

**LESSER-REDACTED FILING**

**EXHIBIT 18**

**EXPERT REPORT OF**

**Dimitri A. Christakis, M.D., M.P.H.**

**May 16, 2025**

The undersigned hereby certifies their understanding that they owe a primary and overriding duty of candor and professional integrity to help the Court on matters within their expertise and in all submissions to, or testimony before, the Court. The undersigned further certifies that their report and opinions are not being presented for any improper purpose, such as to harass, cause unnecessary delay, or needlessly increase the cost of litigation.

A handwritten signature in black ink, appearing to read "Dimitri Christakis", written in a cursive style.

---

**Dimitri Christakis, M.D., M.P.H.**

**Highly Confidential (Competitor)**

## Table of Contents

I.	Introduction.....	3
II.	Executive Summary of Key Opinions .....	3
III.	Qualifications.....	5
IV.	Methodology.....	7
V.	Detailed Statement of Opinions.....	8
	A. Strength of the Evidence.....	8
	C. Review of Publicly Available Medical Literature .....	18
	D. Review of Defendant Documents & Research .....	23
VI.	Social Media Use and Mental Health Harms.....	28
	A. Conceptual Models from Defendant Platforms .....	29
	B. Other Foundational Concepts .....	34
	i) The Psychology of “Flow”.....	34
	ii) Active vs Passive Social Media Use.....	37
	iii) Measuring “Screen Time”.....	41
VII.	Pre-Teen and Teen Brain Development.....	44
VIII.	Addiction.....	49
	A. Validated Social Media Addiction Scales.....	52
	B. The Mechanism of Addiction .....	53
	C. The Formation of Habit—Distinguishing Addiction from Habit.....	55
	D. Problematic Social Media Use Can Occur as Habitual, Compulsive, or Addictive .....	60
	E. Social Contagion via Social Media Platforms .....	64
	F. Fear of Missing Out (FoMo).....	69
	G. Internal Documents from Defendant Platforms Recognizing Problematic and Addictive Usage Among Users .....	75
	i) Meta.....	75
	ii) TikTok.....	87
	iii) Snap.....	94
	iv) YouTube .....	98
IX.	Design Features of Social Media that Drive Usage and Addictive Behavior.....	103
	A. Likes, Comments, and other Metrics .....	105
	B. Algorithmic Recommendations .....	107
	C. Auto-Play .....	109
	D. Infinite Scroll .....	109
	E. Beauty Filters.....	110

F.	Safer Alternative Design.....	111
X.	Internal Documents Connecting Features to Harm.....	112
XI.	Social Media and Specific Harms.....	157
A.	Body Dysmorphic Disorder .....	158
B.	Eating Disorders.....	183
C.	Sleep Problems.....	192
D.	Depression and Anxiety.....	206
i)	Depression .....	208
ii)	Anxiety .....	240
E.	Suicide, Suicidal Ideation, and Self-Harm.....	244
F.	Attentional Capacity .....	269
G.	School Performance .....	278
H.	Cyberbullying and Risky Behaviors .....	281
I.	Unwanted Interactions from Adults.....	289
XII.	Selected High Profile Dissenting Studies .....	319
J.	Orben and Przybylski (2019).....	319
K.	Ferguson (2024).....	322
L.	Ferguson, Kaye, Branley-Bell, Markey (2025) .....	326
M.	Hancock et al. (2022, last revised 2025).....	332
N.	NASEM Report.....	336
XIII.	Prevention Measures.....	339
O.	At What Age Should Adolescents Use Social Media?.....	349
P.	Inadequate Age Verification.....	353
i)	Meta .....	353
ii)	TikTok.....	370
iii)	YouTube .....	374
iv)	Snap.....	379
Q.	Inadequate Parental Controls .....	381
i)	Meta .....	381
ii)	TikTok.....	384
iii)	YouTube .....	386
iv)	Snap.....	391
v)	Exogenous Parental Controls Efforts.....	395
R.	Other safety features .....	397

i)	Lack of default limits on the length and frequency of sessions.....	397
ii)	Defective opt-in restrictions to the length and frequency of sessions, FTW .....	398
iii)	Defective self-limiting tools .....	398
iv)	No blocks to usage during certain times of day.....	399
v)	Defective barriers to deactivation/deletion of accounts.....	400
XIV.	Conclusion .....	405

**I. Introduction**

1. I was asked to provide an expert assessment of what, if any, role social media (SM) usage plays in youth mental health and function. I was also asked to review the specific features of SM that adversely contribute to mental health and function. Lastly, I was asked to opine as an expert in public health and pediatrician as to the necessity to fully inform parents and their children as to the full risks associated with use of SM. For the purposes of this report, I will classify children ages 8-12 as “pre-teens,” children 13-17 as “teens,” and the 18-26 age group as “emerging adults.” As discussed below, teens and preteens are especially vulnerable to the effects of SM. My report and opinions will primarily focus on SM’s impact on teens and preteens.

2. A copy of my CV is attached (Exhibit A), Materials List (Exhibit B), Compensation Statement (Exhibit C), and Prior Testimony (Exhibit D).

**II. Executive Summary of Key Opinions**

3. Problematic social media use and addiction are disorders defined in part by the compulsive use of social media. They are well recognized in the scientific community and peer-reviewed literature.

4. Pre-teens and teens are particularly vulnerable to the problematic use of social media and its resulting negative health outcomes. Pre-teens are the most vulnerable to effects from social media, including problematic use, addiction, mental health harms, and inappropriate contact from adults.

5. Certain vulnerable adolescents are more at risk for adverse effects on their mental health and function.

6. A review of the available meta-analyses and other relevant literature establishes that social media causes or contributes to addiction, problematic usage, anxiety, depression, body dysmorphia, eating disorders, sleep deprivation, suicide, and self-injury.

7. Specific design features of Facebook, Instagram, Snap, YouTube, and TikTok work in concert to promote both the addictive nature of social media and its associated harms. Platform features and platform algorithms create and amplify mental health problems for pre-teens and teens.

8. Defendants' (Meta, Snap, TikTok, and YouTube) internal studies and documents show the harmful effects of their social media platforms, including addiction and negative mental health outcomes. The documents also reveal that the resources Defendants put towards mitigating these harmful effects were weighed against user engagement and the risk of subsequent loss of revenue. Perhaps as a consequence, Defendants only instituted minimal change prior to the initiation of this litigation.

9. Regardless of any safety changes, the ongoing research and literature shows that children and teenagers continue to be harmed by social media use.

10. In other instances, Defendants had internal data regarding potential harms and the ability to further investigate those harms, but did not do so.

11. Despite Defendants' internal data showing that their social media sites are addictive, promote problematic use, and result in an increased risk of anxiety, depression, suicidality, sleep deprivation, body dysmorphia, and eating disorders, as well as other mental

health issues, Defendants did not provide meaningful information about these harms to parents or children.

12. For parents and children to make an informed decision regarding the risks/benefits of social media, social media companies need to fully disclose the nature and risk of harms to them.

