

A D D E N D U M
to
HIDDEN HARMS: EXAMINING
WHISTLEBLOWER ALLEGATIONS THAT META
BURIED CHILD SAFETY RESEARCH

This Addendum is available at:

<https://www.govinfo.gov/content/pkg/CHRG-119shrg62327/pdf/CHRG-119shrg62327-add1.pdf>

Submitted by Ranking Member Klobuchar:

Horizon-Worlds-Exposed, statement	2
Further Declaration of Jason Sattizahn, statement	22

HORIZON WORLDS EXPOSED

**Bullying, Sexual Harassment of
Minors and Harmful Content are
Rife in Meta's Flagship VR Product**



Meta fails to protect minors from explicit content and harassment in its flagship VR app

Minors are being routinely harassed and exposed to adult content on Meta's flagship virtual reality social network, Horizon Worlds, new research has found.

The research follows Meta's announcement that Horizon Worlds will be opening up to 13–17 year olds as soon as March, showing that Meta is already failing to prevent minors from accessing mature content despite a supposed ban on minors accessing its VR apps.

Researchers logged into Horizon Worlds 100 times, recording footage from each visit. Horizon Worlds is the flagship social app in Meta's family of 'metaverse' VR products, and is currently supposed to be limited to users aged 18 and up.

Analysis of recording showed that minors already routinely use the platform, with minors present in 66 of the 100 recordings collected. Within these recordings researchers identified 19 incidents of abuse directed at minors by adults, including sexually explicit insults and racial, misogynistic and homophobic harassment.

Minors were also spotted in multiple 'Mature Worlds' where Meta permits sexually explicit content, legal drugs and gambling. Mature Worlds must be marked as 18+ but there are no further safety measures and they are easily accessible from the main menu or in-world 'portals'.

Sexually explicit insults were not uncommon, with researchers encountering four incidents of adults harassing minors in this way. One adult repeatedly shouted at a group of young girls "I don't want to cum on you", continuing even after the girls said they were minors.

Minors also faced harassment based on race, sexuality and gender with six incidents recorded. In a virtual courtroom a young girl with a black avatar was told "you're black, you're sentenced to death" by an adult player.

Researchers entered Mature Worlds 19 times and found minors present on five occasions. One minor was found inside a virtual strip club named *The Rabbits Den* serving alcoholic drinks for in-app money. In a separate incident, a portal to *Chicago Penthouse*, a strip club, was opened in front of a minor, creating direct access to the Mature World.

Researchers collected this data by entering Horizon Worlds and filming 100 visits to the most popular worlds. Worlds were selected from the *Top 100* menu which ranks worlds by

popularity. Researchers explored the worlds fully, spending a maximum of five minutes inside.

Recordings were then analyzed to identify underaged users and incidents of abuse. Horizon Worlds does not provide information on the declared ages of users, and so users were considered to be minors if they stated that they were under 18 or if two researchers agreed that the voice of the user sounded under 15 years old.

Summary

- Meta will soon be opening up their flagship VR social network Horizon Worlds to teens aged 13 to 17, but new research has found that minors are already being failed by the VR social network.
- Despite Horizon Worlds currently being limited to users aged 18+, researchers found that young teens are joining the network where they have easy access to virtual sex clubs and drug boutiques.
- Researchers made 100 recordings of visits to the most popular social spaces in Horizon Worlds and found minors were present in 66 out of 100 worlds.
- Analysis of the recordings identified 19 incidents where minors were harassed by adult users, including sexually explicit harassment, racist abuse and misogyny.
- Researchers also found minors in 1 in 4 *Mature Worlds*, Meta's name for worlds that are sexually suggestive, promote legal drugs or contain gambling.
- Three Mature Worlds were found to have sexually explicit content, including wearable naked bodies, naked pole dancers and group sex games. These worlds do not have any additional controls to prevent minors from accessing them.

Minors present in nearly 7 out of 10 worlds

- Researchers logged into the most popular 'worlds' in Meta's flagship Horizon Worlds social app for VR 100 times and found minors present in 66 out of 100 worlds.
 - On each log in to Horizon Worlds, researchers visited worlds with a high number of active users from the app's 'Top 100' worlds menu, exploring the world fully for a maximum of five minutes.
 - Worlds are virtual social spaces, some of which have been created by Meta but the majority have been created by users.
- Meta does not provide information on the declared ages of users in Horizon Worlds. To conduct our analysis two researchers assessed recordings to determine if minors were present, marking a World as containing minors if the recording contained:

- Users who stated that they were younger than 18
- Users with voices that both researchers agreed were aged 15 or younger
- This analysis revealed that users determined to be under 18 were present in 66 of the 100 recordings.
- Meta has stated that Horizon Worlds is currently only open to users aged 18 and up.

