

Senate Judiciary Committee Hearing  
“Hidden Harms: Examining Whistleblower Allegations  
that Meta Buried Child Safety Research”

**Questions for the Record for Cayce Savage**

QUESTIONS FROM SENATOR BLACKBURN

**1. Was Meta truthful in its responses to the Washington Post story about Meta’s child safety practices in its virtual reality (VR) devices?**

No, Meta was not truthful.

*“In a statement to The Post, Meta spokeswoman Dani Lever said that the allegation that Meta curtailed research is based on a few examples “stitched together to fit a predetermined and false narrative...””*

This is not a predetermined, nor a false narrative.

*“...and that the company has had no blanket prohibition on research about people under 13. Meta has produced research on youth safety in virtual reality and the company consulted children and their parents as it created the tween accounts, she said.”*

We are not claiming that Meta has a “blanket prohibition” on research about individuals under 13. The issue is not whether research was done at all, but how difficult Meta made it to do research with *anyone* under the age of 18, despite there being a grievous need for it. In particular, required legal reviews significantly lengthened research timelines, which is often unacceptable for researchers, given that product teams and leadership demand they usually conduct and report research as quickly as possible. This process became even more arduous for researchers needing to conduct research with people under 13. Despite my title being the lead - and only - youth-focused VR UXR, I had to request special permission to conduct research with people under 13 to support Project Salsa - the purpose of which was to lower the minimum age of VR headset use from 13 to 10, meaning research with 10-12 year olds was *necessary*. Despite this, and despite the fact that leadership was pushing the Project Salsa team to move as quickly as possible, it took several months for my legal counterpart to even identify a path by which I would be allowed to conduct research with people under 13, despite this being something we had been discussing since the beginning of my time on the team many months prior. legal never permitted me to conduct research with individuals under 13.

*“Lever added that Meta’s virtual reality devices have long had safety features including the ability to block problematic users, and over time the company has used research to develop additional protections for young people, including parental supervision measures and default settings that allow teens to communicate only with people they know.”*

We are specifically discussing the presence of safety features *for children*, which Meta has only had for its VR devices since mid-2022, eight years after its acquisition of Oculus and six years after it released its VR headset to the market, despite copious existing research which indicated that children would likely be heavily early adopters.

With regards to the parental supervision measures and default settings - While these are good things to have, our documentation and declarations show that these are insufficient so long as Meta fails to (1) educate parents about the risks their children face in VR and (2) improve the quality of their age data to ensure teen users are actually using teen accounts - only then will features like parental supervision and social default settings be triggered.

*““We stand by our research team’s excellent work and are dismayed by these mischaracterizations of the team’s efforts,” Lever said.”*

We are the research team in question and we are not mischaracterizing.

*“Meta did not directly dispute or confirm the events in Germany described by the researchers but said such a deletion would have been meant to ensure compliance with a U.S. federal law governing the handling of children’s personal data and with the General Data Protection Regulation, a landmark European privacy law that broadly prohibits companies from collecting personal information from anyone without consent. “Global privacy regulations make clear that if information from minors under 13 years of age is collected without verifiable parental or guardian consent, it has to be deleted,” Lever’s statement said.”*

GDPR and COPPA (As I assume Ms. Lever meant by “a U.S. federal law”) only apply when collecting data directly from a child under 13 (i.e. interviewing them directly) or when dealing with personally identifiable information (i.e. their address). Neither of these circumstances were present in this case.

With respect to “If information from minors under 13 years of age is collected without verifiable parental or guardian consent,” we were interviewing the child in question’s *mother*, who gave explicit consent both before the interview and during, by urging her elder son to continue speaking about the younger son’s usage.

*Meta told The Post that the headsets were meant only for people 13 and older and emphasized that the product packaging made that clear.*

Here is everything Meta’s VR box says about child usage (on the bottom of the box in small font): “Parents: Not all children are ready for Meta Quest. See child safety guidance at QR code or [meta.com/quest/parent-info](https://meta.com/quest/parent-info). Meta accounts for 10+ (varies by region).” and, in another paragraph, “Not recommended for younger or smaller-sized children or if headset cannot be adjusted to fit properly.”

Nowhere on Meta’s box does it explicitly state the *minimum age of use for the headset itself*. I’ll also note that Ms. Lever appears confused regarding Meta’s minimum age of use - It’s 10, not 13.

*“Meta said that lawyers have long partnered with research teams at the company and that there is nothing controversial about Meta lawyers advising researchers on attorney-client privilege or explaining that some work might require legal advice.”*

That would be true if that was what Meta’s lawyers were doing. As clearly presented with first-hand knowledge and documentation submitted to Congress, Meta’s lawyers were actively interfering with research, including preventing some work from being done and editing some work after the fact to obscure findings.

*“Changes instituted after Haugen’s disclosures were intended to make sure that research is high-quality and accurate, the company said.”*

In context, this reads like Meta believes its lawyers are qualified to assess the quality of and manage the research process. They are not.

This is also *not what Meta’s lawyers were doing*. If that were the case, they wouldn’t have been prohibiting certain topics from being researched, but rather providing guidance as to the best approach.

