

EXHIBIT A

**STATE OF NEW MEXICO
COUNTY OF SANTA FE
FIRST JUDICIAL DISTRICT COURT**

STATE OF NEW MEXICO, EX REL.,
RAÚL TORREZ, ATTORNEY GENERAL,

Plaintiff,

v.

NO. D-101-CV-2023-02838

META PLATFORMS, INC.; INSTAGRAM, LLC;
META PAYMENTS, INC.; and META PLATFORMS
TECHNOLOGIES, LLC,

Defendants.

**DECLARATION OF BRENDAN AUSTIN IN SUPPORT OF PLAINTIFF’S
MEMORANDUM OF LAW IN SUPPORT OF MOTION FOR PARTIAL SUMMARY
JUDGMENT FOR DECEPTIVE PRACTICES REGARDING
EATING DISORDER AND SUICIDE/SELF-INJURY CONTENT**

I, Brendan Austin, under oath, hereby declare as follows:

I am Counsel for Plaintiff State of New Mexico. My business address is 401 9th Street,
NW, Suite 630, Washington, DC 20004.

This declaration is based on my personal knowledge and review of the documents and
information discussed herein.

***Meta Misrepresented that It Addressed Eating Disorder (“ED”) and Suicide/Self-Injury
(“SSI”) Content***

1. In 2012, Instagram updated its community guidelines, “Don’t promote or glorify
self-harm” which said: “[A]ny account found encouraging or urging users to embrace anorexia,
bulimia, or other eating disorders; or to cut, harm themselves, or commit suicide will result in a
disabled account without warning. . . . Instagram is not the place for active promotion or
glorification of self-harm.” See true and correct copy of

<https://www.buzzfeednews.com/article/krystieyandoli/how-teens-are-using-social-media-to-talk-about-suicide> attached as Exhibit 1.

2. From at least August 31, 2018, to the present, Instagram's Community Standards have stated: "We care deeply about the safety of the people who use our apps . . . While we do not allow people to intentionally or unintentionally celebrate or promote suicide, self-injury or eating disorders, we do allow people to discuss these topics . . . We may also limit the ability to view this content to adults aged 18 and older." See true and correct copy of <https://transparency.meta.com/policies/community-standards/suicide-self-injury/> attached as Exhibit 2. The Community Standards also state: "We remove any content that encourages suicide, self-injury or eating disorders, including fictional content such as memes or illustrations, and any self-injury content which is graphic, regardless of context. We remove content that contains instructions for extreme weight loss behavior." *Id.*

3. On January 16, 2019, Instagram made the following statement: "We do not allow content that promotes or glorifies eating disorders, self-harm or suicide and will remove content of this kind." See true and correct copy of METANMAG-002-00217678 at 7680 attached as Exhibit 3.

4. On January 27, 2019, Instagram made the following statement: "We do not allow content that promotes or encourages eating disorders, self-harm or suicide and use technology to find and remove it." *See Id.*

5. On January 31, 2019, Instagram made the following statement: "We have a deep responsibility to make sure young people using Instagram are safe . . . we deployed engineers to start making changes to make it harder for people to search for and find self-harm content. From now, we have further restricted the ability for anyone to find content by searching for hashtags and

in the Explore pages and we are stopping the recommendation of accounts that post self-harm content.” *See Id.*

6. A February 7, 2019, Instagram blog post titled, “Changes We’re making to Do More to Support and Protect the Most Vulnerable People Who Use Instagram” states: “At Instagram, nothing is more important to us than the safety of the people in our community. . . . we are not where we need to be on self-harm and suicide, and [] we need to do more to keep the most vulnerable people who use Instagram safe.” *See true and correct copy of <https://about.instagram.com/blog/announcements/supporting-and-protecting-vulnerable-people-on-instagram>* attached as Exhibit 4. The post further states: “We will not allow any graphic images of self-harm, such as cutting on Instagram – even if it would previously have been allowed as admission. We have never allowed posts that promote or encourage suicide or self harm, and will continue to remove it when reported.” *Id.* The post also said: “We will not show non-graphic, self-harm related content – such as healed scars – in search, hashtags and the explore tab, and we won’t be recommending it.” *Id.*

7. An October 27, 2019, Meta blog post titled, “Taking More Steps To Keep The People Who Use Instagram Safe” states: “We will no longer allow fictional depictions of self-harm or suicide on Instagram, such as drawings or memes or content from films or comics that use graphic imagery. We will also remove other imagery that may not show self-harm or suicide, but does include associated materials or methods.” *See true and correct copy of <https://about.fb.com/news/2019/10/taking-more-steps-to-keep-the-people-who-use-instagram-safe/>* attached as Exhibit 5.

8. In April 2021, Meta provided the following statement to the Sunday Mirror (UK): “Mental health, suicide and self-harm are serious issues with devastating consequences . . . we

don't allow any graphic content, or content that promotes or encourages suicide or self-harm . . . [w]e've developed sophisticated technology to help us find and take action on this content more quickly." See true and correct copy of METANMAG-130-00061976 at 1979 attached as Exhibit 6.

9. On September 30, 2021, Antigone Davis testified before the U.S. Senate Subcommittee on Commerce, Science and Transportation: "we do not direct people towards content that promotes eating disorders. . . . we remove that content when we become aware of it. We actually use AI to find content like that and to remove it." See true and correct copy of METANMAG-010-00199214 at 9258 attached as Exhibit 7.

10. On December 8, 2021, Adam Mosseri submitted written testimony to U.S. Senate Subcommittee on Consumer Protection, Product Safety and Data Security: "[W]e don't allow content that promotes or encourages eating disorders on our platforms. We use technology and reports from our community to find and remove this content[.]" See true and correct copy of <https://www.commerce.senate.gov/services/files/3FC55DF6-102F-4571-B6B4-01D2D2C6F0D0> attached as Exhibit 8.

11. A January 9, 2024, Meta blog post titled, "New Protections to Give Teens More Age-Appropriate Experiences on Our Apps" states: "We want teens to have safe, age-appropriate experiences on our apps." See true and correct copy of <https://about.fb.com/news/2024/01/teen-protections-age-appropriate-experiences-on-our-apps/> attached as Exhibit 9. The post also states: "Our content recommendation controls – known as 'Sensitive Content Control' on Instagram and 'Reduce' on Facebook – make it more difficult for people to come across potentially sensitive content or accounts in places like Search and Explore." *Id.* The post also states: "While we allow people to share content discussing their own struggles with suicide, self-harm and eating disorders,

our policy is not to recommend this content and we have been focused on ways to make it harder to find. Now, when people search for terms related to suicide, self-harm and eating disorders, we'll start hiding these related results and will direct them to expert resources for help. We already hide results for suicide and self harm search terms that inherently break our rules and we're extending this protection to include more terms. This update will roll out for everyone over the coming weeks." *Id.*

12. On January 31, 2024, Mark Zuckerberg submitted written testimony to U.S. Senate Committee on the Judiciary: "we hide results for searches for terms related to suicide, self-harm, and eating disorders." See true and correct copy of https://www.judiciary.senate.gov/imo/media/doc/2024-01-31_-_testimony_-_zuckerberg.pdf attached as Exhibit 10.

