information, and noting Meta's "apparent inability" to end those mistakes, noting "Meta must do better.")

[310] METANMAG-011-01633788 at 793; *see also, e.g.*, MT-IG-AG-NM-000038264 (internal "Content Governance and Enforcement Report (CSER) v2 -Transparency Pre-Read"received by Zuckerberg, which noted: "Explaining the nuances and types of content, and how we work with National Center for Missing and Exploited Children (NCMEC), will be critical to mitigating Comms/PR/Legal risks. We should also be prepared to deal with questions suggesting that the high number of take downs reflects a high prevalence of child exploitation on our platform," and that "the biggest challenge here is that … but people (especially policymakers and law enforcement) do think we create the opportunity for children to be groomed. Until we have really strong mechanisms for preventing inappropriate adult/minor contact, I think we risk generating more negative press than positive – especially when we do very little to keep U13s off our platform.")

[311] *See also, e.g.*, METANMAG-002-00032490 (Internal document where child safety employee notes that "[i]nternal child safety data analysis found that the daily number of minor victims having sex talk conversations in IGDM with adults is *38x compared to FB (~200k victims/day).* When normalized for the number of teens, the relative prevalence on Instagram compared to Facebook is about 5x," with another stating that "key findings around both comments and DMs: Children are regularly exposed to sexualizing content and comments – girls in particular have come to just expect that they're see[ing] this type of content and have these experiences," and noting that "another IG researcher who was recruiting teens for a messaging study met a potential recruit who described a predatory grooming message she received from an older man. The potential recruit was disturbed by this but also mentioned that she regularly gets these kinds of messages, usually at least once a day.")

[312] See, e.g., Bejar Dep. Tr. 206:5-11 ("But what is out there is profoundly misleading. And as a parent, this is a number that I would want to know so that I can take appropriate measures and have the right conversations with my kid.")

[313] *See* METANMAG-059-00001065.

WhatsApp Enforcement Summary" noted that "[r]eactive enforcement depends on review of reports," but that "[c]urrent resource constraints allow for review of ***just over half*** of submitted reports," and that "[i]ncreased reporting friction designed to limit inflow" had been designed into Messenger and Instagram Direct: "[w]e have ***deliberately added steps and difficulty*** to reporting flow to ***reduce the overall number***, increase % of actionable reports, and prioritize reporting of higher severity violations" because "[f]rictionless reporting ***would require 50-100x current review capacity***."[314]

This reduced ability to report represents a risk that is not included in the CSER reports. It also degrades the meaningfulness of several metrics included in the CSER reports. For instance, the rate of violating content "reported by users" was already very low, as noted in, for example, Meta's CSER Q3 2021 CSER report discussed above, which showed that: for "Jul – Sep 2021" (i.e., Q3 2021), for Facebook, including presumably Messenger, only "2.90%" of "violating content" related to "child nudity and physical abuse and "0.9%" of "child sexual exploitation" was "reported by users," while for Instagram, including presumably Instagram Direct Messages, "7.7%" of "violating content" related to "child nudity and physical abuse, and "3.7%" of "child sexual exploitation" content was "reported by users."[315] The balance for each category was purportedly "found and actioned by us [Meta]," although roughly contemporaneous internal documents state that Meta's "current CEI [child exploitation imagery] classifiers," or the automated tool Meta used to proactively detect (without user reporting) that type of content, "are <20% accurate," and that "improving these classifiers was deprioritized last year."[316]

Accordingly, the reassuring numbers provided in Meta's CSER report are suspect given that Meta intentionally introduced friction into the reporting process, meaning additional user reports were actively discouraged, and that its tools for proactively catching abuse on its platforms were woefully inaccurate. This intentional effort to reduce reporting, or Meta's silence on its issues with actively detecting violating content in messaging content, likely resulted in disingenuously increasing the proactive metric in the CSER reports, despite that rate itself being illusory. These truths regarding Meta's CSER statistics were, to the best of my knowledge, never publicly mentioned or referenced by Meta.

In my opinion, Meta's CSER reports did not make a good faith "best effort at sharing our policies and our enforcement and our progress openly with the public."[317] Rather, they appear to be a strategically curated set of information that downplayed the risks of using Meta's products. Overall, I found from my review of documents that Meta had internal knowledge of many significant risks to children using its apps that it chose not to include in its CSER reports. By

---

[314] METANMAG-059-00001065 at -72 (emphases added).

[315] See, Meta, *Child Endangerment: Nudity and Physical Abuse and Sexual Exploitation: proactive rate*," tables showing "Of the violating content we actioned for child endangerment, how much did we find and action before people reported it?," https://transparency.meta.com/reports/community-standards-enforcement/child-nudity-and-sexual-exploitation/facebook/

[316] METANMAG-079-00005582 at -584.

[317] Rosen Dep. Tr. 404:1-4.

omitting known internal risks, Meta's CSER reports likely gave parents and children a misleading sense of safety when deciding whether to use Meta's apps.

III. Conclusion

Through my analysis, I have identified that Meta's executives had knowledge from internal research, investigative journalism, and external research of a wide array of safety issues linked to product design decision which were implemented to increase short-term usage metrics or reduce expenses. Tragically, Meta's increased ad revenue and cost savings from its dangerous product design decision often came at the expense of its users' safety.

Meta executives publicly stated that its products were safe for teens to use and that user safety was being prioritized while in many instances withholding knowledge of unaddressed safety issues and continuing to prioritize usage growth over safety. Internal research and communication from Meta executives often acknowledged the harms experienced by its users and in particular its child and teen users. These internal researchers devised technically and practically feasible design changes that they identified would improve the safety of Meta's product offerings. Unfortunately, Meta often did not work in good faith to address safety issues. Recommended design changes to improve safety were often never implemented or delayed until large amounts of public pressure were applied. In many instances, watered down versions of these safety design changes, which were not championed by internal research, were adopted.

Instead of adopting a culture of safety and working in good faith to improve safety, Meta choses to conceal and downplay known safety issues through a number of methods. These included releasing limited incomplete transparency reports, interfering with internal and external integrity and safety research, and publicly undermining or threatening researchers.

Meta's executives often abdicated their responsibility to design safe products and are exposing their users to avoidable safety risks. Tragically, Meta chooses to move fast and break its users, which included many teens and children in New Mexico.[318]

* * * * *

[318] *See, e.g.*, State of New Mexico's Amended Complaint for Abatement and Civil Penalties and Demand for Jury Trial, at ¶¶ 246-255 (describing specifics of how "Children in New Mexico have fallen victim to distribution and solicitation of CSAM or CSEC on Meta's platforms.")