*“Meta told The Post that the company does not view app reviews as a reliable source of information about users’ ages and that it has better ways to determine how old users are.”*

We did not assert that app reviews were being used to assess the ages of users. I used app reviews to gather qualitative evidence that children were being harmed using Meta’s VR devices and to identify what kinds of harms those children were experiencing. These were a public source accessible to me after Meta’s legal team had prevented me from conducting original research into this topic.

With respect to “It has better ways to determine how old users are”...not really. That’s why I had proposed Project Horton to research age assurance. Until Meta released its age verification solution, which only applies after the fact to require users to provide evidence of their age *if* Meta suspects they are underaged, it had relied *entirely* on stated age, which research across Meta’s products shows is highly unreliable. At this current juncture, using app reviews as an age signal might be a good idea until Meta develops a more robust age assurance ecosystem.

*“The company disputed that a lawyer would be in a position to approve or a (sic) reject a study and said that lawyers never edit research results. Lawyers offer advice and suggestions, the company said, but research leaders ultimately determine which studies to pursue and the scope of their projects.”*

Unfortunately, Meta’s assertions are wholly untrue. I and many of my colleagues repeatedly experienced Meta Legal explicitly approving/rejecting studies, editing research results, and determining the scope of our studies.

*“Asked about the rush, Lever said the effort to create tween accounts came “on top of all the protections we had already in place” for teens.”*

“All the protections” already in place at Meta were under adopted, ineffective and insufficient - and Meta knows this.

*“Meta told The Post that it chose a 10-year-old cutoff so it could make sure it was offering age-appropriate content.”*

Before Meta leadership made the decision to lower the minimum age to 10, *no* research was conducted on the age appropriateness of VR content, only on the headset hardware, despite myself and other researchers flagging this as a serious safety gap. When I and other researchers raised concerns that Meta’s VR app ratings were insufficient to help parents make meaningful decisions about age appropriateness, Meta flatly told us that it did not want to invest in this area.

*“The company said it later implemented an initiative that asked headset users to confirm their birth dates. Those who said they were under 10 were removed.”*

This Meta assertion is misleading and the company knows it. Any of Meta’s so-called protections that rely on self-disclosure are insufficient. The issues contributing to underage usage are (1) a lack of parent education and (2) a lack of reliability in stated age data. Asking users to verify their age under these circumstances means the majority of under age users are still going undetected, as parents may provide their own information to “verify” the account. Moreover children have shown that they are more than capable of finding clever work arounds (using someone else’s ID, using AI images for the face scanner, etc).

*“Lever told The Post that leaders of the Reality Labs research team decided not to move forward with Project Horton because the company was already developing parental control tools and the initiative that asked headset users to confirm their ages by entering their birth date.”*

Project Horton was formally reviewed and approved *after the parental supervision tools in question had already been released to the public*. There was no new information about the parental *supervision* (not “control”) tools which would have affected anyone’s opinion regarding the necessity of this project.

Asking users to confirm their ages by entering their birth date is a low confidence method for gathering age data. It is very easy and common for users to simply give false information - Especially when users and their parents aren’t aware of the risks which make providing an accurate birthday important. Reliance on “stated age” (as this was called internally) is the reason I had proposed Project Horton in the first place.

*“An internal FAQ explaining the changes said the company’s “culture of openness” must now be balanced against “the risks that naturally stem from conducting and sharing research on sensitive topics and populations.” Meta told The Post that the changes were meant to ensure that research projects are accurate and findings are incorporated into the company’s product decisions.”*

Meta's statements here are false. Meta research leadership and my legal counterpart both told me that this shift in culture and process was to "protect" myself and other researchers in the event our work was to be "leaked" to the public. I was also explicitly *told* by Legal that this shift was to "mitigate risk" to Meta.

*"Meta told The Post that it requires users to verify their age with an ID or credit card if it suspects they are lying and that it created a tool to help third-party VR developers understand their users' ages."*

Yes, *if it suspects they are lying*. This means the user will have already been using the product for some time, and that this will only happen for *some* of the underaged users on Meta's platform. A company prioritizing user safety should aim to identify underaged users *proactively* and *en masse*.

*"In his affidavit submitted to Congress, the researcher wrote that he thought he was being told to avoid gathering data that could "implicate the company in future engagements with regulators." Meta said that characterization was inaccurate and that the lawyer was trying to help the researcher achieve the approved objective of his study, which did not involve collecting information about users under 13."*

To be clear, a user above the age of 13 mentioning that someone - not in the room - under 13 is using a VR headset is appropriate and standard research practice, and is neither in violation of COPPA nor GDPR. If Meta were leveraging research in good faith, it would *want* to collect these insights, as they would indicate that there was an issue of underage usage which should be addressed.

**a. A Meta spokesperson stated that Meta added more protections for young people on its VR devices. Are you aware of any protections that Meta added?**

Meta's continued pattern of vaguely referring to "protections" and the like indicates that they are not approaching this conversation in good faith. Any features Meta has already added to its product are now public facing, meaning there is no reason for Meta to avoid citing precisely *which* "protections" it has added. If Meta's goal is truly to deliver safer experiences, surely it would wish to educate the public how to use its products more safely at every opportunity.