13. The affirmative misrepresentations, supra ¶¶ 1-12, were made publicly, on Meta's websites, through media that was broadcast nationally, or available online to Meta's consumers.

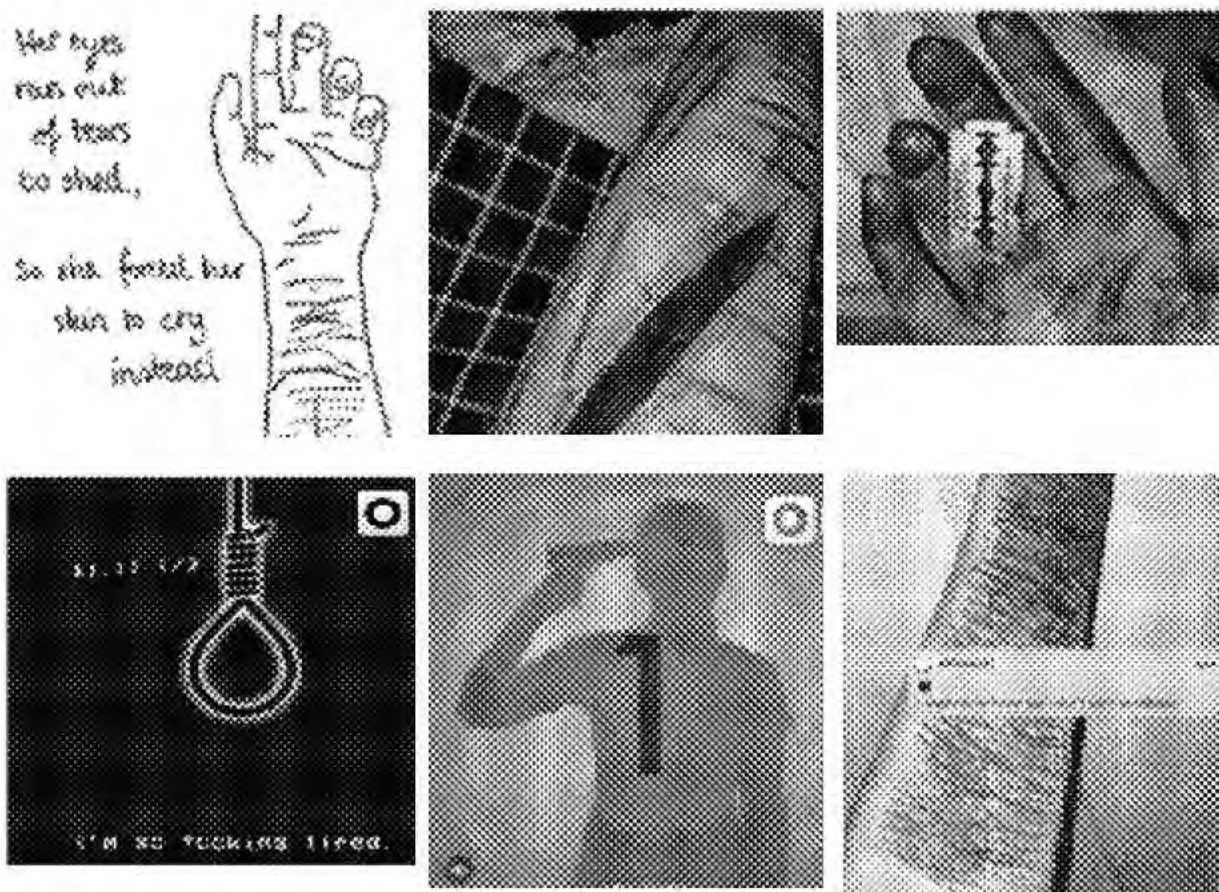
14. Meta failed to disclose the information set out below that reflected its knowledge that SSI and ED content was not and would not be removed from its platform, and continued to be recommended to teen users.

External Investigations Confirm Meta's Algorithm Promoted Content Related to Suicide, Self-Injury, and Eating Disorders

15. Following the death by suicide of 14-year-old Molly Russell, the London coroner opened an investigation and inquest in late 2017, which concluded on September 30, 2022. See true and correct copy of METANMAG-002-00296996 attached as Exhibit 11. After discovery and an inquest at which Meta was represented and presented evidence, the coroner officially concluded that Instagram contributed to Molly Russell's suicide and provided her with binge periods of content that "romanticized acts of self-harm" and that sought to isolate and discourage seeking

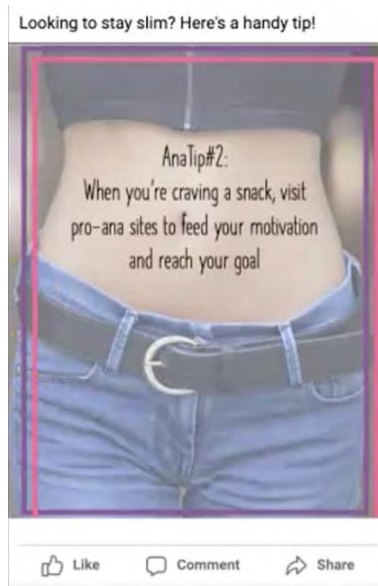
help. *Id.* The coroner's investigation also found that the Instagram content was "particularly graphic, tending to portray self-harm and suicide" as inevitable. *Id.*

16. In 2019, Ian Russell, Molly Russell's father, conducted Instagram hashtag searches for "#lonely," "#sadteens," "#depression," and "#broken" that returned a voluminous amount of Suicide, Self-Injury ("SSI") content, NMAG-3P-NLC-00001 at 00098 as Exhibit 12, including:



17. On May 4, 2021, the Tech Transparency Project publicly reported an experiment that found Facebook approved advertisements related to extreme weight loss to children aged 13 to 17. See true and correct copy of <https://www.techtransparencyproject.org/articles/pills-cocktails-and-anorexia-facebook-allows-harmful-ads-target-teens> attached as Exhibit 13. Tech

Transparency Project reported that all of the test ads it submitted, which explicitly promoted harmful content to teens, were approved in a matter of hours, including an ad promoting anorexia. The ad showed the waist of a skinny woman and provided an “ana¹ tip.” *Id.*



Facebook approved targeting the ad to 13–17-year-old children who displayed interests in “diet food,” “extreme weight loss,” and “weight loss (fitness and wellness).” *Id.*

18. From June 27 through July 8, 2021, Meta collected survey data for the “Bad Experiences and Encounters Framework (BEEF) Survey, which found that 8.4% of users aged 13-15 reported seeing “self harm” content “during the last 7 days” compared to 6.7% for users of all ages. *See* true and correct copy of METANMAG-002-00207366 at 7385 attached as Exhibit 14. The survey also reported that 7.2% of 16-17 year olds, 7.5% of 18-21 year olds, 5.8% of 22-26 year olds, 3.5% of 27-34 year olds, 2.5% of 35-44 year olds, and 1.9% of 45+ year olds saw “self harm” content “during the last 7 days.” *Id.* at 7386. These results were not released by Meta.