Minors are accessing 'Mature Worlds'

- Across our recordings we logged in to 'Mature Worlds' 19 times, found that minors were present on five occasions.
- "Mature Worlds" is Meta's name for worlds where users can create sexually suggestive content, promote legal drugs, or contain gambling. These worlds must be labeled 18+.
- Mature Worlds are recommended through Horizon World's *Top 100* menu where they mingle freely with non-mature worlds.
- Minors were identified as being present in four different Mature worlds, three of which were virtual strip clubs. In these worlds we recorded:
 - A minor serving alcoholic drinks in a strip club for in-game money.
 - A minor convinced adult users he was over 18 and was allowed to stay in a world with drugs and alcohol. Researchers found his Facebook profile which confirmed he was underage.
 - Two adults offered minor sexual images of their feet before realizing he was a minor. The young user had set his username to "SendToePics".
- Throughout our investigation of the Mature Worlds, we found sexually explicit content created by users, including:
 - A group sex game, with questions such as "Have you ever sent a nude photo to someone in the Metaverse?".
 - A lockable bedroom for "seven minutes in heaven", where two users enter for seven minutes together.
 - A "Hot Girl Playhouse" where users are awarded points if they complete a list of increasingly suggestive tasks, complete with a leaderboard and strip pole.

- A naked body in a sexually explicit position, wearable naked body parts, and rotating naked pole dancers.
- We recorded two instances of portals to Mature Worlds being opened in normal worlds.
- Portals are gateways to other worlds and Meta allows users to place them anywhere, leading other users to any World of their choosing, including Mature Worlds.
- We recorded:
 - A portal to a Mature World dedicated to Marijuana opened at the entrance of the “Venues” World, the main World users first enter. A Meta employed moderator was present and encouraged the users who placed the portal.
 - A portal to a Mature World was opened in front of an underage user, leading anyone who entered to a strip bar serving alcohol and marijuana.

Minors subjected to harassment and abuse in 19 incidents

- Researchers recorded their visits to popular worlds 100 times and found 19 incidents of minors being harassed or abused by adult users.
- In four of these incidences minors were on the receiving end of sexually explicit insults:
 - In a world modeled on a stand-up comedy club, an adult asked a young user “Do you have a cock in your mouth?” after being heckled by the minor.
 - One adult user told a minor he was going to “suck his balls”.
 - An adult user repeatedly shouted “I don’t want to cum on you” at a group of young girls, even after they asked him to stop and said they were underage.
- Researchers recorded three instances of racist harassment, including:
 - In a virtual courtroom, a minor with a black avatar was told “you’re black, you’re sentenced to death, get out of here”.
 - A user at a stand-up comedy club called a minor with a black avatar a “monkey”.
- Researchers recorded five instances of gendered harassment. Including:
 - An adult man shouted “Yeah boys, run most of the country and protect y’all, dumb ass hoes” at a group of young girls.

How we conducted this study

- We recorded 100 instances of in-app footage inside the most popular worlds.
- Worlds were selected from the *Top 100* worlds menu, which ranks worlds by how many likes they have.
- Some high ranking worlds were empty, and so the researcher only entered worlds that were populated at the time of recording.
- In each recording the researcher entered one world and stayed there until the world had been fully explored, staying for a maximum of five minutes.

How we determined if a user was under 18

- Meta does not provide information on the declared ages of users in Horizon Worlds. To conduct our analysis two researchers reviewed the recordings and marked the world as containing minors if either:
 - A user declared they were under 18.
 - Both researchers agreed that a user sounded under 15 years old.

Horizon World's age policy

- Meta announced in December 2021 that Horizon Worlds was open to all users aged 18 and up.

Horizon Worlds Opens to Those 18+ in the US and Canada, Meta Quest, 9 December 2021, <https://about.fb.com/news/2021/12/horizon-worlds-open-in-us-and-canada/>

- Meta is planning to open up the social network to teens aged 13 to 17, potentially as soon as March 2023.

Meta Pursues Teen Users as Horizon Metaverse App Struggles to Grow, Salvador Rodriguez, The Wall Street Journal, <https://www.wsj.com/articles/meta-to-revamp-horizon-metaverse-app-plans-to-open-for-teen-use-as-soon-as-march-11675749223>

- Meta’s app store for its Quest headsets gives Horizon Worlds a ‘Parental Guidance’ maturity rating.
- Researchers tested this and found that Horizon Worlds can only be downloaded by users who set the age of their Meta account to 18 or older.
- However, researchers found cases where minors spoke to Community Guides, in-app moderators employed by Meta, and were not removed from Horizon Worlds.

Meta’s safety commitments

- Mark Zuckerberg committed to building safety into the Metaverse from day one when launching the Metaverse.