Meta has released additional features since my time on the team. We are not asserting that Meta has not released *any* "protections," but that the "protections" it has released are retroactive and insufficient - and Meta is aware of and disingenuous about the efficacy of these "protections." For example, the website behind the QR code on the VR headset box delivers concise and actionable supervisory and safety guidance for parents, *but* this information is needed in more discoverable places (for example, in a flyer *in the box itself*), as it is well established that putting information behind a QR code decreases the likelihood of people accessing it. Meta is aware of this through my own assertions, the assertions of my colleagues, and because this is general best practice within the UX field.

The core safety risks to young users which our disclosure has raised still have not yet been meaningfully addressed. If Meta's defense is that it has added new protections, why have those protections not meaningfully addressed the largest risks to the safety of young users? This too reflects Meta's bad faith approach to this conversation. At the time I left the company children were rampant in VR, and I regularly observed and heard accounts of children being propositioned for sex acts, groomed, harrassed, getting doxxed and more. The vague mentions of "protections" could refer to something as simple as moving the parental education page in front of the Meta account log in requirement (as they have recently done), a change which is certainly a good one, but which does not meaningfully address the fundamental issue that *user awareness remains low about the minimum age of use of their headsets and the risks of harm on platform*. For example, Meta has still not chosen to clearly and concisely state the minimum age of headset use on the box itself. Their own spokesperson did not even seem to know the minimum age of headset use when responding to this whistleblower disclosure.

Finally, Meta has a history of adding features and using their presence as proof that Meta is behaving responsibly, when these features do not necessarily make the experience safer. For example, Meta bragged about its parental supervision features for teen, then tween VR users, when existing research of which Meta was aware clearly showed that parental supervision features alone are not sufficient because they often go underutilized, and need to be paired with meaningful conversation (and co-usage for younger users). Meta's own internal data reflected this, showing low adoption rate of their VR parental supervision features.

Meta's spokespeople perhaps forget that some whistleblowers that are party to this disclosure are *still at the company*, and so the information we have about what "protections" Meta has added is current. I would like to see a list of the "protections" Meta's spokesperson speaks of, as well as a detailed explanation of why the company believes these efforts are effective, given the grievous nature and frequency of harm to which they are responding.

- b. Meta disputed your assertion that lawyers were able to approve or reject research projects, and they said that lawyers never edit results.**
  - i. Is this statement consistent with your experience at Meta?**
  - ii. Would a company lawyer being involved in child safety research interfere with such research?**

[b.i.] This statement is wholly inconsistent with my personal experience at Meta, as well as the experience of many of my colleagues, with whom I regularly discussed this issue. During my time in Meta's Reality Labs, legal interference was so frequent and disruptive that my UXR team spent numerous meetings discussing this issue. I was also told by researchers supporting a variety of Meta's products, including Instagram, Horizon, and Reality Labs more broadly (mixed reality, wearables, etc) that Legal interfered with their research, and witnessed this occurring on more than one occasion.

I will also note that Legal's leadership was well aware that Legal was acting in this capacity, as I experienced Legal directors engaging in this behavior themselves. UXR leadership was also aware that Legal was behaving in this capacity and told our research team on numerous occasions to defer to Legal when making decisions about research.

Legal interference I (and my colleagues) experienced included:

- Rejecting proposed areas of inquiry, stating that they were too “sensitive” or “risky” (This included investigating what kinds of harms children were experiencing in VR.)
- Editing specific research questions (meaning, within a specific area of inquiry, the more specific things we could investigate)
- Editing discussion guides and survey questions
- Editing research reports to remove or abstract information
- Controlling with whom research reports could be shared

To make this very clear, *Legal was given the power to behave in these ways by both legal and research leadership*, and there are *no other bodies that engage with research in this way, nor have this power at Meta*.

[b.ii.] Yes, “youth” in VR was my explicit area of study, and I was given a legal counterpart to review *everything* I wrote down which I intended to share with anyone other than my manager, and with whom I was instructed by my manager to discuss any plans for future research. Legal’s interference prevented me from investigating (1) the prevalence and causes of underage users in VR, (2) the harms these users experienced and potentially caused, (3) what could be done to address points 1 and 2.

**c. Andy Stone, Meta Spokesperson, stated that Meta approved almost 180 studies on safety and well-being on its virtual reality platforms since 2022. How would you respond to this assertion?**

This is an absurd statement which belies, at best, a fundamental lack of research literacy. That Meta would allow someone displaying such a lack of knowledge to speak on their behalf is itself concerning.

Firstly, the number of studies *approved* is not the same as the number of studies *conducted*. As we saw from the age assurance project (Project Horton) example, a study could be approved and still not be conducted.

Secondly, we have not claimed Meta’s Reality Labs has done no safety and well-being research. That was my colleague Dr. Sattizahn’s specific area of study for which he conducted numerous research studies. My specific concern is about the lack of due diligence with respect to the safety and well-being of *children* using Meta’s VR headsets, given their heavy presence in Meta’s virtual spaces. Children have meaningfully different safety and well-being needs from those of adults with respect to the usage of digital media, so dedicated research is necessary.