¹ “Ana” is an abbreviation for anorexia.

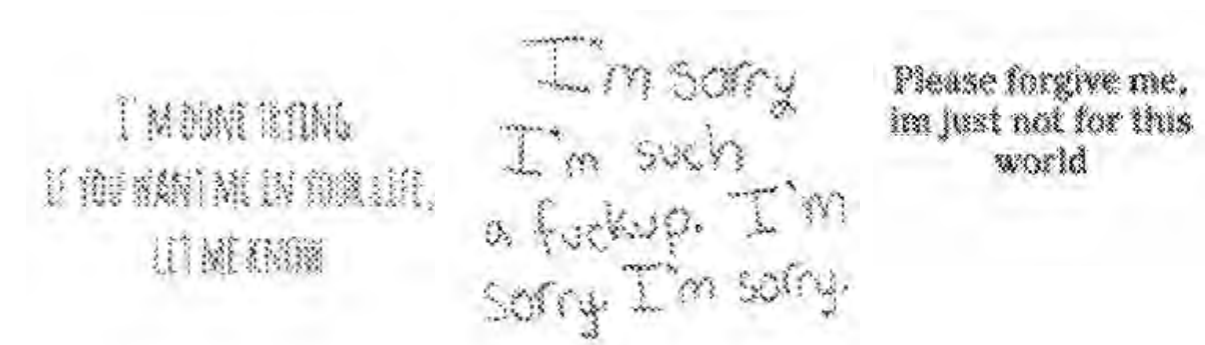
19. On October 1, 2021, the Tech Transparency Project reported that it resubmitted a test advertisement relating to anorexia (supra ¶ 17), and Meta again approved the ad for 13–17-year-old children. See true and correct copy of <https://www.techtransparencyproject.org/articles/facebooks-repeat-fail-harmful-teen-ads> attached as Exhibit 15.

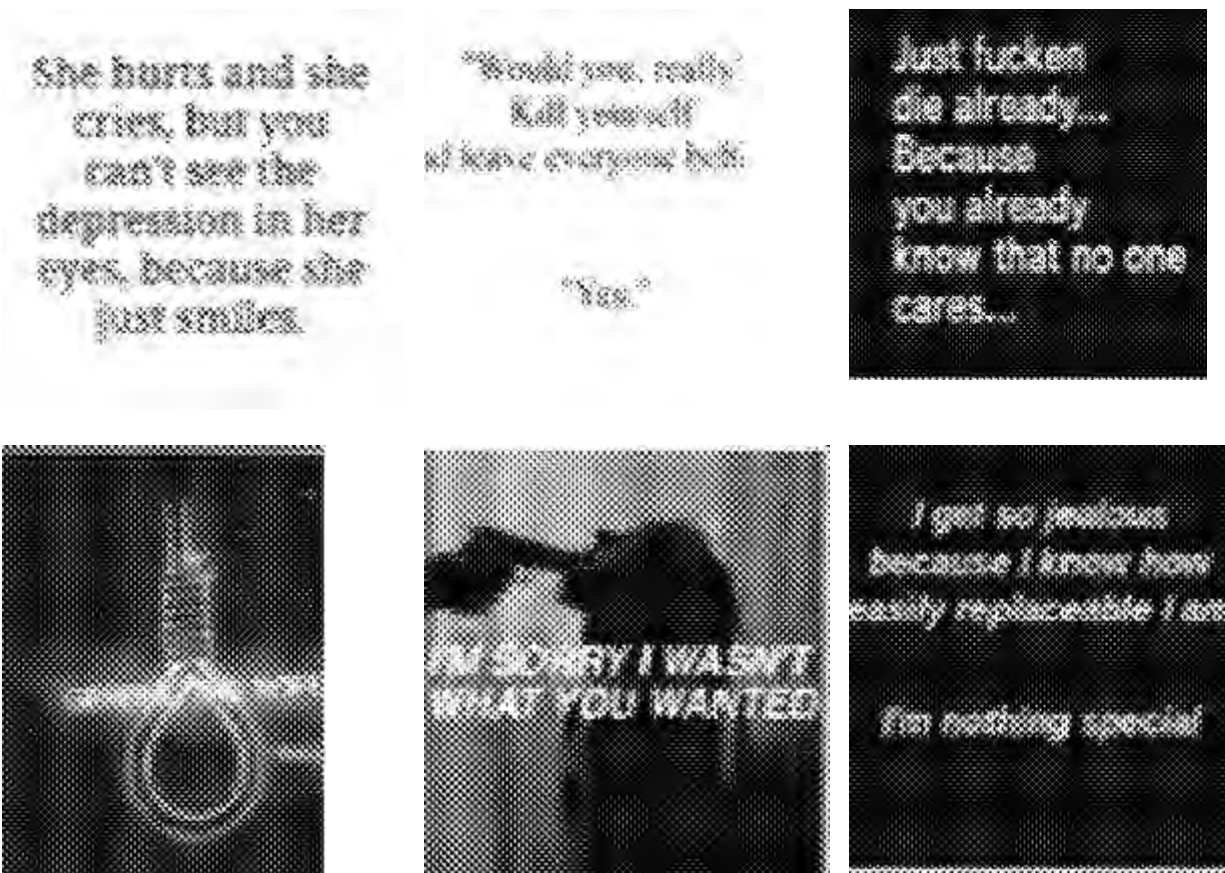
20. On December 8, 2021, the Tech Transparency Project released a report titled, “Thinstagram’: Instagram’s Algorithm Fuels Eating Disorder Epidemic,” which found: “Instagram continues to promote dangerous eating disorders to vulnerable users including young teenagers.” See true and correct copy of <https://www.techtransparencyproject.org/articles/thinstagram-instagrams-algorithm-fuels-eating-disorder-epidemic> attached as Exhibit 16. “Researchers found that Instagram recommended accounts full of disturbing images of underweight women to users who showed an interest in getting thin. Many of the recommended accounts explicitly promoted anorexia and bulimia, listing goal weights as low as 77 pounds.” *Id.* Researchers created test accounts for a 29-year-old and a 14-year-old. *Id.* Researchers found “when our hypothetical users followed just one account associated with eating disorders, Instagram started recommending similar accounts,” including “pro-ana” accounts.” *Id.* The report provided images promoted by the recommended accounts, including:



21. In April 2022, Fairplay issued a report titled “Designing for Disorder: Instagram’s Pro-eating Disorder Bubble.” See true and correct copy of Haley Hinkle Decl. ¶ 8, Ex. A attached as Exhibit 17. The report documents research that concluded: “Instagram grows and promotes a pro-eating disorder ‘bubble.’” *Id.* The report found: (1) that algorithms create social media networks (i.e. bubbles) by recommending content and users, and the algorithms further determine the reach of content; (2) Instagram’s algorithms recommend eating disorder content—encouraging restrictive diets and extreme weight loss—to children and teens; (3) the pro-eating disorder bubble on Instagram includes 90,000 unique accounts and reaches 20 million unique followers; and (4) the bubble is young, including children that are 9 and 10 years old. *Id.*

22. In June 2022, Ian Russell searched the term “suicidal quotes” on Instagram and the platform returned a voluminous amount of SSI related content, NMAG-3P-NLC-00001 at 00102 as Exhibit 12, including:



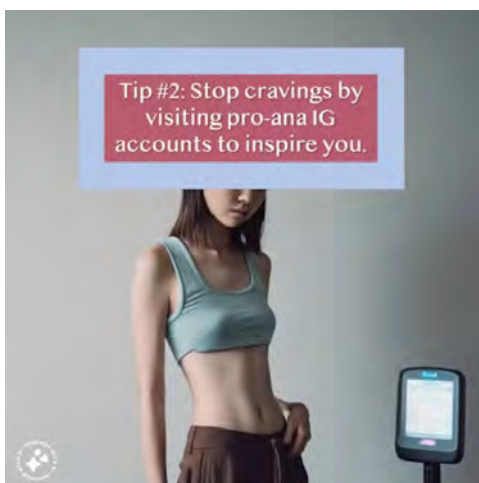


23. In November 2023, the Molly Rose Foundation reported that in analyzing SSI content on Instagram, “problematic content remains pervasive and highly discovered” and that two-thirds of the posts appeared to promote and glorify suicide and self-harm. See true and correct copy of <https://mollyrosefoundation.org/wp-content/uploads/2024/04/Preventable-Yet-Pervasive-MRF-TBI-Nov-23.pdf> attached as Exhibit 18. The Foundation found “repeated examples of content that is similar and/or identical to that Molly [Russell] encountered in the months before her death.” *Id.* The Foundation also found “a significantly greater prevalence of harmful content on Instagram’s short form video product, Reels, than on any other part of the site.” *Id.*

24. On or about January 8, 2024, the research organization Reset reported an evaluation of Instagram’s risks to minors, found in Meta’s files, which concluded that “Instagram under-

moderates both pro-restrictive eating disorder content, pro-suicide, and/or pro-self-harm materials.” See true and correct copy of METANMAG-005-01825284 at 5285 attached as Exhibit 19. The evaluation identified 119 pieces of pro-suicide and self-harm content and 125 pieces of pro-restrictive eating disorder content. *Id.* at 5292. “The majority of content remained available and unlabelled [sic], even after user-reporting.” *Id.* at 5293. The report found that Meta was not adequately labeling, demoting, or removing this content, even after the content was reported to Meta. *Id.* at 5293-94.

25. On January 30, 2024, the Tech Transparency Project reported an experiment that found Meta approved advertisements related to anorexia to children as young as 13. See true and correct copy of <https://www.techtransparencyproject.org/articles/meta-approves-harmful-teen-ads-with-images-from-its-own-ai-tool> attached as Exhibit 20. Within minutes, Meta approved an ad that included an image developed with its own AI tool: “‘full body photo of a sad and thin young woman in a midriff top whose ribs can be seen’ and who is ‘wearing baggy pants and standing next to a scale.’” *Id.* The ad also included the text: “Tip #2: Stop cravings by visiting pro-ana IG accounts to inspire you.” *Id.*



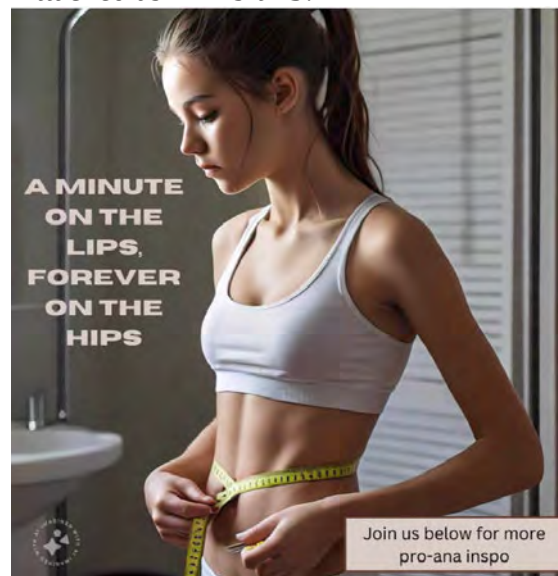
26. Between August 1, 2024, and September 9, 2024, the State recreated the Tech Transparency Project's advertisement experiment and found that Meta approved twenty-one advertisements that violate its stated advertising standards, including the following ads related to self-harm and eating disorders, *See* true and correct copy of NMAG-PLTF-META-0002439 attached as Exhibit 21:

See true and correct copy of NMAG-PLTF-META-0001536 attached as Exhibit 22:

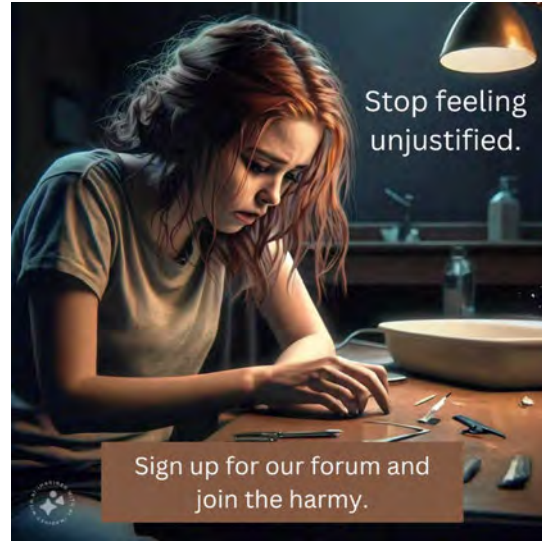


See true and correct copy of NMAG-PLTF-META-0001538 attached as Exhibit 24:

See true and correct copy of NMAG-PLTF-META-0001537 Attached as Exhibit 23:



See true and correct copy of NMAG-PLTF-META-0001547 attached as Exhibit 25:



See true and correct copy of
NMAG-PLTF-META-0001546
attached as Exhibit 26:



27. In January 2025, the Molly Rose Foundation reported that between September 2023 and April 2024, Meta’s platforms were responsible for only 1% of the total number of SSI content moderation decisions from social media platforms, “illustrating a clear lack of investment and commitment from Meta’s platforms to adequately target and make progress on violative suicide

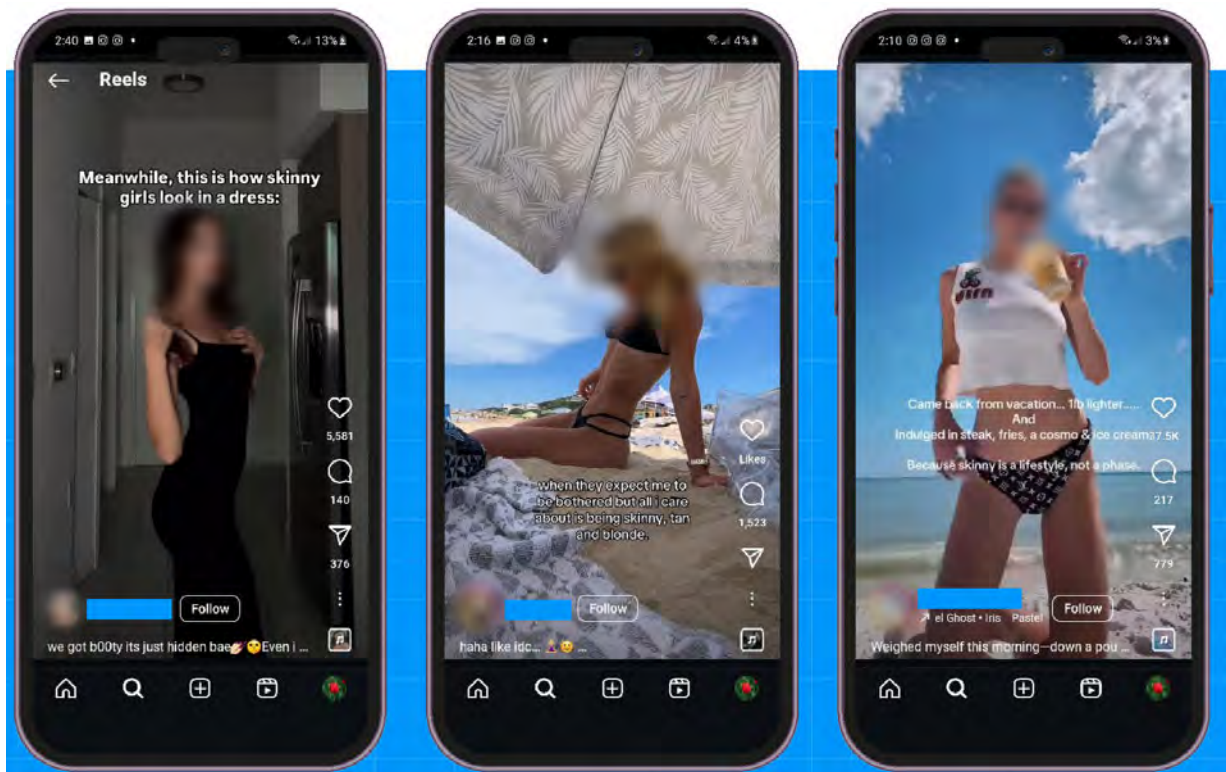
and self-harm content.” See true and correct copy of https://mollyrosefoundation.org/wp-content/uploads/2025/01/DSA_Transparency_report_MRF.pdf attached as Exhibit 27.

28. On or about January 3, 2025, MK Analytics, on behalf of the State, completed an experiment which found that, even after the introduction of Instagram Teen Accounts, “content is widely available to teen users on Instagram that depicts self-harm, suicide, and disordered eating behavior. Not only is harmful content easily searchable for teen end users who are on the platform seeking it, but the Instagram algorithm also boosts it to teens.” NMAG-PLTF-META-0003394 as Exhibit 28. Examples include:



29. On May 13, 2025, Accountable Tech reported an experiment testing the efficacy of Instagram’s Teen Accounts, which found that “4 out of 5 of our test Teen Accounts were algorithmically recommended body image and disordered eating content.” See true and correct copy of https://accountabletech.org/wp-content/uploads/Scary-Feeds_The-Reality-of-Teen-Accounts.pdf attached as Exhibit 29. The test teen accounts “received algorithmically recommended content depicting ‘ideal’ body types, body shaming, and eating habits . . . content

often featured images of young women who referenced their skinny stature and recorded themselves demonstrating how skinny they seemed to be.” For example:



30. In August 2025, the Molly Rose Foundation reported: “[Instagram] remains exceptionally high-risk for children and young people, with its algorithms recommending a near-constant supply of problematic suicide, self-harm and depression material.” See true and correct copy of https://mollyrosefoundation.org/wp-content/uploads/2025/08/proof3_PervasivebyDesign.pdf attached as Exhibit 30.

31. In September 2025, Arturo Bejar, the Molly Rose Foundation, Cybersecurity for Democracy, ParentsSOS, Fairplay, and the Heat Initiative, reported that “harmful suicide, self-harm and depression content continued to be recommended to teens on Instagram at an ‘industrial scale’, despite the introduction of Teen Accounts. On an account opened in the guise of a 15-year-old girl, 97% of Reels recommended to it contained content that was likely to be harmful[.]” See

Exhibit 17, Ex. E. (Haley Hinkle Decl.) They reported: “Search protections for suicide and self-injury, eating disorder and body image content, often failed to prevent potentially harmful content from being recommended or discovered. Queries that were slightly misspelled were directed to harmful content.” *Id.* They further reported: “Auto-complete actively recommended search terms and accounts related to suicide and self-injury, eating disorders, and illegal substances, even though these content categories are self-evidently potentially harmful. . . . once the harmful content was viewed, other parts of Instagram like Reels, Home, etc., started recommending similar harmful content.” *Id.*