13. Social media has also changed the school environment. The same addictive design features of social media that drive user engagement result in its use during the school day. The school environment has been negatively impacted by the mental health problems social media causes in kids, and by increases in distraction and behavioral issues linked to social media use.

14. Because of the increased risk of harm to children and adolescents, in my opinion, social media platforms, as designed, are not reasonably safe for children. At a minimum, informed parental consent should be required for use of social media under the age of 16.

15. Due to the risks to children, effective age verification and parental controls are necessary.

16. Due to the risks to children, including the risk of addiction, better user controls are necessary.

### **III. Qualifications**

17. I am the George Adkins Professor of Pediatrics and an adjunct Professor in Psychiatry and in Health Services at the University of Washington. I have been studying children and media for 27 years (including social media since it was launched) and have secured millions of dollars in federal and foundation grants as a principal investigator or co-investigator. I have served as a mentor to over 15 junior faculty and post-doctoral students who also study children and media. In addition to clinical and teaching duties, I am a prolific researcher. I have published

over 275 peer reviewed scholarly articles including over 80 related to children and media. My current h Index (measure of scholarly impact) is 102 (>60 = “Exceptional”). I am the editor in chief of *JAMA Pediatrics*, the world’s leading pediatric scientific journal with an impact factor of 24.7.

18. I received the Academic Pediatric Association Research award for lifetime contribution to pediatric research. I received the Holroyd-Sherry Award from the American Academy of Pediatrics (AAP) for my outstanding contributions to research related to children and media. I was asked to give the University of Washington Distinguished Scientist Lecture (Highest Faculty Honor 2021). I served on the AAP Executive Committee on Children and Media for six years and have been the lead author of several AAP guidelines on children and media. I was a member of the National Academy of Sciences Board of Children Youth and Families for six years. I served on a National Academy of Medicine expert panel on children and media. I was a member of the 2023 National Sleep Foundation expert consensus panel on screens and sleep. I am the co-editor of a recently published 87-chapter comprehensive *Handbook on Children and Screens* (Springer 2025). I served on the Advisory Board for Children and Screens, a not-for-profit foundation with a mission to help children live healthy lives in a digital world, from 2009-2022, and have served as its Chief Science Officer since 2022. In that capacity, I review all of the current research and oversee our grants related to the foundation’s mission.

19. Currently, I serve on the DSM-VR expert panel that is seeking to add gaming addiction to the manual. Finally, as a board-certified pediatrician and Professor of Pediatrics, I have provided direct patient care to children 0-26 (including those with eating disorders, depression, anxiety, suicidality, and addiction) in both inpatient and outpatient settings. As part of

my clinical work, I have directly seen the impact social media has had on adolescents and their mental health.

#### **IV. Methodology**

20. I approached my evaluation by drawing upon my multidisciplinary expertise as a Professor of Pediatrics, Psychiatry, and Health Services, which combines both medical training and public health education. My analysis employs a systematic review of meta-analyses of existing literature, individual studies where relevant, and internal industry documents and studies done by some of the Defendants. My systematic approach evaluated the “strength of the evidence,” which aligns with clinical frameworks used in pediatric practice, while incorporating epidemiological principles from the public health field. Throughout my academic career and clinical practice, I have routinely evaluated research based on this methodology.

21. In forming my opinions regarding the potential causal relationship between social media platform use and adolescent mental health outcomes, I have relied upon my medical training, training in public health and epidemiology, my clinical experience, and my own research into media as well as an extensive review of academic literature. I have also reviewed and considered internal documents from the Defendants and depositions of current and former employees of the Defendants that were provided to me.

22. As will be discussed in detail throughout this report, even the best designed studies have limitations, and scientists love to harp on them. For one thing, doing so motivates further study which is the business scientists are in. But the truth is science is both iterative and imperfect and with each successive study more of the “truth” is revealed. In offering my “expert” opinion for this report, I am bringing to bear the totality of my knowledge of the literature, my ability to

interpret it, and my clinical experience to reach a conclusion. This is consistent with the way I (and others) practice clinical care.

23. I hold all the opinions stated in this report to a reasonable degree of scientific and medical certainty.

**V. Detailed Statement of Opinions**

**A. Strength of the Evidence**

24. Scientific discovery is an iterative process with a generally accepted hierarchy of the strength of evidence.

**Figure 1: Hierarchy of Types of Scientific Studies**



25. **Editorials, expert opinions.** These lie at the bottom of the “pyramid” as they do not “add” any new evidence but rather render opinion(s) on existing data.

26. **Case series, case reports.** These are descriptive studies ranging from one to several cases. They do not test specific hypotheses and lack a comparison (“control”) group. Accordingly, their broader implications are difficult to contextualize.

27. **Case-control studies.** These studies are designed especially for “rare” outcomes. They retrospectively compare exposures among “cases” that have a particular outcome or condition to “controls” that do not to try and identify antecedent risk factors for developing a condition of interest.

28. **Cohort studies.** These are studies of “exposed” and “unexposed” populations. They can be prospective, retrospective, or cross sectional. Cross sectional studies compare the groups at a single point in time. In general, prospective longitudinal (cohort) designs are stronger than cross-sectional ones. They begin with a population with (or without) a particular exposure and follow them through time. Retrospective cohort studies compare groups as well but look backwards rather than forwards. Cohort studies can generate (or corroborate) hypotheses and reliably estimate the prevalence of conditions in populations.

29. **Randomized Controlled Trials (RCTs).** Long considered the “gold standard” of evidence, these entail experimental manipulations of a prespecified exposure or treatment (e.g. a drug) as well as a prespecified outcome of interest. RCTs have the most robust control (comparison) group, since the randomization of the exposure ensures that any differences between the “exposed” and “unexposed” (or the “treatment” and the “control”) group are due to random chance. However, there are many situations in which RCTs are neither practical nor ethical (e.g., exposing a comparison group to a known carcinogen), and indeed a great many accepted causal relationships are based on observational studies alone. The vast majority of social media studies that are available in the general scientific literature are observational, as randomizing people to sites, platforms, or exposure time is impractical in real world settings.

