AMENDED Exhibit 121

PLAINTIFFS' OMNIBUS OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

Case No.: 4:22-md-03047-YGR
MDL No. 3047
In Re: Social Media Adolescent Addiction/Personal Injury Products Liability Litigation

Instagram Appearance Comparison Decision - A/C Priv

Decisions Needed:

1. Can IG move forward in development of Appearance Comparison metrics, including the deployment of user surveys, for potential future goaling provided that IG addresses CWB feedback on survey development, risk mitigations, and metric validation plans?

Deciders:

Central Social Impact / Research: **Pratiti Raychoudhury**
IG / Youth: **Miki Rothschild**

Full RACI [HYPERLINK "https://docs.google.com/document/d/1E1ZUnyPzriP5U3CLHO6HGaRE7hAKbjQvLuFJrqxha8Q/edit" \h]

Timeline:

- Needed for a decision: **mid-March** (email review to start March XYZ, decision triggers urgent H1 execution or roadmap replanning)

Summary:

- **Motivations/Expected Benefits:** Given current IG Well-being strategy focus on Appearance Comparison, IG would like to continue investment in development of an appearance comparison metric to assess success of in development products and potentially for future goaling.
- **Risks:** Central Well-being (CWB) believes that the risks associated with continuing this sensitive-topic metric development outweigh the benefits given lack of clarity on theory of change for current products, lack of clear alignment with company well-being strategy, and in development alternative metrics for problematic use.
- **Mitigations:** If IG is able to continue pursuing AC metric development, CWB strongly recommends multiple survey quality and risk mitigation strategies.

Decision Details

AC Measurement Motivations and Expected Benefits

1) Instagram decided to focus its mental well-being strategy on appearance comparison because:
    a) **Appearance comparison is experienced by many in our communities.** An estimated 33% of Instagram DAU across 10 countries (FCI + Mexico) compare their appearances to others' on Instagram "often" or "always"; 48% "occasionally" or "never" compare their appearances. An estimated 48% of teen girls compare

their appearances to others “often” or “always”, and 30% “occasionally” or “never” compare their appearances ([HYPERLINK "https://fb.workplace.com/notes/1065141417573177" \h]).

b) **Academic research suggests a causal link between Instagram use and appearance comparison and/or body image issues.** Longitudinal studies indicate a bidirectional link between social media use and body image issues ([HYPERLINK "https://www.sciencedirect.com/science/article/pii/S1740144521001327" \h], [HYPERLINK "https://www.sciencedirect.com/science/article/pii/S0747563217302157" \h], [HYPERLINK "https://doi.apa.org/doiLanding?doi=10.1037%2Fppm0000206" \h], [HYPERLINK "https://onlinelibrary.wiley.com/doi/10.1002/eat.22640" \h], [HYPERLINK "https://journals.sagepub.com/doi/10.1177/1461444819894346" \h]); lab experiments in which participants were assigned to browse appearance-relevant content also demonstrated increases in body-related concerns compared to participants assigned to browse appearance-neutral material ([HYPERLINK "https://www.sciencedirect.com/science/article/pii/S1740144519303754?via%3Dihub" \h], [HYPERLINK "https://psycnet.apa.org/record/2019-14939-001" \h], [HYPERLINK "https://pubmed.ncbi.nlm.nih.gov/25035882/" \h], [HYPERLINK "https://www.sciencedirect.com/science/article/pii/S1740144517305326" \h], [HYPERLINK "http://www2.psy.unsw.edu.au/Users/Ivartanian/Publications/Fardouly,%20Diedrichs,%20Vartanian,%20&%20Halliwell%20(2015).pdf" \h]). Appearance comparison is typically believed to be the link between social media use and body image ([HYPERLINK "https://www.sciencedirect.com/science/article/pii/S1740144514001375" \h], [HYPERLINK "https://isiarticles.com/bundles/Article/pre/pdf/36346.pdf" \h], [HYPERLINK "http://www2.psy.unsw.edu.au/Users/Ivartanian/Publications/Fardouly,%20Diedrichs,%20Vartanian,%20&%20Halliwell%20(2015).pdf" \h]), particularly on image-based social media platforms ([HYPERLINK "https://www.sciencedirect.com/science/article/pii/S0736585319308111" \h], [HYPERLINK "https://www.sciencedirect.com/science/article/pii/S0747563220300832" \h]).