Thirdly, even the number of studies conducted is not a sensible metric by which to gauge Meta’s dedication to user safety. It matters what the studies’ areas of inquiry area are, with whom the research is conducted and at what scale, the truthfulness and directness with which researchers are able to report findings, whether other staff are made aware of the research findings, whether the research findings are actioned upon, and whether the actions taken are meaningful. As our evidence, disclosures, and testimony show, Meta Legal and leadership actively undermine these

steps. Given the grievous nature of the concerns under discussion, it would benefit the public if Mr. Stone could stay on topic in future communications.

## **2. What is the relationship between Meta's Reality Labs and Instagram?**

Meta shares a variety of resources across its wide diversity of products, including Reality Labs and Instagram. The bulk of Meta's resources are dedicated to whatever area it considers to be its strategic area of focus at the time. Prior to 2021, Instagram was Meta's area of focus, such that the majority of its funding, talent, and headcount resources were concentrated there. When I joined Reality Labs in January of 2022, it was widely acknowledged internally that Meta's focus had shifted from Instagram to Reality Labs, such that the funding, talent, and headcount were being increasingly concentrated there instead. The implications of this were that people were transferring internally from Instagram to Reality Labs at an increasing rate, bringing with them established processes, knowledge, etc.

Internal transfers from Instagram to Reality Labs were very common, and there was no contextual training to make this transition. This actively caused issues, as VR is a general computing hardware device often used for gaming, and Instagram is a social media app. Folks coming over from Instagram often struggled to make the conceptual shift; I had to repeatedly explain to fellow Reality Labs employees, including leadership in C-suite, that VR was different from a two dimensional software experience and that this had strong implications for user safety and well-being (i.e. The immersive and embodied nature of VR means that harm has greater impact than equivalent experiences on any two dimensional app, and so establishing robust protections is even more critical in VR.).

Individuals coming over from Instagram to Reality Labs brought with them established processes, such as the rapid six month development cycle used for app development. This rapidity contributed to and, in some cases caused, user safety issues. For example, leadership demanded the VR Parental Supervision product team develop net-new parental tools for VR in less than a year, when this is something which needed significant research and careful thought.

Many employees work across products (i.e. Product Compliance and Policy), including VR/Reality Labs and Instagram, and make decisions which affect both of these areas. Some of these teams are even user-facing - For example, the Avatars team creates avatars which are used across both Instagram and VR, as well as a number of Meta's other products.

Some employees dedicated to Reality Labs' products also supported Instagram. For example, during my time in Reality Labs, I shared my VR research with Instagram teams to guide their decision making, and vice versa. I also partnered with Instagram researchers to conduct net-new research relevant to both of our products. Such knowledge sharing and collaboration is common and encouraged across Meta's products - Or rather, it was when I worked there.

Most of Meta's products are connected on the backend. For example, when I worked with the VR Integrity team to identify those users behaving maliciously towards children, the team noted that they were significantly hindered in their ability to do this. The problem was, the information VR users gave when setting up their VR headsets and accounts didn't correspond to information



they used when setting up their other Meta accounts (Facebook, Instagram, etc), so the team's access to existing information about the person (such as age) was more limited than usual. This is an example of just how normal it is for Meta's teams to collaborate and share information across products, including about its users.

Finally, Meta is actively pursuing avenues by which to create shared experiences across VR and Instagram in order to boost engagement. For example, Meta has been showing short video clips of VR on Instagram, and has been pursuing the creation of an "Instagram feed" within the VR homescreen. Not only are these products connected in a myriad ways on the backend, but Meta is actively seeking ways to connect them for the user as well.

### **3. What implications are there for user safety when Instagram or other social media content is integrated into VR?**

Firstly, this is a privacy issue, given the high prevalence of VR headset and account sharing. Only one user's Instagram is going to be shown within the VR homescreen, and introducing an Instagram feed into VR means anyone in the household using the headset can easily access this person's private account.

Secondly, as I mention in my documentation and declaration, VR introduces new and more severe kinds of risks. Reading a bullying message in Instagram in real life may feel painful, but reading it in VR may feel even more severe - Research into the effects of this kind of experience don't yet exist within or outside of Meta.

Thirdly, overuse is a serious concern for VR, as the headset covers your eyes, making it difficult to perceive time passing in the real world. We already know Instagram can be addictive for both adults and children, but introducing Instagram (or any social media) into VR could potentially exacerbate this effect. Again, the research on this risk has not been done.

### **4. Meta has said publicly that its VR headsets are intended for children aged 10 and older. But based on the internal documents you shared, as many as 80 to 90 percent of users in some virtual rooms were under ten years old.**

#### **a. How are underage users able to access Meta VR?**

Underage users are able to access Meta VR in two ways:

- Sharing a headset/account with an older user
- Using inaccurately aged accounts (i.e. their parent's account or an account the child intentionally made to have an older age)

Headset sharing is common, as VR headsets are expensive and most households do not own a headset for every household member. Account sharing is common, especially for younger users, for a variety of reasons, including convenience and the perception that this enables supervision.