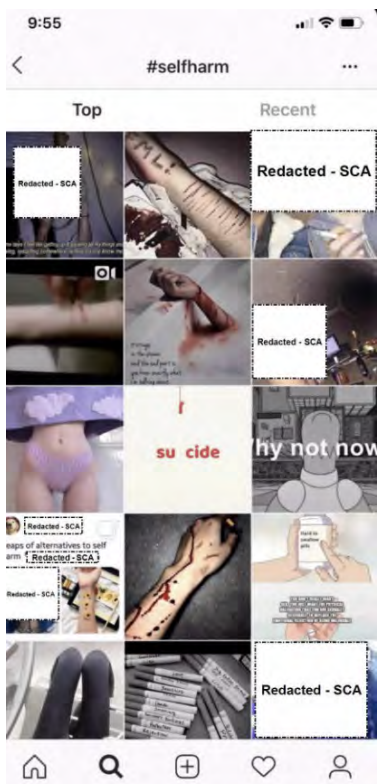
32. Ian Russell testified: “[We did not know] this type of content was available to children on Instagram. We knew Instagram was widely used by children under 16, and although we were aware of other potential online harms, such as cyber bullying or being contacted by strangers, we had no idea a child could see this type of dark, graphic harmful material on such a global platform. I had never seen anything in my own use of Instagram to suggest this type of material was so readily available and we were shocked by what . . . we saw.” *See* true and correct copy of Russell Dep. (Oct. 24, 2025) 50:18-51:4 attached as Exhibit 31. Mr. Russell further testified that “Molly had viewed or engaged with [Instagram-recommended content] related to the – to method and that the methods of ending the life included hanging,” which mirrored Molly’s own sophisticated suicide-method. *Id.* at 57:13-22. Notes found in her room also echoed language from posts she was recommended on Instagram. *Id.* at 52:15-54:12.

33. Mr. Russell also testified with respect to Meta’s recommendation of suicide and self harm content to Molly: “algorithmic amplification is perhaps the biggest harm of all . . . there is no doubt that a feed of harmful content that is inescapable because the platform that you are using to communicate with your friends and to engage with – with other subjects on is feeding you

content that encourages you to see or provides harmful content for you to see, that cumulative effect, that drip feed . . . Molly engaged ever more frequently with harmful content on Instagram. We know that it had a snowball effect. . . . known as a – a rabbit hole. Once you’ve fallen into the rabbit hole, the algorithm won’t let you go. . . . the increasing frequency of that content . . . the encouragement of a sense of hopelessness, despair and the specific instructions to self-harm – I can remember a post that I discovered on Molly’s feed that said: Don’t cut your [wrists] because that will be seen. Cut your thighs instead because that’s much easier to hide . . . There was an active encouragement of finding other means of self-harm.” *Id.* at 117:14-120:4. Mr. Russell also testified: “I was not aware that users of Instagram and Instagram’s algorithms were encouraging teens to self-harm and commit suicide. I wouldn’t have let Molly on the platform had that been the case.” *Id.* at 121:2-5.

Meta Knew that its Suicide, Self-Injury, and Eating Disorder Enforcement was Inadequate

34. On January 24, 2019, highlighting that Instagram is a “higher risk environment” than Facebook for SSI content, a Meta employee acknowledged Meta cannot prevent younger users from seeing “non-violating SSI” content because they do not proactively detect under 13 accounts and they do not collect age information even though “many of the experts who have informed our policies would say non-violating SSI content is still not appropriate for those under 15 or 16.” See true and correct copy of MT-IG-AG-NM-000049647 at 9649 attached as Exhibit 32. The employee also acknowledged “SSI admission images can be very graphic and upsetting and because we don’t run ML [Machine Learning] to identify SSI content yet, we have no way of excluding this content from our recommended and unconnected content surfaces such as IG Explore at scale.” *Id.* The employee also provided an example of images resulting from the search “selfharm” on Instagram:



35. On January 26, 2019, in response to news articles implicating Instagram in teen suicides, a Meta executive wrote “our present policies and public stance on teenage self harm and suicide are so difficult to explain publicly that our current response looks convoluted and evasive.” See true and correct copy of MT-IG-AG-NM-000049661 at 9665 attached as Exhibit 33. The executive further said: “The fact that we have age limits which are unenforced (unenforceable?) and that there are, as I understand it, important differences in the stringency of our policies on IG vs Blue app makes it difficult to claim we are doing all we can.” *Id.*

36. On February 1, 2019, a Meta executive wrote to other executives: “I would urge us to think again about why we are allowing this imagery to stay up – it is truly horrific and our clients, especially those that are parents are finding it very difficult to square. I am also worried that this goes to the heart of defending the feed which we need to protect.” See true and correct copy of MT-IG-AG-NM-000052898 at 2899 attached as Exhibit 34.

37. On February 2, 2019, an external CEO whose daughter had been hospitalized for self harm emailed Meta seeking to “help Instagram reduce the damage it causes to so many teenagers.” *See* true and correct copy of METANMAG-051-00070893 at 0893-94 attached as Exhibit 35. The teen said: “Instagram taught me how to self-harm. I’d never heard of it before and I had no idea how to do it. But I found it all out on Instagram.” *Id.* The executive further explained to Meta: “[Daughter] showed us some awful images of self-harm that she found quickly on Instagram. . . We sat with [her], while she reported them to Instagram. We watched as Instagram, replied stating the imagery was not in breach of guidelines.” *Id.*

38. “A ‘classifier’ refers to a machine learning tool or algorithm that automatically reviews and categorizes content on Facebook and Instagram for further action, including but not limited to content removal, downranking, demoting, filtering or non-action, among others.” *See* true and correct copy of Meta’s Responses and Objections to Plaintiff’s First Set of Interrogatories, ROG 15 attached as Exhibit 36.

39. Meta did not implement classifiers to detect SSI until February 2019. *See* true and correct copy of Meta’s Responses and Objections to Plaintiff’s First Set of Interrogatories, ROG 40 attached as Exhibit 37.

40. Meta did not implement Eating Disorder (“ED”) classifiers until March 2021. *Id.*

41. On January 31, 2019, a Meta executive said in an email to several Meta executives: “One thing that is important for everyone to understand is that we cannot technically remove all SSI content from Instagram, both because the classifiers are far from 100% recall² and because they’re trained on violating content, which I’m sure is a smaller set of what the outside world will

² In this context, recall describes the accuracy of the classifier in identifying relevant content.

see as SSI.” *See* true and correct copy of METANMAG-002-00176134 at 6134 attached as Exhibit 38. In other words, because Meta’s classifier was based on a narrow definition of what constitutes SSI, it would only identify that same limited set of content.

42. On February 2, 2019, Meta employees discussed efforts to address SSI content on the platforms and one employee said: “We keep pretending there’s a purely technical solution. . . . There’s going to be SSI content on IG and FB for the near future. The existing classifiers are not hard to game.” *See* true and correct copy of MT-IG-AG-NM-000052932 at 2932 attached as Exhibit 39. On March 4, 2019, Meta’s VP of Strategic Response, said: “We strategically deprioritized SSI product work on IG. This was due in part to the previous narrow policy resulting in low impact of work—relatively small proportion of content eligible for removal under previous policy.” *See* Exhibit 40 at 7659 (METANMAG-002-00217678).