c) **This is one of the most significant external criticisms we’re facing with regard to teen well-being.** Regulators and press are pushing us to address social comparison and negative appearance comparison *specifically* ([HYPERLINK "https://www.blumenthal.senate.gov/newsroom/press/release/blumenthal-and-blackburn-introduce-comprehensive-kids-online-safety-legislation" \h], [HYPERLINK "https://www.whitehouse.gov/briefing-room/statements-releases/2022/03/01/fact-sheet-president-biden-to-announce-strategy-to-address-our-national-mental-health-crisis-as-part-of-unity-agenda-in-his-first-

state-of-the-union/" \h]). We've also made our commitment to addressing social comparison a primary proof point in our recent external narratives ([HYPERLINK "https://about.instagram.com/blog/announcements/using-research-to-improve-your-experience" \h], [HYPERLINK "https://about.instagram.com/blog/announcements/pausing-instagram-kids" \h]).

2) There was leadership [HYPERLINK "https://www.dropbox.com/s/ql6eicl4kygszzh/Re_%20Answering%20your%20question%20from%20Q%26A%20on%20Social%20Comparison.eml?dl=0" \h] between Adam and Pratiti in June 2021 on IG proceeding with negative social comparison IG proactive work through the lens of agency and control, one of the core mechanisms for how Meta can help reduce Problematic Use.
3) IG is defining success for its teen mental well-being investments to focus on reducing appearance comparison, with IG leadership alignment.
4) IG's H1 bets were framed around generalized agency & control solutions due to uncertainty around its use of AC. These bets are: (1) Take a Break (long-session reminder) and (2) Topic Nudge (in-feed nudge on long scrolling in recs to alternate topics). The AC measurement that is being discussed here is partly intended to inform H1 decision making by enabling IG to weigh likely engagement regressions with user benefit (as measured by AC sentiment).
5) IG aims to unblock new product bets beyond H1, where AC is more deliberately targeted with appropriate theories of change. Having clear approval to measure AC will enable IG to design more targeted interventions both within the current H1 bets and to pursue other AC reducing pathways beyond agency & control (current [HYPERLINK "https://docs.google.com/presentation/d/19JpzbdOFo6PwDnRF-AGLOL-cbtyDbNGgXQ8UJSLvpjc/edit" \l "slide=id.g1048090c7e9_8_37" \h]).

AC Measurement Risks

1) **Sensitive Topic Risks.** Asking users on platform about sensitive topics (like appearance comparison) come with a number of risks that must be balanced against likely benefits. These include (1) negative user experiences (i.e., users do not come to our platforms to be asked about high sensitivity topics like AC); (2) concerns about having data on such experiences retained in our systems; and (3) comms risks (e.g., screenshot risks). (4) Unclear or unknown consequences of optimizing for this type of metric (e.g., it may reduce important and healthy content too, like motivational and inspirational posts).
2) **Benefits may not outweigh these risks:** The CWB team has concerns that the benefits of AC survey measurement do not outweigh these risks. These concerns include:
    a) **Appropriate theory of change:** There is not clear theoretical or empirical basis to expect that Take a Break should impact AC. While there may be more of a path for Topic Nudges, this depends highly on product design specifics and would require that the product target and reduce the prevalence of appearance-

Commented [1]: If we are pursuing more longitudinal measurement, maybe we can remove this?

comparison content. If we are taking on the inherent risks here, there should be a very clear expected linkage between intervention and measured outcomes. If IG has a path forward to building substantially more targeted interventions, and will not be relying on the AC metrics for core decision-making on current generally scoped products, concerns are substantially lessened.

b) **(Mis-) Alignment with Core Company Wellbeing Strategy.** Unless anchored in Problematic Use, appearance comparison is not an active core topic prioritized in the x-Meta strategy. This does not mean that teams cannot work on this topic or that it is a bad topic to work on (as per previous leadership alignment), but it does mean that expected benefits to the company are lower in the overall risk calculation.

Commented [2]: I think we should remove this bullet. Just because it's not in the x-meta WB strategy does not mean expected benefits to the company are lower.