30. **Systematic reviews.** Even strong study designs, including RCTs, however large or well done, are always subject to limitations of generalizability as they almost certainly focus on a particular subset of the population at a particular point in time. Accordingly, the “pinnacle” of the pyramid is occupied by “systematic reviews.” A systematic review is a scientifically reproducible search that aims to acquire and summarize all of the studies published on a particular topic within a specified date range. Once the relevant studies are collected and summarized, researchers can,

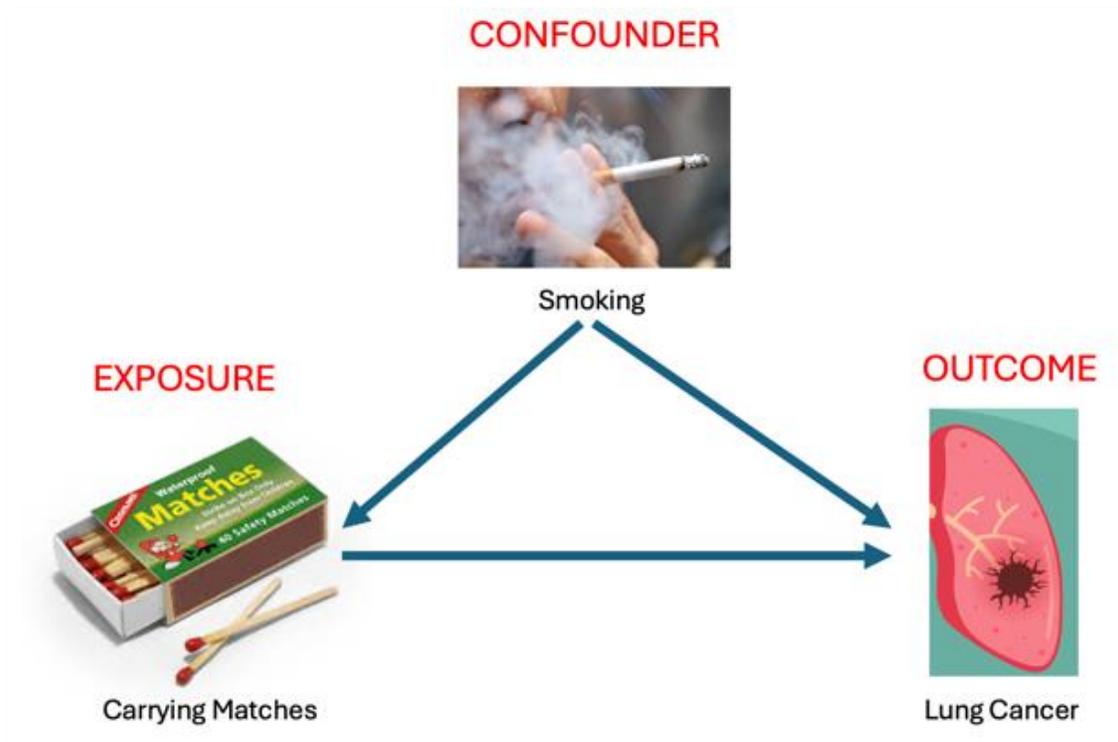
when possible, conduct a “meta-analysis” wherein the results of multiple studies are combined to establish a summary estimate of an effect.

31. Residual confounding is always a potential limitation of even the best-done observational studies. In statistical parlance, a “confounding” variable is something that is associated with both the exposure and the outcome of interest. Consider the (“true”) finding that people who carry matches are 10 times more likely to die of lung cancer than people who do not carry them. (Figure 2). While the association is real, it is “confounded.” Smokers are also 10 times more likely than non-smokers to carry matches and smoking increases the chances of developing lung cancer. (Figure 3). Indeed, once “smoking” is added as a variable to the statistical model, the relative risk of “carrying matches” leading to “lung cancer” is dramatically reduced and is no longer statistically significant.

**Figure 2: Carrying Matches and Lung Cancer**



Figure 3: Carrying Matches, Smoking, and Lung Cancer



32. Conceptually, confounding is different from mediation. While confounding creates spurious associations, mediation explains the mechanism by which an exposure leads to an outcome. Mediating variables are *in the causal pathway* between an exposure and an outcome.

33. The distinction between a confounder and a mediator is critical to understanding how to interpret results of statistical models. Adding a “confounder” to a statistical model eliminates a spurious association; adding a “mediator” to a statistical model is a way of testing whether it is potentially in the causal pathway. It explores the *mechanism* by which an exposure leads to an outcome. Table 1 below summarizes the differences between confounding and mediation.

**Table 1: Confounders and Mediators**

<b>Aspect</b>	<b>Confounding</b>	<b>Mediation</b>
Role	External distortion of relationship	Explains causal mechanism
Association	Related to both variables but not on pathway	Lies on causal pathway between variables
Effect on Analysis	Biases the estimate of causal effect	Decomposes effect into direct/indirect paths
Goal	Remove or control for it	Understand the causal mechanism

34. When it comes to interpreting effect sizes, larger effect sizes from observational studies are generally thought to provide stronger evidence for a given association being causal because they make it less likely that a residual confounding variable might explain the observed association. Nevertheless, there are several reasons why small effect sizes still have serious implications.

35. **First**, while the probability of an exposure leading to an outcome might be low, it is crucial to remember that for the individuals affected should the outcome occur, the consequences can profoundly affect their life. That basic fact has been acknowledged by at least one of the platforms at issue in this litigation. In an interview with the *Wall Street Journal* responding to Meta documents leaked by whistleblower Frances Haugen, Adam Mosseri, CEO of Instagram is quoted as saying, “In no way do I mean to diminish these issues,” including teen body image issues among others. “Some of the issues mentioned in the story aren’t necessarily widespread, but their impact on people may be huge.”<sup>1</sup>

36. **Second**, even very small effect sizes can represent the truth and can have serious implications when applied at a population level. For example, in 2023, *JAMA* found that exposure to air pollution for more than one year increased the chances of developing cardiovascular disease

---

<sup>1</sup> Alison Lee Deposition Exhibit 7 at 4

over one's lifetime by 8%.<sup>2</sup> Although the relative risk is numerically “small,” it results in a sizeable number of additional cases of heart disease (800,000 of them).

**Figure 4: JAMA - Relative Risk - Air Pollution Exposure**

**Relative Risk (RR):** Long-term exposure to fine particulate matter (PM2.5) increases the risk of heart disease by 8% (RR = 1.08).

**Baseline Risk:** Suppose the baseline risk of heart disease in a population is 10% (i.e., 10% of adults develop heart disease over their lifetime without exposure to PM2.5).

**Population Size:** Imagine a population of 100 million adults.

**Risk Without PM2.5 Exposure:**  $10\% \times 100 \text{ million} = 10 \text{ million cases}$ .

**Risk With PM2.5 Exposure:**

- Relative risk increases to  $10\% \times 1.08 = 10.8\%$ .
- Cases =  $10.8\% \times 100 \text{ million} = 10.8 \text{ million cases}$ .