Research indicated that parents are most involved in their child's usage of Meta's VR headsets during account creation, and that parents may create an account on behalf of their child, or instruct their child to create an account using the parent's information, to protect the child (by not

giving Meta the child's information) and to facilitate supervision. It is also very well established in research across digital media/experiences that children, especially tweens and teens, seek increasingly mature experiences, and so are likely to create adult-aged accounts for a variety of reasons (i.e. to have access to M- rated games).

**b. Why is this problem still ongoing?**

Firstly, because Meta has not sufficiently invested in parental education. Parents are largely unaware of Meta's stated minimum age of use for their VR headsets (it's not even on the box) and are largely unaware of the risks their children face when using VR. For these reasons, many parents do not understand why adhering to the minimum age of use or using accurately aged accounts for tweens or teens matters, and so they continue to allow underaged usage and inaccurately aged accounts. I want to emphasize that this education is not difficult to do - Simply doing things like putting a flyer in the VR box and putting the minimum stated age on a discoverable part of the box itself would go a long way.

Secondly, because Meta released this technology without setting the precedent for appropriately aged usage in the first place. Even the best approach to education about minimum age and risk is going to be less impactful now that the headset is in so many households and underaged usage is normalized.

Thirdly, because Meta refuses to invest in meaningful age assurance. As I discuss in my declaration, age assurance is comprised of stated age, predicted age, and age verification. To meaningfully improve their age assurance ecosystem, Meta needs to:

- Educate users and parents about why providing an accurate stated age matters (The reason is the risk children face)
- Investigate age prediction technologies, such as pupil distance, height determination, etc. This will require Meta to work closely with industry experts and parents to ensure this approach feels appropriate (This was the work proposed in my "age assurance" study, called Project Horton, which was shut down by Meta.), given that VR is a newer form of technology which enables novel forms of age prediction. It is of paramount importance that any new age prediction approaches feel appropriate, given concerns about user privacy.

Meta continues to assert that they now require age verification if they suspect an account is being used by a minor, but Meta knows that, without these other two steps, verification can be easily foiled and is unlikely to catch the majority of underaged users or users who should be using a tween or teen account.

These are the tactical reasons underaged usage continues, but all of these are things Meta could readily address - If it wanted to. The reason the issue persists is because Meta does not want to deal with the business downside. Meta is required to remove accounts when it becomes aware they are being used by underaged individuals. Doing so would decrease their usage metrics, which Meta relies on to demonstrate shareholder value. Meta chooses not to meaningfully address underaged usage in favor of maintaining inflated metrics for their shareholders.

## 5. What can Meta do to make VR safer? Why isn't Meta implementing these measures?

### What Meta can do to make VR safer:

1. Meta should listen to its users and prioritize their safety.
2. Meta should begin investigating safety risks and mitigations *proactively*. Today, Meta tends not to conduct good-faith research on the safety of its products - including VR - until after it's already in a rush to get the product out to market. For example, Meta could have begun developing parental supervision features for VR when it first acquired Oculus, but instead waited eight years until it was already the subject of regulatory oversight to invest in this much needed feature set. Having multiple years to research these features would have ensured that Meta could produce an industry-leading experience, but they chose to develop something in a rushed manner and release a sub-par experience.
3. Meta should establish a practice of proactively referencing existing research and speaking with relevant experts - Many of whom they already employ. I witnessed Meta regularly ignore its *own employees* for whom they paid for their expertise, when it was perceived that engaging with the concerns the employee raised would generate "risk" or slow the team down. I also witnessed Meta ignore experts outside of the company. For example, Meta's VR policy team regularly engaged with industry experts to get their thoughts on Meta's approach to various aspects of its product. The concerns of these experts were regularly ignored.
4. Meta should enable its researchers to do their job (without interference from legal) in order to identify the risks users face and the paths towards addressing those risks.
5. Meta should *meaningfully* invest in accurately identifying the ages of its VR users. Any other safety features related to a user's age are meaningless without this.
6. Meta should *meaningfully* invest in parental education, especially regarding the minimum age of use, the risks that children face, and how best to keep young users safe (i.e. time limits, co-usage, etc).

### Why Meta is not implementing these measures:

The things I've listed above would arguably be *more efficient* than any frantic, after-the-fact scrambling on Meta's part. Having a safer experience is better for the health of both the product and the company, as it increases usage, time used, company trust, sentiment, etc. Even Walt Disney insisted that investing in user safety as a pillar of development is just good business sense. The only conclusion I'm able to come to here is simply that Meta's desire to "move fast" and get products out in front of the public as quickly as possible is more of a priority than user safety. I might say that perhaps Meta doesn't understand the business value of trust and safety, but they employ *hundreds* of researchers who have repeatedly told them otherwise. Unfortunately, prioritizing speed over safety is a conscious choice.

## 6. Did the Meta legal team understand the safety implications when it instructed you to change studies and findings from your research?

Yes. In some cases, I explained the safety implications to them. In other cases, their instructions were *because* they understood the safety implications. For example, instructing me to remove mentions of users under the age of 13 or 10 meant they had to first be aware that underaged users

were being organically mentioned in research. If Meta responds here by saying their legal team didn't have a tidy list of safety risks, I will remind them that Legal's interference was the reason *they did not have that list*.