3) **Alternative Measurements.** Given these concerns, CWB recommends that IG pursue different metrics for currently proposed products.
   a) **[*Short term*] Usage Metrics:** IG switches to focus on usage metrics (ex. how many people click Topic Nudge to end long-scrolling on one topic). This tells us whether people find value in product interventions but not if the interventions reduce outcomes like appearance comparison. This is the status quo.
   b) **[*Medium term*] Switch to Problematic Use (PU) measure.** IG switches to actively invest sooner than planned in Central Well-being's PU measure. CWB has started a new initiative to develop company-level well-being metrics around the five prioritized well-being topics, including Problematic Use (PU). A PU v0 measure will be identified at the end of H1 or soon after.
      i) Benefits of using PU instead of AC include (1) clearer theory of change for current product portfolio and (2) clearer alignment with cross-company well-being strategy.

Commented [3]: [redacted]@fb.com this is an important detail that can materially change the outcome. H1 work is well planned out, so if we want to be more PU focused in H2, we'll need to start planning in Q2. Design planning will work backwards from how PU is defined. What would it take to develop a clear enough definition of PU so enable H2 planning within IG?
Assigned to [redacted]@fb.com

Commented [4]: CWB is standing this up now from a defintional perspective and putting together the associated measurement framework, which we expect ~ within next several weeks. This will enable standing up of metrics in Q2 (pending some staffing arrangements)

Appendix

Overview of AC Metric Development Plan

1. The measurement plan consists of two metrics: NAC ground truth (survey) and pNAC (statistically estimated ground truth proxy). We expect to have a v0 version ready for validation and understand work in Q2. We will follow [HYPERLINK "https://fb.workplace.com/groups/1540209922820872/permalink/1785747918267070/" \h] to determine the appropriateness of the proxy.
   - The key criteria we'll anchor success to are:
     - Construct validity - demonstrate that operational metric correlates well with the Ground Truth metric.

        - Sensitivity - demonstrate that the metric can be moved in individual experiments and the holdout
    - Demonstrating construct validity requires re-launching the survey in experimental holdouts
2. IG was also planning to iterate to improve the measurement apparatus used for goal setting. In particular, we intend to launch post-session surveys so that we can better correlate key session events, content and context to the ground truth collected by the survey. If we continue this plan, IG can amend our total plan as follows:
    - Use v0 pNAC primarily for understand work in H1 products
    - Incorporate changes required by CWB into the next survey iteration, and base the v1 pNAC on this ground truth.
    - Follow above validation steps on v1 pNAC, incorporating additional validations as requested by CWB.

CWB Execution Recommendations and IG Responses

Risk Mitigations. If IG does move ahead with AC measurement development, CWB recommends multiple strategies to improve survey quality and minimize risks / maximize benefits. Below are the specific recommendations and planned mitigations/responses from IG:

Recommendations	Description	IG's response
Consent	Clear and explicit consent; ideally parental consent from minors. Consider implementing on Viewpoints to enable robust consent and data protection practices.	include explicit consent at the begi of the survey; discuss with policy whether to include parental conser minors. (This should likely be a cer policy that includes surveys on PU other well-being-related topics.)
Survey	Remove emotional terms from survey questions (and start with validated items).	Done
	Audit survey questions; process is TBD, but something like [HYPERLINK "https://fb.quip.com/1DrgA2DTGhmK" \h], including, but not limited to validating instrument is catching the problematic content (e.g., if a classifier exists, pull content being demoted and then see if it correlates with desired [HYPERLINK "https://fb.workplace.com/100051002047003/p	Included in [HYPERLINK "https://docs.google.com/documen 5z5y3IR2zjkdOp1zZhWzWIQlJYA cdXuIzPJc/edit" \h]

Commented [5]: Our discussions with policy partners around doing work of this kind strongly advise AGAINST doing work on such sensitive topics with minors UNLESS parental consent is obtained. The message has been: this type of work can be very valuable, but only if done well/ethically, and parental consent is a cornerstone of research with minors, especially if we're asking about health/well-being

Commented [6]: Appearance comparison is common among women of most ages, it would still be possible to measure the phenomenon among that population and extrapolate learnings.

Viewpoints (one methodology under consideration) actually doesn't allow minors to be surveyed since they require a payment method.

Commented [7]: Are you suggesting extrapolating findings from adult women to teen girls? I wouldn't suggest that, from a developmental perspective. Do you have evidence suggesting the phenomenon manifests similarly on our platforms across the lifespan? And it seems like the strategic motivation is drawn from a desire to address this phenomenon specifically among teens (also what we've committed to doing publicly).

Right, Viewpoints isn't currently enabled for minors (but I think they'd like to be!).