**Additional Cases:**  $10.8 \text{ million} - 10 \text{ million} = 800,000 \text{ additional cases}$ .

37. Similarly, at his deposition, Mark Zuckerberg (the CEO and founder of Meta) acknowledged he was sent an email from another Meta executive indicating that the company had a “deep understanding” that the prevalence of severe problematic use among Facebook users was 3.1%.<sup>3</sup> Mr. Zuckerberg acknowledged the obvious, which is that “3 percent of billions of people is a lot of people....It’s not – not the majority, but it’s – obviously, it’s millions of people.”<sup>4</sup>

38. *Third*, effect sizes—however small—can be contextualized by comparing them to other known effect sizes that are deemed “significantly impactful.” For example, mammography

---

<sup>2</sup> Alexeeff SE, Deosaransingh K, Van Den Eeden S, Schwartz J, Liao NS, Sidney S. Association of Long-term Exposure to Particulate Air Pollution With Cardiovascular Events in California. *JAMA Network Open*. 2023;6(2):e230561-e230561. doi:10.1001/jamanetworkopen.2023.0561

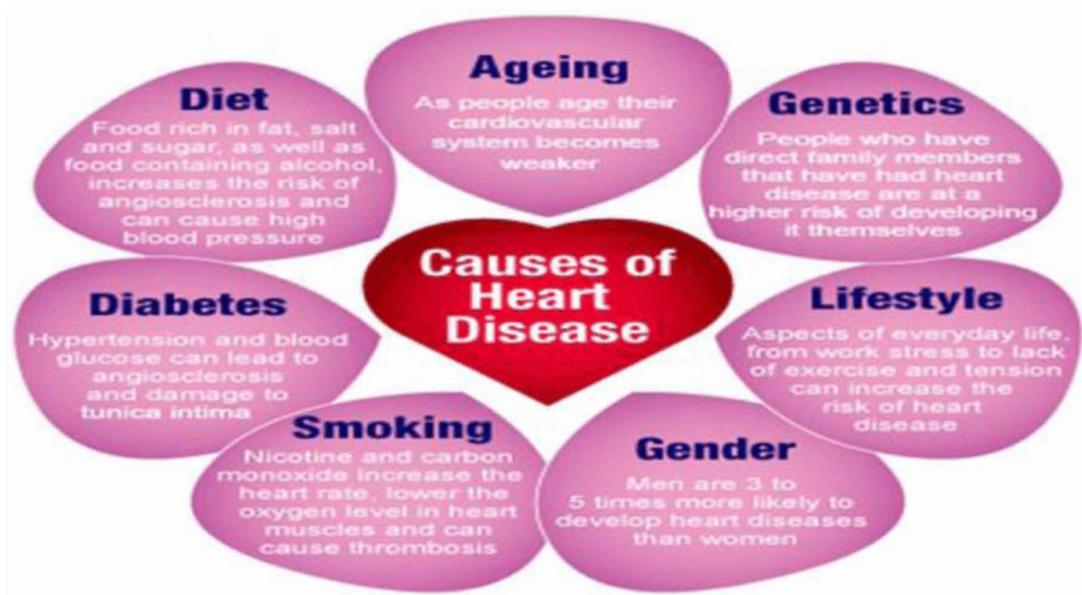
<sup>3</sup> Mark Zuckerberg Dep. Exhibit 39 at -0761; Mark Zuckerberg Deposition Transcript at 262:5-266:13.

<sup>4</sup> Mark Zuckerberg Deposition Transcript at 275:2-6

screening for women 40-75 years of age decreases the risk of death from breast cancer by 15% in relative terms and .05% in absolute terms yet it is routinely done.<sup>5</sup>

39. **In addition, although** some health problems have a singular cause, many do not. For example, Sickle Cell Anemia is the result of a single gene mutation affecting hemoglobin formation. But most health problems are multifactorial in that they have several contributory causes and risk factors. One such condition is heart disease, the leading cause of death in the US. The figure below illustrates these causes and the roles they play.

**Figure 5: Causes of Heart Diseases from Kadam et. al.<sup>6</sup>**



40. Changes in any of these would affect the probability of developing heart disease.

Unfortunately, some of them are immutable. For example, gender and genetics are fixed at birth

<sup>5</sup> Force UPST. Screening for Breast Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2024;331(22):1918-1930. doi:10.1001/jama.2024.5534

<sup>6</sup> Kadam, Kalyani & Kamat, Pooja & Malav, Amita. (2019). Cardiovascular Disease Prediction Using Data Mining Techniques: A Proposed Framework Using Big Data Approach. 10.4018/978-1-5225-7338-8.ch007.

and ageing is inevitable. Others are potentially modifiable (e.g. smoking, lifestyle, or diet). In fact, a foundational tenant of Public Health is focusing resources on modifiable risks as a means to improve the health of a population, and although the more risk factors reduced the better, changes in any would be expected to result in benefits. In other words, an obese, diabetic, elderly male would still reduce their risk of heart disease by quitting smoking

41. Each of these independent risk factors required years of research to isolate and were incorporated into the model over time. Similarly, the outcomes discussed in this report are influenced by a complex interplay of genetic, biological, psychological, and environmental factors—social media platforms are one additional contributor. While this report details how and why social media plays a role, that does not alter the fact there are multiple potential causal influences for the harms discussed. At the same time, the existence of known risk factors does not rule out the identification of new ones, just as occurred in the evolving understanding of heart disease over decades. When considering the role that Social Media plays in problematic usage, sleep, depression, anxiety, eating disorders, suicide and self-harm, I also considered the alternative explanations including COVID-19, socio-economic status, family stress, exposure to violence, Adverse Childhood experiences etc. In the best designed studies, those alternatives are explicitly accounted for either as covariates in the model or because of the design of the study itself. Just as smoking is not the *only* cause of heart disease but it is still a risk factor for it, social media is not the only cause of the harms detailed in this report, but it is an important contributor to them, and for some children, it is the determinative factor.

#### **B. Applying Population Estimates to Individual Cases and Generalizability of Studies**

42. Whether observational or experimental, research studies frequently compare differences between the means of two samples, measure the effect size of that difference, and then

perform a test of statistical significance to see if that difference has a less than 5% chance of being due to chance alone. These differences are measured at the population level—an exposed one compared to an unexposed one. Population estimates are statistical tools that provide insights into the characteristics of groups within a population, such as age distribution, income levels, or health variables. While these estimates are invaluable for understanding trends and making decisions at the community or national level, applying them to individuals requires caution because population estimates represent averages or generalizations that may not accurately reflect the unique circumstances of a single given person. For instance, knowing that a certain percentage of a population has a specific health condition (even if it's 99%) doesn't *ensure* that a particular individual drawn from that population has it. Or knowing if someone is exposed to something doesn't necessarily pre-determine their outcome. Therefore, while population estimates can inform policies and programs, individual assessments should rely on personal data and direct evaluations to ensure accuracy and relevance.

43. Put another way, epidemiologic studies using population estimates answer a question writ large: does smoking increase the chances of developing lung cancer? Or does it cause *some* cases of it? They do not tell us if smoking *caused* lung cancer in a particular patient. For example, given two people with the exact same smoking history, one might develop lung cancer and the other may not; and of course, some people who never smoke also get lung cancer. Seen in this light, the associations discussed in this report must ultimately be applied to the individual attributes and the context of each individual plaintiff.

44. **Generalizability.** Consider two substances which are commonly accepted to be harmful to health: alcohol and cigarettes. Significant studies have been conducted establishing that exposure to alcohol and cigarettes can increase the risk of certain cancers, heart disease, and

premature death. As such, the U.S. Surgeon General and other public health officials caution Americans against consumption of these substances. There is no need for the warnings to be specific to the various types of alcohol or brands of cigarettes. The underlying mechanisms that lead to these harms are the same regardless of small differences between each product (e.g. “light” cigarettes were no safer than regular ones it turns out). In much the same way, given the similarities between the salient features of social media platforms and the mechanisms by which they lead to harms discussed below, it is reasonable and appropriate to extrapolate findings from one site to another. In epidemiologic terms this is called the principle of generalizability, which allows scientists and clinicians to make meaningful policy and treatment recommendations without requiring time consuming, expensive, and unnecessary studies. Whether one compulsively watches TikTok, Instagram, or YouTube, the effects are analogous. The truth is, as competitive as SM landscape is, the internal documents clearly evidence that they all seized on any feature that others deployed effectively to increase engagement and time on the platform and emulated it.