“Like I said earlier, interoperability, open standards, privacy and safety need to be built into the Metaverse from day one. And with all the novel technologies that are being developed, everyone who’s building the metaverse should be focused on building responsibly from the beginning.”

“The Metaverse and How We’ll Build It Together – Connect 2021”, Meta, YouTube, 28 October 2021, <https://youtu.be/Uvufun6xer8?t=3104>

- Meta has detailed their standards for user behavior in the “Code of Conduct for Virtual Experiences”, violations include the sexual harassment of minors, bullying, harassing, stalking or hateful behaviour, and any form of non-consensual intimate behaviour.

Code of Conduct for Virtual Experiences, Meta, November 24 2022,

<https://www.meta.com/en-gb/help/quest/articles/accounts/privacy-information-and-settings/code-of-conduct-for-virtual-experiences/>

Minors in Mature Worlds

- “Mature Worlds” is Meta’s term for worlds where users can create sexually suggestive content, promote legal drugs, or contain gambling. These worlds must be labeled 18+.

Meta Horizon Worlds Mature and Prohibited Worlds Policy, Meta, September 2022,

<https://www.meta.com/help/quest/articles/horizon/create-in-horizon-worlds/restrictions-to-worlds-in-horizon/>

- Despite Meta’s claims that Horizon Worlds is open only to users aged 18 and over, our investigation found minors present in five of the 19 Mature Worlds that researchers visited, equivalent to finding minors in 1 out of 4 Mature Worlds.
- The instances were across four worlds, three of which were virtual strip clubs:
 - The Vibe (18+) – A club with an interactive strip pole, bar serving alcohol, and private rooms.
 - The Rabbit's Den – A club with an interactive strip pole, private rooms, and a bar where users can serve alcoholic drinks for in-game money.
 - Black Magic Comedy, Karaoke & Chill – A bar with alcohol and marijuana.
 - Coyote Fugly Bar – A bar with interactive strip poles, a bar serving alcohol, and interactive sex toys.



Coyote Fugly Bar, Horizon Worlds

- When an underaged user entered Coyote Fugly Bar two adult users offered to sell him sexual pictures of their feet before realizing he was under 18. His username was 'SendToePics'.
- We found a minor serving alcoholic drinks in The Rabbits Den, following recipes provided by the world’s creator.



The Rabbits Den, Horizon Worlds

- An underage user was allowed to stay in Black Magic Comedy after convincing the other users that he was 23. Researchers found the Facebook profile connected to his Oculus account, confirming that the user was under 18.

Explicit content in Horizon Worlds

- Throughout this investigation we found sexually explicit content made by other users in Horizon Worlds.
- We found group sex games with questions such as “Have you ever sent a nude photo to someone in the Metaverse?” and “Show the group your sexiest moan for 10 seconds”.



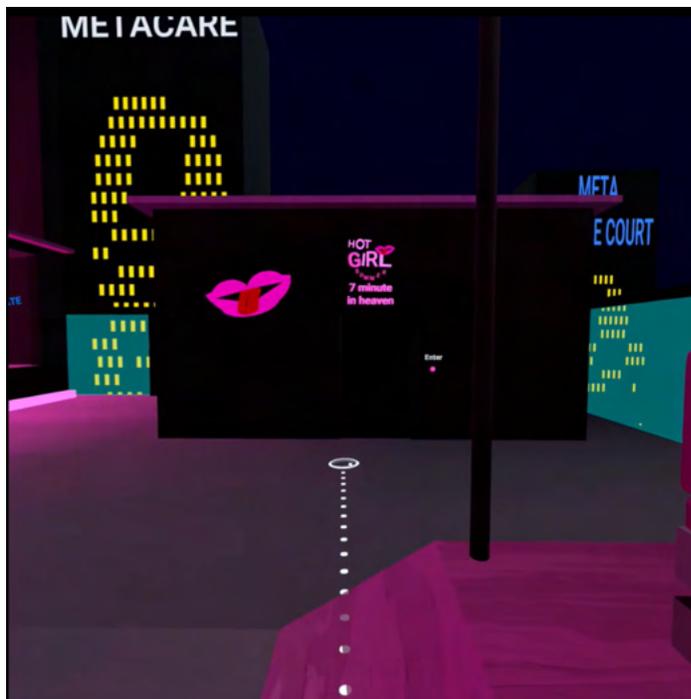
The Adult Club, Horizon Worlds

- A ‘Hot Girl Playhouse’ with a strip pole, where users are awarded points if they complete a list of increasingly suggestive and risky tasks, such as 45 points for going “all the way in front of an audience”, complete with a leaderboard for high-scoring users.



Hot Girl Summer, Horizon Worlds

- A bedroom that can be locked for two users to have '7 minutes in heaven' together.