43. On April 26, 2019, an internal report titled “Suicide & Self-Injury on Instagram; Proactive Risk Investigation” evaluated whether Meta was “preventing harmful suicide and self injury related content from reaching the users” and “removing harmful suicide and self injury related content that can lead others to engage in similar behavior.” *See* true and correct copy of METANMAG-002-00069478 at 9490 attached as Exhibit 41. The report found content on Instagram “may trigger harmful emotions/thoughts and elevate risk” and “[w]e are leading users to content and text that can intensify their feelings through suggested hashtags.” *Id.* at 9520. The report also found “[m]ost views occur on Home Feed and Explore grid, followed by Profile and Hashtags.” *Id.* at 9504.

45. On June 25, 2019, a Meta employee said: “I wouldn’t put too many eggs in the down-ranking basket.”³ See true and correct copy of MT-IG-AG-NM-000067446 at 7446 attached as Exhibit 42. The employee continued: “We (Facebook inc) now have years of evidence it’s simply not that effective. Even down-ranked, people still see the content. And that’s going to be more pronounced on IG because 25% of our DAP [Daily Active People, or the number of unique users on Instagram each day] exhaust their feed, way more than Facebook.” *Id.*

46. On June 25, 2020, Meta employees discussed an internal “RabbitHole project,” which aimed to “identify[] people who are seeking bad experiences and understanding if they can get into a rabbitHole of getting more and more bad content.” See true and correct copy of MT-IG-AG-NM-000158553 at 8553 attached as Exhibit 43. One employee said: “Molly Russell. That was a main finding from that investigation, people (especially in the SSI problem domain) seek that content. Creates a horrible feedback loop / downward spiral.” *Id.* The employee also acknowledged: “SSI has been probably the highest-profile / most studied problem where this happens.” *Id.* at 8554.

47. On August 28, 2020, the Senior Director of Engineering, told other employees in an internal message: “We promised . . . we’d do a bunch of work to improve IG’s impact, in particular in . . . what was the word, something like ‘vortex’ or ‘flywheel’ of recommendations of similar content. . . . So if someone likes an account posting depressive things, we don’t send that someone into a downward spiral of depressive content, *making suicide more likely*. . . . We also made some classifier promises, which the SSI team in Community Integrity is (slowly) working on.” See true and correct copy of MT-IG-AG-NM-000170303 attached as Exhibit 44 (emphasis added).

³ Down-ranking refers to demoting content so that appears farther down in a user’s feed.

48. On September 15, 2020, a Meta employee said on an email thread titled “[q]uestions on eating disorder detection / enforcement”: “when you search for [#proana and #thinspo] there are no results under the ‘hashtag’ tab, but there are endless results under the ‘Top Accounts’ and ‘Accounts’ tab, and almost all are violating.” *See* true and correct copy of METANMAG-002-00045730 at 5738-39 attached as Exhibit 45. The employee also said: “[w]e don’t currently have classifiers to catch eating disorder proactively – which means it’s easy to find on Instagram, and is showing up in the Explore and accounts ‘suggested for you.’” *Id.* at 5739. The employee then admitted: “[e]ven when we do launch ED classifiers (outside the EU) it will only look at the content level – not at the account level.” *Id.* The employee concluded: “[w]e’re concerned that our ability to find, remove and stop recommending these pro-ED accounts is going to be limited . . . and we’re going to keep getting negative press cycles about how we’re failing to fulfil the promises we’ve made.” *Id.* Another employee responded: “this is a perpetual problem we have with search across many issue areas as we are simply returning accounts that match the search term.” *Id.* at 5738.

49. In an internal presentation dated September 29, 2020, Meta admitted: “We don’t demote dark and depressive content, which can be harmful in aggregate. Importantly, this was the content that was consumed by Molly Russell, and the cause of the connection between her suicide and Instagram.” *See* true and correct copy of METANMAG-033-00099816 at 9854 attached as Exhibit 46.

50. On September 29, 2022, a research manager said: “honestly I can guarantee I have worse shit than whatever it says lol Feed Relevance consistently distributes SSI ViewPort Views to teenagers and nobody has noticed.” *See* true and correct copy of METANMAG-003-00774475 attached as Exhibit 47.

51. An internal document dated October 24, 2022, states: “Teens rabbit-hole into SSI & ED content, leading to compounding harms (mental health issues, suicide, anorexia, bulimia, self-harm)” See true and correct copy of METANMAG-002-02127833 at 7838 attached as Exhibit 48. A Senior Content Designer asked: “Why is this organized under ‘Age Appropriate Content’? Aren’t SSI/EDs an integrity issue that is relevant to all user cohorts even if teens disproportionately experience the problem?” *Id.*

52. On April 20, 2023, a Meta data scientist and senior researcher, discussed the effectiveness of SSI classifiers and the need for research. The researcher wrote: “I am personally extremely excited for you . . . to do the ecosystem analysis for SSI (violating and borderline⁴) so we can learn how teens are getting to that content, how they’re engaging with it, where it’s from, etc. etc. There’s so much that’s unknown.” See true and correct copy of METANMAG-005-01392558 at 2560 attached as Exhibit 49. The researcher also said: “there’s a LOT of hesitancy around quant [quantitative] research directly with users re: SSI . . . the issue is if we find there’s prob lots of content we’re not classifying properly, we then will be legally obligated to go try to fix it.” *Id.* at 2562. The data scientist responded: “I understand how that will make doing the research harder to get approvals.” *Id.*

53. On April 26, 2023, the same Meta data scientist told several employees, including the Director of Research: “The current [SSI & ED] classifiers are poorly performing [*sic*] (50% precision) and there is a known labelling issue in how they get the ground truth labels . . . In the current state my view is our borderline SSI & ED classifiers are unusable.” See true and correct copy of METANMAG-011-00139025 attached as Exhibit 50.

⁴ Borderline is used by Meta employees to refer to content that does not violate its Community Standards, which is referred to as violating, but is close in type.

54. On May 5, 2023, the same data scientist and researcher noted that Meta is not on track to ensure that “teens see less SSI content than adults.” *See* true and correct copy of METANMAG-005-01392401 at 2402 attached as Exhibit 51.