### **C. Review of Publicly Available Medical Literature**

45. Given the multitude of studies that have been performed relevant to the questions of interest, I prioritized systematic reviews and/or metaanalyses of each of the possible pathways. My searches primarily focused on publications from the past 5 years, as these studies are inclusive of the most recent scientific data and include metaanalyses and citations to foundational studies and scientific research from earlier years. Where appropriate, I have searched and reviewed relevant studies published earlier. Table 2 below summarizes my search terms.

**Table 2: Search Terms and Approach Deployed**

<b>Outcome</b>	<b>Search Terms</b>
Addiction	<i>("Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking") and ("addiction" or "problematic use" or "habitual use") and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>
Body Image/Eating Disorder	<i>("Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking") and ("body image" or "body dysmorphic disorder" or "eating disorder" or "body dissatisfaction" or "anorexia" or "bulimia" or "disordered eating") and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>
Sleep	<i>("Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking") and ("sleep" or "insomnia") and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>
Depression	<i>("Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking") and ("depression" or "depressive symptoms") and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>
Anxiety	<i>("Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking") and "anxiety" and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>
Suicide	<i>"Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking" and ("suicide" or "suicidal ideation" or "self-harm") and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>
School Performance	<i>"Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking" and ("school" or "school performance") and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>
Risky Behaviors	<i>("Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking") and ("risk taking" OR "risk*behave*" OR sex* OR smoke* OR substance use OR aggress* OR alcohol OR viol*) and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>
Cyberbullying	<i>("Social Media" OR "Facebook" OR "Instagram" OR "snapchat" OR "TikTok" OR "YouTube" OR "social networking") and ("cyberbullying" or "students") and "systematic review" Since 2020; include citations; (exclude patents). Sort by date.</i>

46. It should be noted that most of the extant literature uses symptoms rather than clinical diagnosis as outcomes. There are several reasons for this. First, symptoms can be self-

reported using existing validated screening instruments for depression, anxiety, eating disorders etc. Clinical diagnoses require structured interviews by trained experts which are cumbersome and expensive—in many cases prohibitively so. This is why national surveillance data bases such as the Youth Risk Behavior Survey conducted by the CDC asks about symptoms and those symptoms formed the basis for the Surgeon General’s Report on the Mental Health of Teens.<sup>7</sup> For example, the fact that in 2021, 57% of females reported “persistent feelings of sadness in the past year” and 30% report “seriously considered suicide” is alarming. No serious scientist should doubt we have a mental health crisis in the US based on those data.

47. Second, there are also ethical issues related to making formal psychiatric diagnoses especially if the study does not have the means to provide ongoing mental health care and medication etc. Clinical diagnoses are also impractical as many research subjects might not agree to get an in-person assessment by a professional leading to missing data and a non-generalizable sample.

48. Third, symptom level data treats mental health as a continuum rather than a dichotomy which is more consistent with how it is experienced. A person who develops depression isn’t “symptom free” one day and “depressed” the next but rather they accumulate and maintain symptoms over some period of time eventually crossing a clinical threshold. In fact, consistent with the FDA recommendation that industry collect patient reported outcomes, it is common for studies of anti-depressant drugs to measure changes in *symptoms* rather than change in diagnosis

---

<sup>7</sup> <https://www.hhs.gov/sites/default/files/surgeon-general-youth-mental-health-advisory.pdf>

as measures of a drug's effectiveness (with some studies determining the minimal clinically meaningful change in symptoms).<sup>8</sup>

49. Finally, as important and impactful as clinically diagnosed mental health conditions are, many are sufficiently rare that studying them would require enormous sample sizes. For example, death by suicide happens in approximately 9 out of 100,000 teenagers in the United States. To have adequate statistical power to detect a threefold increased risk would require 175,000 children, something that would be virtually impossible to do without industry collaboration. Accordingly, suicidal ideation or self-harm, both of which are more common and are correlated with suicide, are frequently used.

50. I then used the ROBIS approach (summarized below) to assess the completeness and quality of all reviews and for selecting which ones to include.<sup>9</sup>

---

<sup>8</sup> Team TFAWDS. Fluoxetine, Cognitive-Behavioral Therapy, and Their Combination for Adolescents With Depression Treatment for Adolescents With Depression Study (TADS) Randomized Controlled Trial. *JAMA*. 2004;292(7):807-820. doi:10.1001/jama.292.7.807; Turkoz I, Alphas L, Singh J, et al. Clinically meaningful changes on depressive symptom measures and patient-reported outcomes in patients with treatment-resistant depression. *Acta Psychiatrica Scandinavica*. 2021;143(3):253-263. doi:<https://doi.org/10.1111/acps.13260>; Hudgens S, Floden L, Blackowicz M, et al. Meaningful Change in Depression Symptoms Assessed with the Patient Health Questionnaire (PHQ-9) and Montgomery-Åsberg Depression Rating Scale (MADRS) Among Patients with Treatment Resistant Depression in Two, Randomized, Double-blind, Active-controlled Trials of Esketamine Nasal Spray Combined With a New Oral Antidepressant. *Journal of Affective Disorders*. 2021/02/15/ 2021;281:767-775. doi:<https://doi.org/10.1016/j.jad.2020.11.066>; Food U, ADMINISTRATION D. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. *Fed Regist*. 2009;74(35):65132-65133; Whiting P, Savović J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. Jan 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005.

<sup>9</sup> Whiting P, Savović J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. Jan 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005