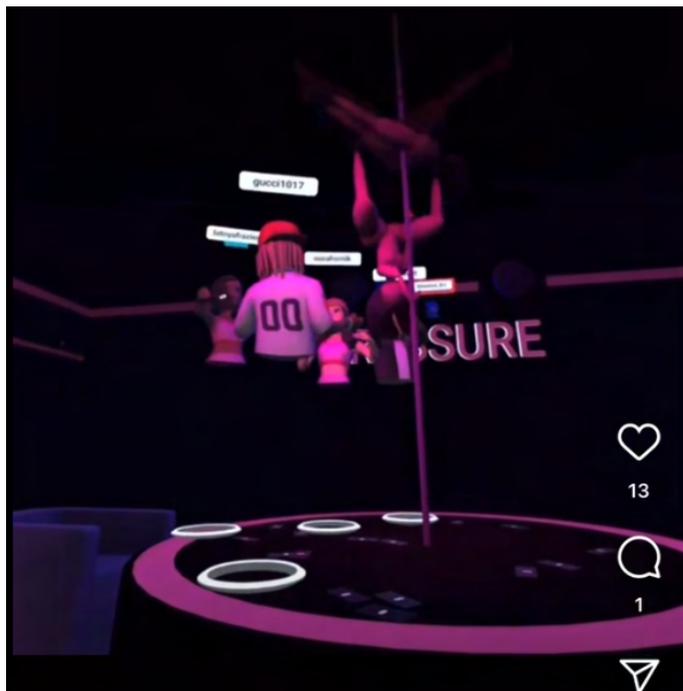


Hot Girl Summer, Horizon Worlds

- Pressure, a strip club with a wearable naked body, rotating naked pole dancers, and naked bodies in sexually explicit positions.



Shadow_bebop, Instagram, 26 October 2022,
<https://www.instagram.com/p/CkMELYzOBON/>



Shadow_bebop, Instagram, 28 October 2022,
<https://www.instagram.com/reel/CkRCYqorpdO/>

- Through our investigation we also found worlds dedicated to marijuana, including interactive bongos and pipes.



Jamaica Paradise Club, Horizon Worlds

Minors receive racial and sexual harassment in the metaverse

- Researchers recorded 100 visits to popular worlds and found 19 instances of minors being harassed by adult users.
- In four instances minors were on the receiving end of sexually explicit insults.
- At a stand-up comedy club an adult asked a minor “Do you have a cock in your mouth?” after being heckled by the minor.



The Soapstone Comedy Club, Horizon Worlds

- In another instance, an adult user told a minor he was going to “suck his balls”.

- After losing a game in one world, an adult user repeatedly said “I don’t want to cum on you”, despite the minors asking him to stop and saying they were minors.



Party House, Horizon Worlds

- An adult told a young girl that “your dad should have swallowed”, to which the girl replied “I’m only 12”.
- We recorded three instances of racist harassment and five instances of gendered harassment directed at minors by adult users.
- In a virtual courtroom a minor with a black avatar was told “you’re black, you’re sentenced to death, get out of here”.



META COURT 2.0, Horizon Worlds

- At a virtual house party an adult user shouted misogynistic insults at a group of young girls, "Yeah boys, run most of the country and protect y'all, dumb ass hoes".
- At a stand-up comedy club a user on the stage called a minor with a black avatar a "monkey".



The Soapstone Comedy Club, Horizon Worlds

Horizon Worlds provides easy access to adult content

- Mature Worlds have to be marked as 18+, but they are otherwise accessible to any Horizon Worlds user through the menu or portals.

"Meta Horizon Worlds Mature and Prohibited Worlds Policy", Meta, September 2022, <https://www.meta.com/en-gb/help/quest/articles/horizon/create-in-horizon-worlds/restrictions-to-worlds-in-horizon/>

- Portals are doorways to other worlds. Meta allows users to place portals anywhere in Horizon Worlds, and these portals can lead to any worlds including Mature Worlds.
- We found that users are opening portals to Mature Worlds in worlds frequented by minors, leading minors directly to adult content.
- Two users opened a portal to *CANNACULTURE*, a world dedicated to marijuana, at the entrance to Venues which is the main hub for Horizon Worlds and the first world most users visit.
- Meta employs in-app moderators called Community Guides. One spoke to the users who opened the portal but encouraged the users and did not mention that a portal to a Mature World was inappropriate for the main hobby.