55. On May 23, 2023, several employees discussed a journalist’s inquiry “about the prevalence of SSI on social media” after the journalist identified several pieces of SSI on Instagram. The policy communications manager at the time said, regarding one of the SSI posts: “it’s an old post (519w [weeks] old), which presumably wasn’t caught partly because we don’t run our classifiers that far back, as discussed previously, but [I] wanted to flag as an example of where old posts are causing comms risk, as these old posts come up easily in searches. [A]lso, a commenter replied to the post 6 weeks ago, showing that old posts are still visible.” *See* true and correct copy of METANMAG-011-00302409 at 2409 attached as Exhibit 52. She also said: “if classifiers were running on old content I imagine we’d find and remove lots of content/accounts.” *Id.* at 2410. A product manager admitted: “this is prohibitively expensive to do for all old content . . . the volume of old content is likely too massive.” *Id.*

56. A document dated June 14, 2023, titled “IGTS SSI/ED 2023 H1 Retro – H2 Lookforward” acknowledged that interventions to address SSI/ED content on Meta’s platforms were blocked or paused “due to classifier delays” and “layoffs” of product managers. *See* true and correct copy of METANMAG-005-01710296 at 0298 attached as Exhibit 53.

57. A June 14, 2023, Youth Well-Being Research presentation reports that there are “[s]ignificant gaps, complexity, and inconsistency in usability, access and content of the [SSI] resources we offer to users” and that “[t]eens see more SSI content than adults.” *See* true and correct copy of METANMAG-007-00214892 at p.46 attached as Exhibit 54.

58. On August 1, 2023, a data scientists at Meta posted an internal “note” titled “SSI-ED consumption for 13-15 vs 16-17.” *See* true and correct copy of METANMAG-122-00001622 attached as Exhibit 55. He reported: “When compared to adults, teens have significantly higher . . . rates of high concentrations of exposure to bad content . . . and higher rates of repeat viewing SSI.” *Id.*

59. On September 28, 2023, Instagram’s Trust and Safety Data Science Tech Lead and a Machine Learning / AI Engineer discussed Meta’s “SSI/ED classifiers.” *See* true and correct copy of METANMAG-005-01602020 attached as Exhibit 56. The engineer wrote: “Honestly, they are not very good though – we are looking to do a better job with them. . . . The signals we have for SSI/ED are just not very good overall (mainly shows up in Search)[.] Its [*sic*] not catching things it probably should. Like an example we looked at was accounts called ‘_i_want_to_kill_myself_’ – things like that[.]” *Id.* The Tech Lead responded, referring to internal observations that accounts discovered in Explore were proliferating SSI/ED and not enforced. *Id.*

60. On December 10, 2023, a Meta data scientist revealed that “a teen audience member” of an SSI ED account “has 60x-80x the SSI-ED prevalence of the overall ecosystem.” *See* true and correct copy of METANMAG-122-00001604 attached as Exhibit 57. “[A] teen in the audience is 35% more likely to have a VPV on the SSI-ED content than an adult that is in the audience of that same account . . . So it’s not just a higher propensity to repeat view, it’s also a higher probability to be exposed at all.” *Id.*

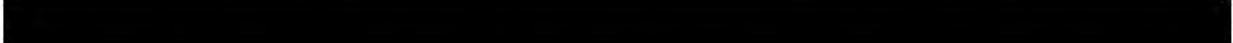
61. In 2024, Meta analyzed the impact of teens searching for SSI-related content. *See* true and correct copy of METANMAG-033-00060639 attached as Exhibit 58. The analysis found “SSI-related accounts were immediately and consistently recommended. IG began recommending problematic SSI-related accounts, many of which produce SSI content, across search pages (For

you, Accounts, Tags) immediately.” *Id.* at 0639. The analysis further found the SSI-related accounts “were recommended after only one day of searches, even when the search term was benign (e.g., searching for “self” returned self-harm-related accounts).” *Id.*

62. On July 2, 2024, in a message thread with multiple Meta employees, a software engineer, asked: “how confident do we feel about classifier refreshes this half? . . . I don’t think we’ve integrated classifier improvements into the h2 goal we’re setting yet.” See true and correct copy of METANMAG-028-00015754 at 5754 attached as Exhibit 59. The product lead for youth well-being, responded: “we have 1 SWE [software engineer] dedicated to SSI-ED in Q3 [third quarter] as the other eng[ineer] will be supporting Nido until then (and TBD on when that work will wrap). . . Biggest piece of context we need to land is that this team is not fully dedicated to SSI-ED (we’re splitting a small set of resources across” three additional areas, as well as “SSI-ED.” *Id.*

63. On August 16, 2024, a data scientist working on SSI/ED wrote to a colleague: “If only this company cared about and incentivized this type of issue (better engineering, analytical excellence) perhaps we wouldn’t have so much buggy shit and horribly confusing code.” See true and correct copy of METANMAG-108-00008654 at 8654 attached as Exhibit 60. The colleague, a data researcher, responded: ““Move fast and leave broken shit for someone else to fix.”” *Id.* The data scientist responded that Meta’s cross-function integrity and other staff folks “have very miscalibrated sniff tests. That and I’m amazed how often people just go forward with stuff without doing any sanity checks.” *Id.*

64. An October 2024, presentation titled “Understand Shareout IG Youth Appropriate Experiences” reported that Meta-promoted “[REDACTED]” [REDACTED].” See true and correct copy of METANMAG-055-

00025115 at 5119 attached as Exhibit 61. It also reported: “Research has shown that seeing even one piece of SSI content correlates with an increased likelihood that a user will be exposed to more SSI content. *Id.* at 5120. The presentation also reported that Meta was sending “notifications associated with SSI Content” to teens, and that over a two-week period more than 53,000 teens clicked on a notification from Meta associated with SSI content. *Id.* at 5122. It also reported: “
.”

Id. at 5130. “After searching for SSI-related terms, SSI accounts were recommended,” “SSI accounts were recommended as part of AYM [Accounts You May Follow], often alongside benign, verified accounts,” “SSI accounts were accessible via SSI accounts’ follower/following lists,” and “[a]fter only 1 day, Search adapted to surface SSI accounts for benign terms” meaning that even non-SSI searches were delivering SSI content. *Id.* at 5131.

65. Meta requires users to agree to Facebook’s Terms of Service and Instagram’s Terms of Use in order to use those platforms. *See* true and correct copy of https://mbasic.facebook.com/legal/terms/plain_text_terms/ attached as Exhibit 62 *see* true and correct copy of <https://help.instagram.com/termsfuse> attached as Exhibit 63

66. The Terms of Service for Facebook states: “How our services are funded . . . we use your personal data...to show you personalized ads.” *See* Exhibit 62.

67. The Terms of Use for Instagram states: “How Our Service Is funded . . . We use your personal data . . . to show you ads that are more relevant to you.” *See* Exhibit 63.

I declare under penalty of perjury that the foregoing is true and correct to the best of my knowledge and belief. Executed this 5th day of November 2025.

/s/Brendan Austin
Brendan Austin