**Figure 6: ROBIS Criteria for Evaluating Systematic Reviews**

	Phase 2				Phase 3
	1. Study eligibility criteria	2. Identification and selection of studies	3. Data collection and study appraisal	4. Synthesis and findings	Risk of bias in the review
Signaling questions	<p>1.1 Did the review adhere to predefined objectives and eligibility criteria?</p> <p>1.2 Were the eligibility criteria appropriate for the review question?</p> <p>1.3 Were eligibility criteria unambiguous?</p> <p>1.4 Were all restrictions in eligibility criteria based on study characteristics appropriate?</p> <p>1.5 Were any restrictions in eligibility criteria based on sources of information appropriate?</p>	<p>2.1 Did the search include an appropriate range of databases/ electronic sources for published and unpublished reports?</p> <p>2.2 Were methods additional to database searching used to identify relevant reports?</p> <p>2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?</p> <p>2.4 Were restrictions based on date, publication format, or language appropriate?</p> <p>2.5 Were efforts made to minimize error in selection of studies?</p>	<p>3.1. Were efforts made to minimize error in data collection?</p> <p>3.2. Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?</p> <p>3.3. Were all relevant study results collected for use in the synthesis?</p> <p>3.4. Was risk of bias (or methodologic quality) formally assessed using appropriate criteria?</p> <p>3.5. Were efforts made to minimize error in risk of bias assessment?</p>	<p>4.1. Did the synthesis include all studies that it should?</p> <p>4.2. Were all predefined analyses reported or departures explained?</p> <p>4.3. Was the synthesis appropriate given the nature and similarity in the research questions, study designs, and outcomes across included studies?</p> <p>4.4. Was between-study variation minimal or addressed in the synthesis?</p> <p>4.5. Were the findings robust, for example, as demonstrated through funnel plot or sensitivity analyses?</p> <p>4.6. Were biases in primary studies minimal or addressed in the synthesis?</p>	<p>A. Did the interpretation of findings address all of the concerns identified in domains 1 to 4?</p> <p>B. Was the relevance of identified studies to the review's research question appropriately considered?</p> <p>C. Did the reviewers avoid emphasizing results on the basis of their statistical significance?</p>
Judgment	Concerns regarding specification of study eligibility criteria	Concerns regarding methods used to identify and/or select studies	Concerns regarding methods used to collect data and appraise studies	Concerns regarding the synthesis	Risk of bias in the review

51. All included studies conformed to Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines. In cases where there were multiple reviews within the same time period, I reviewed them for consistency and completeness. If they were comparable, I chose the more recent one (if it included more studies). If they were inconsistent, I assessed the quality of each and determined which one better represented robust scientific findings.

A relevant example of this is the Ferguson meta-analysis of SM and mental health outcomes which contained multiple fundamental methodological flaws discussed in detail in section XI.B below.<sup>10</sup>

52. In addition to the five-year lookback discussed above, my review and knowledge base also includes published papers and studies from the inception of social media, as I have spent 25 years researching and publishing in this area. In addition, I reviewed the reference lists of selected studies, searched for RCTs published that were not included in the meta-analyses, and used the Web of Science to find studies that cited the studies I used. As discussed below, I also considered some other individual, high-profile studies.

#### **D. Review of Defendant Documents & Research**

53. I also reviewed internal Defendant documents to assess what the companies' internal research demonstrated about the role their products play in the development and/or exacerbation of mental health harms. Additional materials I considered are listed on my materials list that is attached to this report. These materials were made available to me based upon searches and topics that I requested.

54. The table below summarizes a sampling frame of a few company studies that I came across while reviewing documents. In most of these studies, survey data were linked to actual usage statistics and user experiences on the platform.

---

<sup>10</sup> Ferguson CJ. Do social media experiments prove a link with mental health: A methodological and meta-analytic review. *Psychology of Popular Media*. 2024:No Pagination Specified-No Pagination Specified. doi:10.1037/ppm0000541

**Table 3: Select Internal Studies from Defendants**

<b>Sample Size</b>	<b>Country</b>	<b>Reference</b>
22,410 Instagram users	International	META3047MDL-003-0009501
~30,000 adult Facebook users ~30,000 youth Facebook users ~30,000 adult Instagram users ~30,000 youth Instagram users	International	Zuckerberg Dep. Ex. 31
20,000 US Facebook users	US	METATNAG-010-00000060
6,000 US Facebook users	US	META3047MDL-019-00033466
7,471 Instagram; 37,729 FB	International	META3047MDL-020-00588061
~15,000 FB users	International	META3047MDL-020-00588282
100,000 FB cross-sectional; 15,000 longitudinal	International	METAMAAG-011-00000381
50,590 FB users	International	META3047MDL-014-00401897
19,275 Teen Facebook and Instagram users	International	META3047MDL-004-00003256
2,503 Instagram users	International	Bhutada Depo Exhibit 7; slide 3
6,793 Instagram users	International	Bhutada Depo Exhibit 11; slide 8
3,155 Meta users	Unclear	META3047MDL-044-00171351
689,003 Facebook users	International	Kramer et al <sup>2</sup>
1,000,000 TikTok users	Unclear	TIKTOK3047MDL-047-LARK-00510819
238,000 Instagram Users	Unclear	Arturo Bejar Deposition (p. 241)

55. Speaking as a scientist in this research field, the size, scope, and granularity of these data are extraordinary. They are among the largest samples of social media users ever assembled for research purposes. Compared to any independent scientist or the medical community at large, the Defendants had ready access to precise and granular data as well as the ability to deploy robust studies to better understand and mitigate the risks of their platforms. For example, the investigators had the ability to link survey data to actual site usage (something an independent researcher cannot readily do).

56. Publicly, Meta assured parents and physicians that they hoped to share their research through peer-reviewed publications.<sup>11</sup> However, only a small number of these studies (or versions of them) were made available to the medical community, and very few were published in the peer-reviewed literature. Perhaps even more concerning, I found no evidence that Meta’s internal findings and assertions that its products posed risks to children and emerging adults were communicated to parents or children who used their product, whether on the platform itself, the company website, or elsewhere.

57. Further, my review indicates that—to the extent Meta’s research was made public—there were sometimes meaningful differences between what the company found and what it published. As noted above, Mr. Zuckerberg was informed that Meta had a “deep understanding” that 3.1% of Facebook users experienced *severe* problematic use. In the same communication, he was informed that 55% of Facebook users experienced *mild* problematic use.<sup>12</sup> This indicates that, according to Meta’s own research, collectively 58.1% of its Facebook users experience some form of problematic use. Weeks after this email was sent, Meta researchers published a paper concerning problematic use as part of a conference held in May 2019 (the CHI Conference on Human Factors in Computing Systems Proceedings).<sup>13</sup> That paper, titled “Understanding Perceptions of Problematic Facebook Use,” states that, “we estimate (as an upper bound) that 3.1% of Facebook users in the US experience problematic use.”<sup>14</sup> Given that a 3.1% “upper bound” is significantly lower than 58.1%, the published research significantly understates the prevalence and scope of problematic use as known to Meta.<sup>15</sup> Minimizing the problem was a company strategy. As Bejar

---

<sup>11</sup> META3047MDL-020-00253760, -3762

<sup>12</sup> Mark Zuckerberg Deposition Exhibit 39 at -0761

<sup>13</sup> Mark Zuckerberg Deposition Exhibit 89

<sup>14</sup> Mark Zuckerberg Deposition Exhibit 89 at p. 2.