Commented [8]: Just sharing this as a potential approach to mitigate risks. There are certainly tradeoffs- and more work would need to be done to validate that extrapolation would be useful. (This could also be young adult women, say 18-25, to get closer in terms of development stage.) What we do see is that there is plenty of prevalence for women across ages who deal with upward appearance comparison as it pertains to images seen on IG. I'm not sure about how directly this experience translates across development stages- but if the sharp increase in botox usage is any indication it's likely that internalization of beauty ideals is something that persists across age ranges.

The approach of improving the platform to enable teens (perhaps mediated through parents) is another potential approach to mitigate risks.

	osts/432212635155468/" \h]; or include an open-ended question around why they gave their response to a survey question).	
	Seek external consulting on use of items for survey input (either via IRB and/or academic advisory board)	Continue to develop our relationships with external experts including Phillippa Diedrichs.
Strong Data Protection Practices	Document a data storage plan that includes: only sharing data with essential personnel; ensuring data is not connected to any identifiable information (name, contact information, IP address). Consider implementing on Viewpoints to enable robust consent and data protection practices.	Continue to follow existing practice storing de-identified data after 90 c
pNAC	a priori validation criteria (i.e., determine upfront how you'll determine whether the survey items validate pNAC).	Decide on specific survey items pr pAC validation; will likely rely on m surveys. pNAC is a proxy metric, s we're primarily looking for direction consistent metric movement (follov the xMVP framework).
Communication	Any research outputs should be caveated and framed appropriately, based on the robustness of the measure; findings properly contextualized.	Agreed, and will be implemented to shareouts.
Foundational work	Keep work in the foundational bucket, not to be used for goaling – more for understand-related research to help inform product strategy.	IG has not selected using this metr a goal yet, but it is under considera Continue with foundational work or Appearance Comparison to inform product strategy. Re-evaluate goal metrics as we gather more data, including sensitivity to product cha
If goaling (not currently advised by CWB)	Garner central well-being leadership approval	Yes, for deciding to goal on this r we'd invite central well-being stakeholders to the IG review.
	Incorporate guardrails around interpretation of metric, and what product(s) it should be used for and not	Yes, as part of the decision revie adopt a goal on this metric, we'd use cases and standard guardrai those use cases.

Commented [9]: Would want to see and align upon a stronger plan here.

Commented [10]: Agreed. This should be more than just a relationship. This should be a commitment to getting academic advisory about the approach overall and the specific survey items proposed

Commented [11]: You probably know this but we've (IG) been working with Dr. Dietrichs one of our primary expert partners. As has been mentioned before, she's been a key help in developing our strategy. https://fb.workplace.com/notes/776131753309236/ This was early feedback but she also was consulted during our Take A Break design and concept testing. I need to check in with our policy partner (Vaishnavi V) whom has been working with her and soliciting feedback throughout to see our current working relationship with her. I believe it was a formalized.

Commented [12]: Are there other academic advisory boards we're building out for youth or well-being that we might be able to tap into here? I do think it would be helpful to think more deeply about the relationships we have, beyond Dr. Dietrichs and figure out the specifics around how we will engage them. [redacted]@fb.com do you have a more expansive list of experts we might want to consider for something like this from your lit review?

The IG teams haven't yet used IRBs for research, to my knowledge. I think it's worth serious consideration [redacted]@fb.com and [redacted]@fb.com do you have any examples of what this has looked like for teams here that we can learn from and any benefit trade-offs to the academic advisor/panel route?

Commented [13]: I'm not sure this sufficiently addresses concerns here.

Commented [14]: If we're going to do this on-platform, I'd recommend the team speaks with the Health RL team, who've developed strong data storage practices (including use of a data custodian) for sensitive data.

Commented [15]: [redacted]@fb.com - can the set of validation concerns be more carefully laid out? [redacted]fb.com - lets separate data & privacy [redacted] metric validation please?

Commented [16]: [redacted]@fb.com I thought this row is only about data privacy/security?

Commented [17]: What kinds of criteria would be used to determine if it should be used for goaling?

	Elucidate counter-metrics (i.e., how you will determine that high pNAC content doesn't ALSO play an important other role for users)	We'd cover this in the decision re to adopt a goal on this metric.
	Explore and report on what type of content and/or experiences is being increased or decreased as well as differential impacts by sub-population.	Continue ongoing research to understand (1) what subpopulations may be disproportionately affected by product interventions and (2) what kinds of content are increased/decreased by our metric choice.
	Seek external consulting on use of items for goaling (either via IRB and/or academic advisory board)	See above on external consulting.