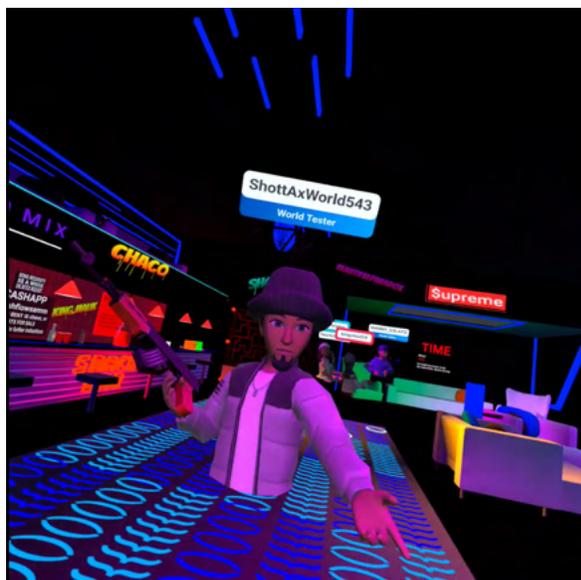
Venues, Horizon Worlds

- In another instance a portal to *CHICAGO PENTHOUSE 18+*, a strip club and bar, was open in front of an underage user.



Party House, Horizon Worlds

- Going through the portal, our researcher was asked to confirm their age by a user holding a gun.



CHICAGO PENTHOUSE (18)+, Horizon Worlds



© Center for Countering Digital Hate Inc

Redacted for Congress

DECLARATION OF JASON SATTIZAHN

I, Jason Sattizahn, pursuant to 28 U.S.C. § 1746, hereby declare as follows:

1. I am a person over eighteen (18) years of age and competent to testify. I make this Declaration on personal knowledge and in support of my whistleblower disclosure submitted to government bodies of jurisdiction.¹

2. In this declaration, I am supplementing my declaration dated March 10, 2025 (“original declaration”) to address three areas of my disclosure with increased specificity:

- a. The ubiquitousness of Meta employees routinely working directly across Meta’s products internally (e.g. working on Virtual Reality and Instagram products), and thus having material knowledge of the work of multiple divisions;
- b. Meta’s years-long, company-wide mandate to prioritize increasing engagement from children (often referred to as “TAYA” or “teens and young adults”); and
- c. Meta’s long history with avoiding appropriate investments into age verification and parental controls.

3. One aspect of Meta that is not well understood to those outside of the company is that internally, Meta has little to no restriction or segregation between who can and does work across their various products.² In fact, Meta’s various major products – whether they are Facebook or Instagram – are not siloed as if they are separate, walled-off corporate entities. While each major product can be said to be a part of Meta’s “Product-Based Divisions”, Meta generally operates like one single integrated company. These divisions share resources, systems, and a top-down strategy employed by Meta. Moreover, every employee at Meta, whether they spend the bulk of their time working on Facebook or Virtual Reality products, all receive a paycheck from Meta. Despite Meta’s

¹ Whistleblower Disclosure package submitted entitled: “Protected Disclosure of Harmful and Deceptive Practices at Meta”.

² “Products” as used here refers to the major Meta product lines; this includes Meta’s “Family of Apps” (i.e. Facebook, Instagram, WhatsApp), Reality Labs hardware (Virtual Reality, Wearables) and other company investments, such as Meta AI

Redacted for Congress

explicit, public efforts to make their products appear separate from one another (e.g. Instagram vs. Virtual Reality), inside the company, employees constantly work across Meta's products. This is due to numerous factors – which I will outline below – but the end result is that internally at Meta, there are no hard walls between “Facebook”, “Instagram” or “Reality Labs” when it comes to what employees and researchers work on. **This means that employees or researchers in areas like Virtual Reality in Reality Labs, including myself and the five other whistleblowers making this disclosure often have direct, hands-on knowledge and impact on other Meta products such as Instagram.** This also means that when Meta attempts to sidestep questioning about Instagram or Marketplace, as it has in congressional testimony, it is an empty argument and misrepresents the reality. One reason employees have to work so intimately across Meta's products is out of sheer necessity because Meta blends and integrates their products and experiences together with one another in the pursuit of making every offering part of a complete social interaction. For instance, Meta often integrates content from one of their products (e.g. Reels from Instagram) into another of their product's experiences (e.g. Facebook) as a means to increase distribution of content and increase user engagement *across products*. This requires their employees internally to work together irrespective of their product designations, learning how each product works, and building these integrated experiences together.

4. Furthermore, Meta employees work across Meta's products because Meta has intentionally structured multiple teams to explicitly work across their products at the same time. Meta has numerous “horizontal” teams internally at the company. These horizontal teams focus their work on concepts or investments that apply to multiple or all of Meta's products (i.e. Facebook, Instagram, Virtual Reality) at the same time. This means that if you are staffed on one of these teams *or* if your work requires collaboration with these horizontal teams, your work is intimately linked to all of Meta's products at once. One example from my own experience is Meta's Avatar team. Meta's

Redacted for Congress

Avatars allow users to create a digital representation of themselves, and these avatars are used across Meta's products such as Facebook, Instagram, and Virtual Reality. Another example – pertinent to user safety – is Meta's "Central Integrity" team, a group at Meta responsible for addressing a range of harmful behavior across Meta's products, including Facebook, Instagram, WhatsApp, and Messenger. I personally worked with both of these horizontal groups in my time at Meta, as did many other researchers I know.