<sup>15</sup> Mark Zuckerberg Deposition Transcript at 682:3-684:8

states in his deposition “you weren’t supposed to use the term addiction. That instead they labeled it “problematic use.” And they had defined problematic use to be very narrow.”<sup>16</sup> Meanwhile, Mr. Zuckerberg ironically acknowledged at his deposition, that published research is “not very helpful if it’s not accurate.”<sup>17</sup> In general, as Arturo Bejar says in his deposition, “I believe that growth and engagement are top priority and safety is an afterthought. (p51. Lines 20-21)

58. My comparison of internal company research and publicly available studies indicates that there are also sometimes meaningful differences between purely internal studies and those that are developed in consultation and collaboration with outside scientists. Perhaps the most notable example of research by a social media company conducted in collaboration with academia is a recently completed study that attempted to assess the introduction of Facebook on well-being using a global sample.<sup>18</sup> The authors (Vuorre and Przybylski) acquired daily (DAU) and monthly active users (MAU) from Facebook and regressed those on to Gallup World Poll (GWP) data from 72 countries. GWP data are collected annually on 1000 noninstitutionalized civilians ages 15 and older per country. For this study, positive emotions included affirmative responses to “did you feel well-rested yesterday?”, “were you treated with respect all day yesterday?,” “did you learn or do something interesting yesterday?,” and “did you smile or laugh a lot yesterday?”.<sup>19</sup> The negative emotions included affirmative responses to “did you experience the following during a lot of the day yesterday: physical pain, worry, sadness, stress and anger?”<sup>20</sup> In brief, they report finding no

---

<sup>16</sup> Arturo Bejar Deposition Transcript at 136:16-21

<sup>17</sup> Mark Zuckerberg Deposition Transcript at 681:8-9

<sup>18</sup> Vuorre M, Przybylski AK. Estimating the association between Facebook adoption and well-being in 72 countries. *R Soc Open Sci.* 2023;10:221451.

<sup>19</sup> Vuorre M, Przybylski AK. Estimating the association between Facebook adoption and well-being in 72 countries. *R Soc Open Sci.* 2023;10:221451.

<sup>20</sup> Vuorre M, Przybylski AK. Estimating the association between Facebook adoption and well-being in 72 countries. *R Soc Open Sci.* 2023;10:221451.

significant association between the rise of Facebook usage and wellbeing changes in the countries studied.

59. But there are multiple limitations to this study. First, the data are ecological, meaning that Facebook usage was not linked to survey respondents. It is not even clear if the respondents used Facebook at all. Second, the age of the sample ranged from 15 years and older. Although the paper does not provide exact numbers, they aggregated data into two strata—15-34 years and 35 plus. Hence, the actual number of teens in their sample is small and not reported separately, making it impossible to discern if they were directly affected. Third, the outcomes measured, while they may have face validity for positive and negative emotions, are not consistent with the constructs believed to be associated with excessive social media usage. For example, no one has ventured a hypothesis that social media use leads to physical pain for a lot of the day.

60. Furthermore, while this might seem like a *prima facie* example of Meta collaborating with an independent scientist, that collaboration is not entirely independent. Przybylski discloses that he has served as an “unpaid” advisor to Facebook in the past, and he is an outspoken critic when it comes to social media’s effects on teenagers. Finally, the article states “that the data are not publicly available, the study was not pre-registered, and researchers can contact Facebook if they wish to reproduce the analyses.” These facts are not consistent with transparent open science.

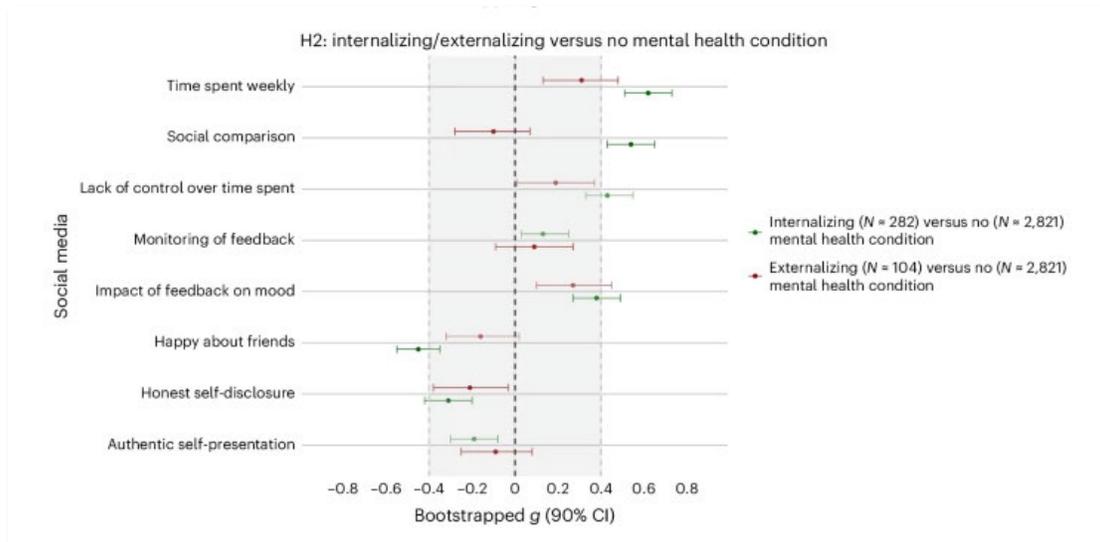
61. My report reviews and synthesizes the existing scientific evidence as well as the industry’s own findings made available through discovery, for each of the eight main outcomes that result from social media exposure in children and emerging adults—Problematic Social Media Use, Body Image, Eating Disorders, Sleep Disorders, Depression, Anxiety, Suicide and self-harm, and School Performance.

## VI. Social Media Use and Mental Health Harms

62. Pre-teens and teens are particularly vulnerable to problematic use of social media and the resulting negative health outcomes. Social media causes or contributes to causing mental health harms such as addiction, problematic usage, anxiety, depression, body dysmorphia, eating disorders, poor sleep, suicide, and self-injury. The paths by which media use in general, and social media use in particular, are related to these mental health outcomes are complex and inter-related.

63. It is recognized the pre-teens and teenagers with pre-existing mental health conditions are particularly vulnerable to harm caused by social media. In a large cross-sectional study of 3340 UK teenagers aged 11-19, Fassi et. al. assessed time spent on social media based on a clinical diagnosis of an underlying mental health condition.<sup>21</sup> The results are below:

**Figure 7: Differences in Social Media Use For H2 Group from Fassi, et. al.<sup>22</sup>**



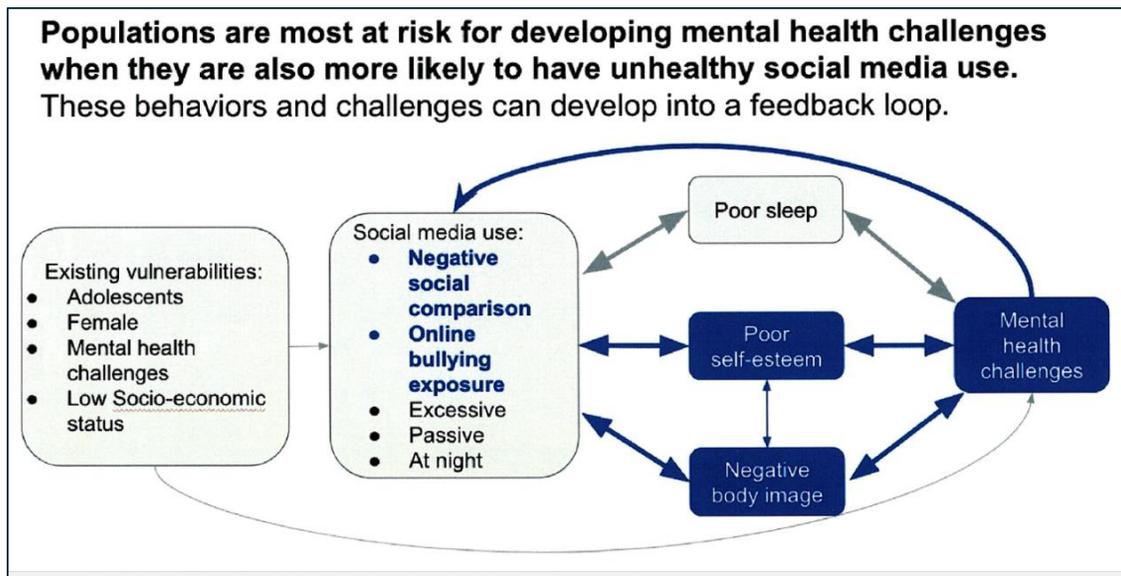
<sup>21</sup> Fassi L, Ferguson AM, Przybylski AK, Ford TJ, Orben A. Social media use in adolescents with and without mental health conditions. *Nature Human Behaviour*. 2025/05/05 2025;doi:10.1038/s41562-025-02134-4