**7. To your knowledge, has Meta ever brought in child health and safety experts to advise on virtual reality products?**

Yes. Firstly, I was aware that Meta's VR policy team hosted a regular meeting (I believe it was monthly) with industry experts regarding VR, some of whom had relevant expertise in areas regarding child health and safety. Secondly, Meta *employed* individuals with this expertise, including former tenured professors of developmental psychology.

**a. Did Meta's findings or recommendations align with what you were seeing?**

No. I consistently saw Meta ignore the recommendations of the experts it consulted. Some of my colleagues who were child health and safety experts even expressed to me that they felt threatened by Legal for expressing concerns about the way Meta was approaching its VR product.

**b. Did Meta implement or act on any recommendations from these experts?**

I am not aware of Meta acting on any recommendations from external child health and safety experts. Meta has acted on the recommendations from some of its internal experts, but only when those recommendations aligned with Meta leadership's existing goal of quick growth. Otherwise, I know internal experts have had to fight hard to get the company to act on even very basic recommendations regarding child safety - for example, *putting the minimum age of headset use on the outside of the box*.

**8. What was Meta Leadership's expressed intention in lowering the minimum age of virtual reality users?**

Minimum age was lowered both for Meta's VR headset and for its VR social app Horizon.

With regards to the headset, I was shown a slide deck when "read-in" to this highly confidential project which included leadership rationale for lowering the minimum age. The rationale in that slide deck stated that Meta was seeking to lower the minimum age as a response to increased FTC pressure, and that they believed doing so would increase engagement.

With regards to the Horizon app, I was explicitly told by Tim Loving that leadership believed it would increase adoption and engagement.

**9. On January 31st, 2024, Mark Zuckerberg testified that Meta does not instruct teams to focus on engagement.**

**a. Is this statement true?**

No. This statement is false, and Mr. Zuckerberg knows that it is false. Throughout my four years

at Meta, I was embedded in Facebook Marketplace, Facebook Jobs, Buy and Sell Groups, and VR. For every one of these products, our *primary* directive was to increase engagement, many times at the expense of good user experience. Every time, I was told these directives came from Mark Zuckerberg himself.

**b. If not, how were these directives shared at Meta?**

These directives were shared by leadership in documents, in Workplace “posts,” and in meetings both big and small. Our success metrics usually reflected these directives. For VR specifically, Tim Loving (UXR director) explicitly told our team that Mark Zuckerberg wanted “every team” to focus on engagement.

**c. Were they shared company-wide?**

It depends on the scope of the directive - Sometimes directives are only relevant for one product (i.e. Just Reality Labs but not Instagram, Facebook, etc). However, (1) the directive from Mark Zuckerberg to focus on engagement was company-wide and (2) even before this specific instance, evidence that the standing directive for *the company as a whole* was to focus on engagement was pervasive.

QUESTIONS FROM SENATOR COONS

**1. I co-lead the *Platform Accountability and Transparency Act (PATA)* which is designed to create mechanisms for independent research of social media platforms, their harms, and the effects they are having on users or society at large. In light of what you observed at Meta regarding their approach to internal research, could you elaborate on the value and importance of having effective independent ways to research platform behavior?**

Meta has clearly demonstrated that they are *unwilling* to conduct good faith research on the ways their products affect users, particularly when they deem the area of inquiry to be “risky” to them. In example, Meta responded to Frances Haugen’s 2021 whistleblower disclosure by making it *more* difficult to investigate the effects of their products on minors, rather than facilitating this research openly and in good faith.

Given how pervasive social media is, and especially how much time people on average spend on social media, meaningfully understanding the effects these products have on us is critical to identifying risks to our individual health, well-being, and so on, as well as the risks to our health and well-being as societies, for social media impacts us both psychologically and sociologically. Meta’s consistent behavior has shown they are not willing to take responsibility for conducting and communicating research to this end on their own. Independent research would be a powerful tool to hold Meta accountable for the ways their products affect the public.

It is also critical to understand the *cumulative effect* social media technologies have on us (i.e. the use of social media across Instagram, TikTok, etc). This research must be conducted

independently to minimize bias - I would argue that it is impossible for for-profit companies to conduct unbiased research on their own products in a competitive context.

**2. During your time at Meta and regarding the businesses you worked on or otherwise had insight into, to what degree did Meta facilitate independent research regarding the harms caused by their products? Was such research permitted or possible?**

No, I am not aware of Meta facilitating independent research regarding the harms caused by their products. I *am* aware of independent research being done on the harm caused by Meta's products, but none of this research was done in association with Meta, and I was at times discouraged from sharing such independent research within Meta.

Meta does have programs which facilitate research through third parties, though none of this work is what I would describe as independent, as Meta enacts control over the process, including:

- The kinds of research questions which are prioritized
- The methodology
- How research participants are sampled
- The way questions are asked
- The way findings are represented
- What research findings are represented
- How research findings are prioritized

To be very clear - If there were good-faith independent research on harms caused by Meta's products, Meta's own employees would have a clearer understanding of the harms caused by Meta's products. And yet, there are rarely reports one can point to internally to confidently and concisely say "These are the ways in which our users are being harmed by our products." Instead, employees must often create a patchwork understanding across numerous documents, rely on word of mouth, or share the information with each other in secret. For example, when investigating possible harms children using VR were facing, I spoke with a Facebook researcher who shared a robust, exhaustive document citing research on bullying children experience on Facebook. The researcher cautioned me not to share the document with anyone, and not to tell anyone they had shared it with me.