5. Meta leadership has prioritized working outside of one's immediate team by tying explicit monetary incentives to collaboration. Employees – and especially researchers – are encouraged and rewarded to work across Meta's products, as it is strongly weighted as an indicator of high-performance in the performance review cycles that determine employees' bonuses, salary increases, and promotions. My consistently high performance evaluations at Meta placed a high value and praised me for my work on other products, collaborating and applying research across multiple areas of Meta. Working across teams and products was piecemeal but common over my six years at the company. Rather than offer exhaustive examples of my own work across Meta's products, I will drill down on two specific examples of work I performed between Instagram and Reality Lab's Virtual Reality products together.

6. As described in my previous declaration (p73-75), I was involved in the VR age assurance project, a research study kicked off in July 2022 to understand how Meta can collect more accurate data on the ages of their users. The purpose was not just to understand how to collect accurate age data, but how to do so while respecting user comfort and privacy. This was key because research had years-long evidence showing that Meta did not have accurate age data on their users *across products*, and that Meta users misrepresented their age largely because people do not trust Meta to collect their personal information. Although the origin of the study came from my colleague on the Virtual Reality Youth team, the research study itself was an intimate collaboration with the Instagram Youth

Redacted for Congress

team as well as Youth teams on other products. The Instagram Youth team and their researcher, [REDACTED], not only helped develop the project, but they committed and planned their own recommendations on how Instagram should alter its age data collection approach based on the findings of the study. Simply put – this was neither an Instagram-only or a Virtual Reality-only project, but rather it was an Instagram *and* Virtual Reality project. This means that in fall of 2022, when Meta Legal interfered and altered the design of the research, they were doing so for both Virtual Reality and Instagram. From the information shared with me by Meta management, including, Ayfer Gokalp, it also means that when the project was cancelled *after* all approvals had been made and a budget had been allocated for the work, there was only one person who had the authority to cancel the project – Mark Zuckerberg. With this sudden cancellation, Meta’s ability to use research to collect better age data on users and children was kneecapped for Instagram and Virtual Reality simultaneously. It is important to note that documented feedback from August 18, 2022 from Elaine Dai from Meta’s Legal leadership) on this age assurance project explicitly referred to Frances Haugen’s whistleblower disclosure in 2021 as a “leak”, further highlighting Meta’s internal campaign to demonize and denigrate researchers and employees who became lawful whistleblowers in the pursuit of holding Meta accountable for their actions.

7. In my original declaration, I also discussed my direct involvement in something called “Horizon Feed” and my work to ensure user safety with this new Meta investment. Generally speaking, Horizon Feed was an effort to have people in Virtual Reality be able to digest “feed content” similar to how you would on Instagram but in an immersive environment. When viewing this “feed” while wearing a Virtual Reality headset, users would see diverse content, such as VR-related videos (i.e. gameplay) and content being pulled in from Instagram (i.e. Reels, videos, etc.). The explicit reason for the integration of Instagram in Virtual Reality – given to me and multiple

Redacted for Congress

personnel across teams – given by Tim Loving himself was that Instagram’s content in VR would provide a strong way to increase user engagement.

8. The integration of new, unrelated content (i.e. Instagram Reels) into a new product (i.e. Virtual Reality) presented significant, foreseeable safety risks as well as likely dangers to users and children that had not been researched (see Alpha declaration p.135). One of Meta’s first steps in creating this Horizon Feed was to first integrate user-generated content (“UGC”; short videos, etc.) into the Virtual Reality experience, alongside content from places like Instagram (i.e. pictures, videos). In response, in early 2024, I collaborated across a number of Meta teams to develop a set of research studies that would help us understand how to keep users and children safe when viewing real-time feed content (UGC, Instagram Reels, etc.) within a Virtual Reality space. The research itself focused on how to develop stronger safety tools for users before bringing this feed into Virtual Reality, including (i) effective parental controls, (ii) direct user controls (e.g. improved reporting and feedback on content shown), and (iii) how Meta could proactively alter their ranking algorithm to prevent show harmful content from reaching Virtual Reality users. Out of an abundance of caution at the outset, I advised the teams involved that there was a chance our research may reveal that Meta would need to slow or halt the integration of feed content (i.e. UGC, IG reels, etc.) into Virtual Reality due to safety concerns and safety investments Meta should make before opening this door. Despite the research being fully designed and emphatically supported by all teams involved, the first research study I was kicking off was inexplicably cancelled by Ayfer Gokalp and Tim Loving from Meta leadership on April 8, 2024, with the only explanation given that the project was “not viable.” After my repeated requests for more information about the basis for the sudden cancellation, Tim Loving stated that the reason the study was shut down was because the project focused on “salsa users” (referring to 10-12 years olds). I emailed Mr Loving, stating, “I’m confused on this. Horizon Feed *isn’t being released to SALSAs users at all (emphasis added)*. This is indicated in the first bullet point