Commented [18]: See comment above about elucidating criteria for goaling. And, what are thoughts about what such counter-metrics might look like?

Theory of Change (High Level)

3. The MWB H1 product bets (TaB & TN) were designed to be intentionally general with an emphasis on agency & control. Appearance Comparison has always been the North Star but due to measurement & allowable intervention uncertainty the team opted to start with the more general framing.
    - The theory of change for H1 bets is centered on awareness. Although many young people can recall NAC experiences when asked about them retroactively, it's typically not top of mind for them while they're using IG. They[HYPERLINK "https://docs.google.com/presentation/d/1PHGJS6j0w-uCKT8J_u-Zz-FpeHKyTdvQ2vrGceI8cgI/edit" \l "slide=id.gd9b8e4fdae_0_288" \h][HYPERLINK "https://docs.google.com/presentation/d/1PHGJS6j0w-uCKT8J_u-Zz-FpeHKyTdvQ2vrGceI8cgI/edit" \l "slide=id.gd9b8e4fdae_0_288" \h] on external interruptions or internal conflict to spark awareness of when they're feeling bad about themselves, and it can be challenging for them to proactively identify the behaviors and experiences that lead to these feelings..
    - Reducing NAC specifically through this mechanism will likely be achieved through targeting. The initial target designs are more content agnostic, so the expected % impact on NAC in H1 may be diluted because the TaM is all active teens, not just those with high NAC potential. By helping the right targeted set to reflect on their current experiences on Instagram and how those experiences make them feel, we can empower them to make informed choices, change their behavior, and take actions that reduce NAC

Background:

1. In early December, Instagram publicly [HYPERLINK "https://about.instagram.com/blog/announcements/raising-the-standard-for-protecting-teens-and-supporting-parents-online" \h] that we're "Providing teens with more tools to better manage their Instagram experience".

   a. This is so that “people feel good about the time they spend on Instagram”, in particular for cases like when someone has been scrolling for a certain amount of time (tool: Take a Break) or dwelling on one topic for a while (tool: Topic Nudge).
   b. To continue to provide public proof points of our progress, Instagram plans to launch Topic Nudge in May and roll out Take a Break improvements to follow the Feb global launch.
2. In early December, Instagram decided that the go-forward [HYPERLINK "https://docs.google.com/presentation/d/19JpzbdOFo6PwDnRF-AGLOL-cbtyDbNGgXQ8UJSLvpjc/edit" \l "slide=id.g1048090c7e9_8_37" \h] in this area is to “reduce experiences of problematic use (PU) and negative appearance comparison (NAC), and land a steady drumbeat of expert-backed visible products that support teen well-being”.
   a. This definition of success assumes that over the long-term (2023-24), we progressively land measurement that covers:
      i. Reducing NAC
      ii. Reducing PU
      iii. Supporting teen well-being (some subset of positive outcomes)
3. In November, Instagram decided on a [HYPERLINK "https://docs.google.com/presentation/d/1vmLxiYdD_UohE5jmq06_H6SYBtzeZ5wHNw9tzd7wmaI/edit" \l "slide=id.gfc3ea09a61_0_0" \h] to complete initial measurement of Negative Appearance Comparison for specific [HYPERLINK "https://docs.google.com/presentation/d/1vmLxiYdD_UohE5jmq06_H6SYBtzeZ5wHNw9tzd7wmaI/edit" \l "slide=id.g1016f4519c4_0_0" \h]. This was informed by the Core Data Science <> IG partnership on pTRIPS and CDS’s recommendation on the [HYPERLINK "https://fb.workplace.com/notes/1025043198241838" \h] (xMVP).
   a. In this phased validation process, the “pNAC” operational metric would be used by mid-2022 for identifying new product opportunities and for evaluating impact for the long-term “reduce experiences of negative appearance comparison (NAC)” goal.
   b. If all goes well, pNAC would be used for team-level goaling in 2023.
   c. pNAC is not proposed as a metric to be used outside Instagram or to be shared externally.
4. In November, the NAC survey at the core of the phased validation process was blocked by central survey reviewers for needing additional review given Facebook Files.
5. In mid-December, Research for societal topics [HYPERLINK "https://fb.workplace.com/groups/userexperience/posts/5253454121348735" \h] to report centrally. Leaders in this new reporting structure were not present in decisions made in the prior month.