**3. Do you think there is more that Meta could be doing now to facilitate independent research into these products? What might that look like if Meta wanted to cooperate?**

Before discussing Meta's facilitation of independent research, I think it is necessary to acknowledge Meta's current relationship to its own internal research. If Meta is actively interfering with good faith research being conducted internally - going so far as to remove findings from reports - I do not believe we can reasonably expect Meta to facilitate independent research in any manner with good faith.

There are ways in which research being conducted independently can protect the fidelity of the findings, but there are other ways in which Meta can actually interfere with the research in even

more significant ways. For example, research being conducted independently means, in theory, that Meta cannot control what research questions are prioritized, but this also means the independent research body will be comparatively limited because they won't have access to Meta's internal data. For example, if the independent research body wanted to investigate bullying on Instagram, they wouldn't have access to Meta's internal data about things like usage of Meta's safety tools (i.e. blocking, reporting, etc) unless, of course, *Meta gave it to them*. Given Meta's interference with its *internal* research, I can't imagine that we could expect them to engage in good faith with requests to reference user data in such ways.

In terms of what cooperation would look like:

- The independent research body would need to be fully in control of the area of inquiry, the specific research questions invested, the methodologies used, the sample used, the research conduct, data analysis, and reporting.
- Meta would need to work with an outside party to determine meaningful, appropriate ways to share existing research and user data with independent research bodies as needed. Data Meta provides would need to include the level of fidelity necessary for the research project, meaning it would not be appropriate for Meta to insist on providing only anonymized or aggregate data for certain studies. Data sharing of this kind is possible to conduct in ways which are respectful of user privacy and is common practice in other areas of research.

**4. Based on your experiences, how would you expect Meta to respond to (or argue against) requests to facilitate such independent research, and how would you respond to their contentions?**

Firstly, I expect Meta to say that they already work with independent research bodies, and to provide a vague quantification of this (i.e. "We work with over 50 independent bodies") or, at best, to cite a few bodies with no context (i.e. "We work with this NGO."). To this, I would say, whatever they're doing now clearly isn't working, given the known harm Meta products cause, and a conversation about user safety isn't the time to argue semantics.

There are *many* independent research and advocacy groups, as well as independent subject matter experts, who have clearly articulated what Meta needs to do differently regarding the safety of its products. If Meta were willing to listen to independent bodies, those changes would already have been made. For example, if Meta had spoken to Fairplay even briefly about parental supervision for VR, they would have prioritized robust, discoverable parental education. And yet, the parental education Meta chooses to provide remains buried behind a tiny QR code on the bottom of the box.

Secondly, I expect Meta will say that they *can't* share user data with independent research bodies because it would violate users' concerns about privacy. To this I would say, first, Meta's users are already deeply concerned about their privacy in Meta's hands. This is one of the main reasons parents don't give Meta their child's accurate age today. Secondly, I would say that Meta engages in this very behavior *constantly* when working with their third party vendors. Throughout the contractual engagement, the third party vendor may have access to relevant Meta log data, to

users' full names, phone numbers, emails, even addresses if the research is in person. Engaging with independent bodies would be no different.

Thirdly, it is possible that Meta will argue against this by making a defensive claim of its internal researchers. For example, Ms. Lever's response to our assertions in the Washington Post, "We stand by our research team's excellent work and are dismayed by these mischaracterizations of the team's efforts." To this, I would say the six whistleblowers in this disclosure *were* and *are* Meta's research team, and we have come forward precisely because we are concerned about the ways Meta weaponizes and suppresses its own internal research. The reason I am advocating for independent research is not because of any concerns with Meta's internal researchers, but rather it is because *Meta* has demonstrated that it will not let their own researchers conduct investigations in good faith.

**5. What would you see as the greatest challenges to creating a process by which vetted independent researchers could study platform behavior in the businesses you have insight into? How could those challenges be addressed?**

For platforms with such wide adoption as Instagram or Facebook, conducting research on their users would often be quite straight forward. Many independent bodies already do it. The biggest hindrance in my opinion will be the lack of access to platform/log data, for a few reasons. Firstly, users' perception or memory of their experience or behavior sometimes differ meaningfully from the reality. For example, a user may claim that they blocked another user, when in fact they did not successfully do so. These discrepancies are critical to understand, particularly when considering user safety and well-being. Secondly, having access to "log data" or internal data about how users engage with the product can be critical to identify trends without needing to survey all users (which is not possible). For example, seeing the log data which indicates that adoption of parental supervision features is very low. This information might be necessary for an independent research body to identify that this is a problem area requiring deeper inquiry, as, clearly, Meta cannot be fully trusted to identify the needed areas of inquiry. This particular challenge can be addressed, quite simply, through good-faith cooperation with the independent research body, provided care is taken to protect users' privacy in the handling and sharing of log data.