Redacted for Congress

of the research brief I wrote, in the section, ‘Business objectives and motivations,’ where it says, ‘SALSA users (10-12yo) will not see HZ Feed or UGC.’ If it wasn’t clear enough to Mr. Loving from my email, the research brief also had a second bullet stating, “Feed/UGC is limited to 13+ yo VR users.” Mr. Loving never responded to my email. Effectively, this research on how to make both Instagram and Virtual Reality safer for children, parents, and any adult was shut down without explanation. I understood these events to reflect the priority goals as previously emphasized by Tim Loving - that the integration of Instagram into Virtual Reality would increase user engagement, including for children to whom the experience would be released.

9. It is imperative to state that Meta had a years-long, company-wide mandate for employees to focus on increasing the engagement of Meta’s products by children (often referred to internally as “TAYA” or “teens and young adults”). Meta’s internal documents explicitly state that it is a “...company-level priority for ‘Meta’s platforms to lead with future generations of youth’” and that they are “...taking proactive approach to Youth (10-17), incl.: ‘building for youth’, ‘legitimacy & advocacy’, ‘winning over parents’, and ‘regulatory readiness’.” Meta’s prioritization to increase engagement of children across their platforms isn’t new, but rather, it has been ongoing for years.

10. After I joined Facebook Marketplace in 2018, I recall when Facebook made an explicit push to increase the engagement of children across all of its products. By the time I moved to VR in 2022, I had seen reports on the importance of increasing the engagement of children (TAYA) across Meta’s products, including Marketplace, Facebook Gaming, and in particular, Instagram. Most disturbing was the breakneck speed Meta put into boosting children’s engagement *after* Frances Haugen’s disclosure shined a light on the negative emotional impacts of Instagram on youth.

11. By the time I left Meta in mid-2024, I had seen dozens of internal documents that Meta had produced on the importance of increasing the use of Instagram by children. There are numerous youth-focused documents showing Meta directly focusing on increasing children’s engagement with

Redacted for Congress

Instagram (see exhibits to the sworn declarations of the six whistleblowers). To illustrate this point,

I will review themes that emerge across these documents:

- a. Engagement. Multiple documents show that it is the youngest users on Instagram who drive the highest engagement, and some documents explicitly outline Meta's growth-goals around youth for Instagram. For example, one document from 2023 states, "We need to drive an additional +3.5M DM Teen DAU in developed markets (+6.3%) on top of what we already ship (+2.6%)". Other documents describe which teens are "hardest-to-reach", the barriers Meta is working to overcome the lack of engagement and outlines paths Meta can make to increase engagement with youth. This pattern for technology adoption is well understood in the industry and drives marketing strategies for product launches.
- b. Emotion. Numerous documents focus on how Instagram can leverage emotional aspects of Instagram's experience to appeal to youth. One report states. "In order to make IG feel more playful, we need to design for both utility AND emotion." Another report from 2024 states, "Humor is like a cheat code that supercharges the connection and validation [for] teen participants desire..." What is unsettling is that there are multiple reports outlining how Meta will readily invest in researching youth emotions if it can be framed positively and will lead to increased engagement by children. However, when my and others' research attempted to research the entire experience including both the positive and negative emotional aspects of Meta's products, Meta's Legal teams and leadership repeatedly shut the research down, explicitly stating that the data was "too risky" to have.
- c. Parental Controls. My own research on Virtual Reality in 2023 showed the low adoption rate of Meta's parental controls, mirroring internal data from Instagram

Redacted for Congress

around the same time showing that only .5% of teen accounts had turned on parental controls. This wasn't a surprise, though – one of my first exposures to cross-product insufficiencies of Meta parental controls was actually when I was still working on Facebook Marketplace from 2018 to 2020. I recall seeing work on Instagram and across Meta's products directly highlighting why parental controls as built were neither sufficient nor effective to keep children safe. Issues included, but were not limited to: (i) the general increase of a burden that digital controls add to parents' (and children's) lives, (ii) parents' lack of time or tech literacy in managing these tools, (iii) a mismatch of how the tools are built with a parent or guardian's preference in managing digital access. More importantly, effective parental controls are predicated on the accuracy of Meta's age data, the assumption that parents have account access, are present with children in parts of set up, and are willing to provide information about their children to Meta.