**6. What would you see as most necessary to get right to have as effective a mechanism for independent research as possible?**

It is critical that the independent research body be

- Protected as much as possible from Meta's interference, intimidation, and disparagement
- Attractive to researchers as a place to work in order to attract some of those experts who are currently under Meta's employ. This will require funding and tooling to enable high quality, expeditious research.
- Able to "move fast" to keep up with the speed of Meta's decision making, especially given this body would likely find out about decisions retroactively
- Able to meaningfully disseminate this research to appropriate bodies (i.e the public, regulatory bodies, etc)



It will be critical for any independent research body to have consistent, good faith access to Meta's usage and user data as appropriate in a manner which is respectful to users, and that the data Meta makes available to them accurately reflects users' real experiences - Not just the picture Meta wants to paint.

I, and the other whistleblowers taking part in this disclosure, are excited by the prospect of such an independent research body and are eager to support it.

**7. Apart from independent research, what transparency (e.g., data, reports, etc.) do you think Meta could and should be providing regarding the businesses you have insight into so that the public, researchers, and policymakers can better understand the effects of its products?**

Generally, Meta must stop censoring its own researchers, stop compromising the integrity of internal research findings, and must cease its bad faith usage of numbers/statistics to demonstrate responsible action. For example, Andy Stone's recent claim that Meta had "approved almost 180 studies on safety and well-being on its virtual reality platforms since 2022" belies either a fundamental lack of understanding of or care for research. The number of studies approved is not the same as the number of studies conducted (as we saw from the cancelled age assurance study), and the number of studies conducted is not a sensible metric by which to gauge Meta's dedication to user safety. This is a wholly meaningless response which underscores Meta's lack of willingness to discuss user safety in good faith.

In terms of more specific practices, Meta must shift to having a more open relationship with the public. Their products permeate the lives of individuals across the world, and I believe the public would greet open dialogue eagerly and actively.

- Sharing research: Today, Meta shares selective reports which are positioned to advocate for the company (i.e. Meta's Trust, Transparency & Control Lab). Given Meta's history, I believe a more robust, regular cadence of public disclosure of research would be appropriate, provided proprietary information is not shared, user consent is given, and PII is removed. This should be inclusive of both past and future research.
- Research for the public: I would also advocate for Meta creating a cadence of "research for the public." These would be studies done on topics the public cares deeply about (i.e. Instagram bullying) in which Meta makes public all relevant aspects of the study, including research questions, methods, sampling approach, timeline, discussion guides or survey questions, analysis approach, and final report. Meta should provide explanations for the decisions made about the studies (i.e. "Here's why we used this methodology"). The goal of these studies would be transparency and public education. I should note that Meta does currently make some research publicly available (i.e. their Trust, Transparency and Control Lab), but this research is cherry-picked to demonstrate Meta's good deeds.
- Public tracking of key metrics: Given Meta's history, I believe making public certain metrics related to user well-being and safety is needed. For example, making public the rate of adoption of its VR parental supervision tools overall, as well as the individual features.

- Enabling third party audits of the product development process and the products themselves: Meta must enable third party insight into *how* it develops products, as well as the outcomes of the products it creates.
- 8. Based on your experiences, how would you expect Meta to respond to (or argue against) requests for such transparency, and how would you respond to their contentions?**

Meta consistently employs a playbook in response to whistleblower disclosures such as ours which I expect to be repeated here. Broadly, this will look like Meta critiquing the validity of the ask, rather than engaging with it directly:

- Meta will likely claim that it is already being transparent, providing vague examples such as Mr. Stone’s “almost 180 studies approved.”
- Meta will attack the credibility of the requester, for example claiming the individuals making this request have an agenda or just don’t understand how much Meta allegedly already does.
- Meta will attack the credibility of any evidence provided, claiming it’s being taken out of context, isn’t sufficient, etc without specificity.
- Meta will play the victim. This may look like Ms. Lever’s statement to the Washington Post that our lawful disclosure was a “mischaracterization” of their research team’s excellent work, thereby positioning the critique as one not of the company but of specific employees who need protection. Or it may look like Meta’s internal communications that Frances Haugen’s 2021 disclosure made it difficult for researchers to do their job, thereby positioning the critique as one which unfairly hurts a company just trying to do the right thing.
- But, most importantly, Meta will not acknowledge the validity of the critique, no matter what.

My response to any of Meta’s contentions is this:

To respond to requests for greater transparency with anything other than meaningful engagement (such as, “Absolutely, how can we help?”) is to clarify that the safety and well-being of your users is not your priority. We are no longer having a conversation about the validity of the critique, because the body of supporting evidence and whistleblower accounts of Meta’s bad acts is too comprehensive, consistent and compelling to be ignored. At some point, highly consistent data becomes significant. Whether greater transparency is needed is no longer up for debate, so we ask Meta to stay on topic. Don’t try to distract us with a series of empty percentages or make this conversation about protecting your employees. We know how hard they work. Some of the whistleblowers that have come forward currently work for you. Help the public understand how your products affect them, so we can work *together* to make them better. Your users are telling you what they want. It’s time to listen.