12. The lack of age verification is the crux of the matter. When Meta offers public assurances about their investments around parental controls (e.g. content controls) and child safety (e.g. age-appropriate content, account restrictions), Meta is misleading the public based on what they don't say. For Meta's efforts to be effective, Meta would need accurate age data. Age data (or associated behaviors triangulating age) would be the explicit way for Meta to flag a user's account as mismatching their age or as some content as being outside of an appropriate range for an account. Unfortunately, to my understanding Meta has never had accurate age data, and from what I was told during my time there, they do not want it because it would trigger additional regulatory compliance inconsistent with their engagement priorities. In fact, while preparing for one of my first research studies in 2018, I was told by a lead Facebook researcher, "Don't trust our internal age data, collect it yourself." Meta claims that they are taking steps to improve experiences for Youth, such as their

Redacted for Congress

September 2024 announcement of “Instagram Teen Accounts.” For a detailed explanation of why Meta’s efforts in this regard are misleading, see my original declaration (Alpha declaration, beginning at p.139). Important to note, though, is that our Age Assurance project was approved to understand how to collect better age data for Instagram and Virtual Reality in mid-2022. After the project was suddenly shut down as discussed in paragraph 6 above, to my knowledge there was no further research or investment on how to collect better age data on Instagram between 2022 and 2024. It was only in response to congressional government pressure in Mark Zuckerberg’s January 2024 Senate Judiciary hearing that something happened. In September 2024, Meta’s Head of Product, Naomi Gleit, stated, “We’ve been working on [Instagram] teen accounts for *several months*” (emphasis added). This statement was significant that for nearly two years after cancelling our age assurance project to improve Instagram age data, *Meta did nothing*. Only after public pressure, did Meta make even meager efforts to appease the public.

13. For years across the company, Meta knew it needed to do more for identity and age verification and chose not to. In fact, nearly a decade ago when I joined Facebook Marketplace, I actually worked directly on the Marketplace identity and age verification project. The idea was simple: create a system where Marketplace users could upload a government ID to validate their identity and age. On one hand, the project was for tax purposes (i.e. someone sells a certain dollar amount on Marketplace). On the other hand, the Facebook Marketing team wanted to use identity and age verification as a way to offer people who verified their identity and age a “trusted person” badge. The goal was that by offering these badges on accounts, people would see that Marketplace users were “Trusted” and user engagement of Marketplace would increase. Unfortunately during my research, I uncovered numerous issues with the age verification process, including: (i) Facebook users simply didn’t trust giving Facebook this information, (ii) ID verification could be easily sidestepped by using fake credentials, and (iii) the process gave users false assumptions and

Redacted for Congress

heightened concerns about their safety (e.g. saved ID could be forwarded to the police in instances of harm to users; this is not true). I attempted to delay the release of the “Trusted” badge on Marketplace, citing these concerns as a reason to first improve the verification process. However, the badge was released anyway. Following the release of this identity and age verification workflow, an audit of accounts who had the “trusted” badge showed that Facebook accounts with the “trusted” badge were actually *more* likely to be bad actors and scammers than accounts without the badge. Quite literally, Meta was telling Facebook users that certain accounts were “Trusted” when they were more likely to be causing harm, all because Meta thought this would increase user engagement.

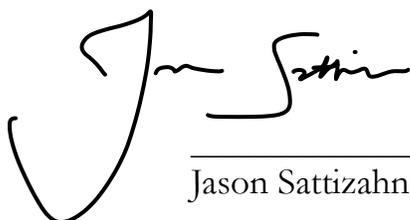
14. One refrain Meta has used to defend claims about its well-known age problem is that verification systems are generally too difficult to develop, which leads to problems like this within their products. Of course, no system is perfect, but it is important to highlight that Meta can do much more to verify the identities and ages of their users for safety reasons, and Meta knows this. The simplest thing to do would be to quite literally pay a statistically representative subset of Meta users to explicitly verify their age across a wide range of ages. In response, Meta could integrate this data and retrain their existing data sets to appropriately flag accounts or content that mismatch a user’s age. I personally discussed this approach internally across multiple teams and researchers in 2018, but that clearly did not come to fruition. In fact, even before I joined Meta and worked on projects such as their verification systems, other companies had made advancements in user and account age verification. While I was employed at Sony PlayStation in 2017, Sony actually rolled out a meaningful age verification system. Any time a new adult PlayStation account was created, the account was required to undergo verification via checking against the account holder’s credit card. Of course, account creation is just one of multiple places in the experience of a user where age verification may need to occur, but as of 2025, Meta’s products do not require consistent verification of adult accounts when new accounts are created. Meta has pointed to their policy to require age

Redacted for Congress

verification if they *suspect* a user is underage as a safety tool but again - what they leave out is that Meta does everything they can to avoid “suspecting” their users are lying about their age.

I do solemnly affirm under the penalties of perjury and upon personal knowledge that the contents of the above statement are true to the best of my knowledge.

Date: September 26, 2025



Jason Sattizahn