<sup>22</sup> Fassi L, Ferguson AM, Przybylski AK, Ford TJ, Orben A. Social media use in adolescents with and without mental health conditions. *Nature Human Behaviour*. 2025/05/05 2025;doi:10.1038/s41562-025-02134-4

The authors had pre-specified the effect size that they deemed to be clinically significant at +/- .4. This is an arbitrary and debatable threshold. Nevertheless, those with internalizing conditions (anxiety and depression) spend statistically (and clinically per the authors) more time on social media and are more likely to be subjected to social comparison. Many of the other effects are statistically different from null and, in my opinion, still clinically relevant (recall the potential relevance of small effect sizes (Section XX). These results highlight the importance of differential susceptibilities and demonstrate how some populations may be at increased risk of SM effects.

### A. Conceptual Models from Defendant Platforms

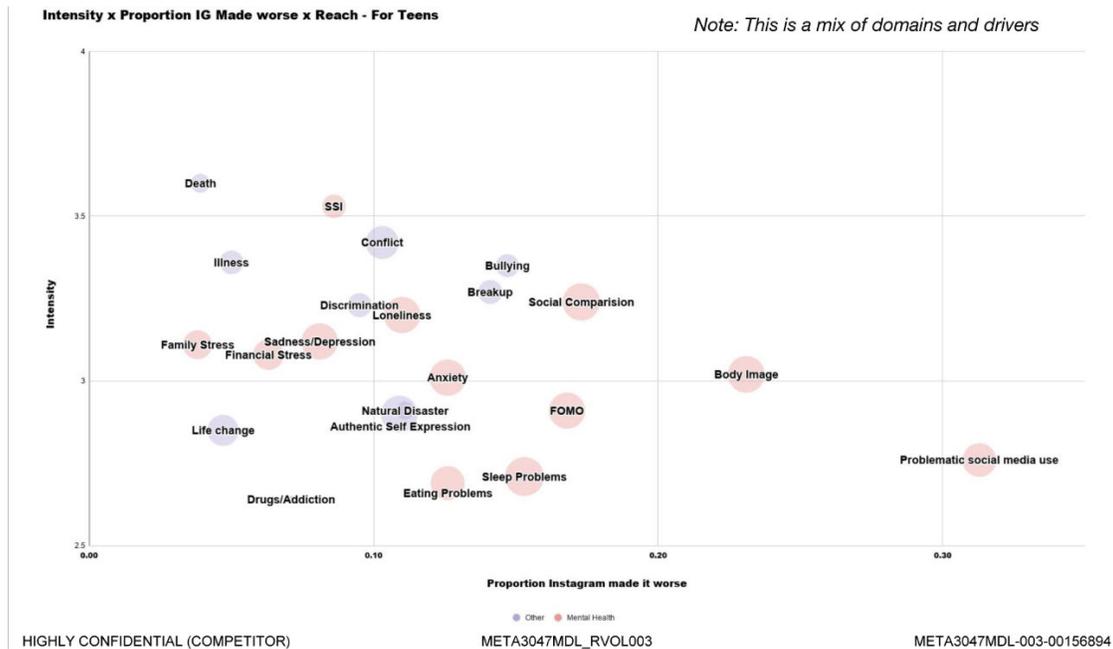
64. Interestingly, Meta has a model illustrating the negative harms caused by social media and acknowledges adolescents and people with mental health challenges as especially vulnerable groups.



*Document 1: Deposition of Diego Castaneda, Exhibit 26 at 3*

65. The existence of this Meta logic model suggests that at least some of its scientists, (Dr. Castaneda was a leader in the Instagram well-being team) are both cognizant of and

conceptually grounded in the current scientific literature. In fact, an internal Meta presentation contained the following slide depicting certain harms that Instagram “made [] worse” for teens:

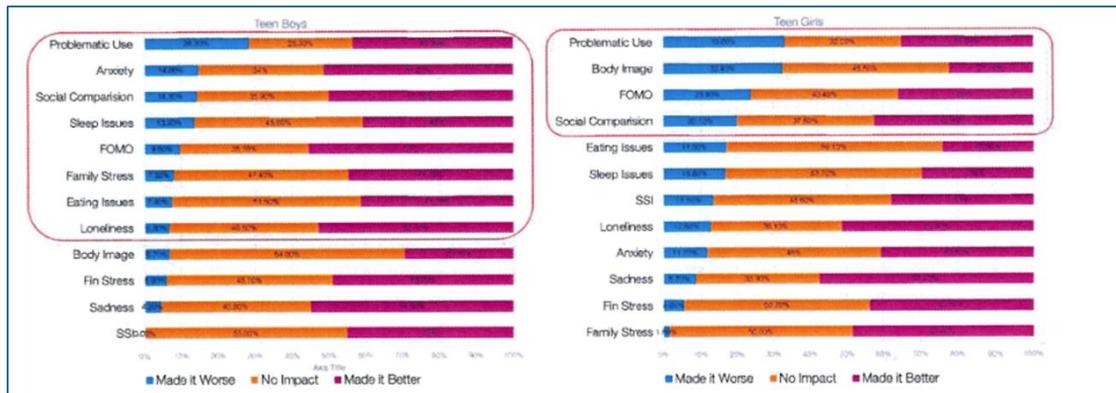


Document 2: META3047MDL-003-00156888, -6894

66. The pink bubbles highlight some of the mental health harms I address in this report. Based on my review of similar documents, it is apparent that the size of the circles corresponds to the “reach” of these issues.<sup>23</sup> The figure indicates Instagram itself acknowledges that it worsens the mental health areas noted.

67. The self-perceived role that social media plays in many mental health outcomes was assessed by Meta in a survey of over 22,410 Instagram users across the United States, Japan, Brazil, Indonesia, Turkey and India. Amongst teens, the results were as follows:

<sup>23</sup> META3047MDL-003-00094811, -4828.

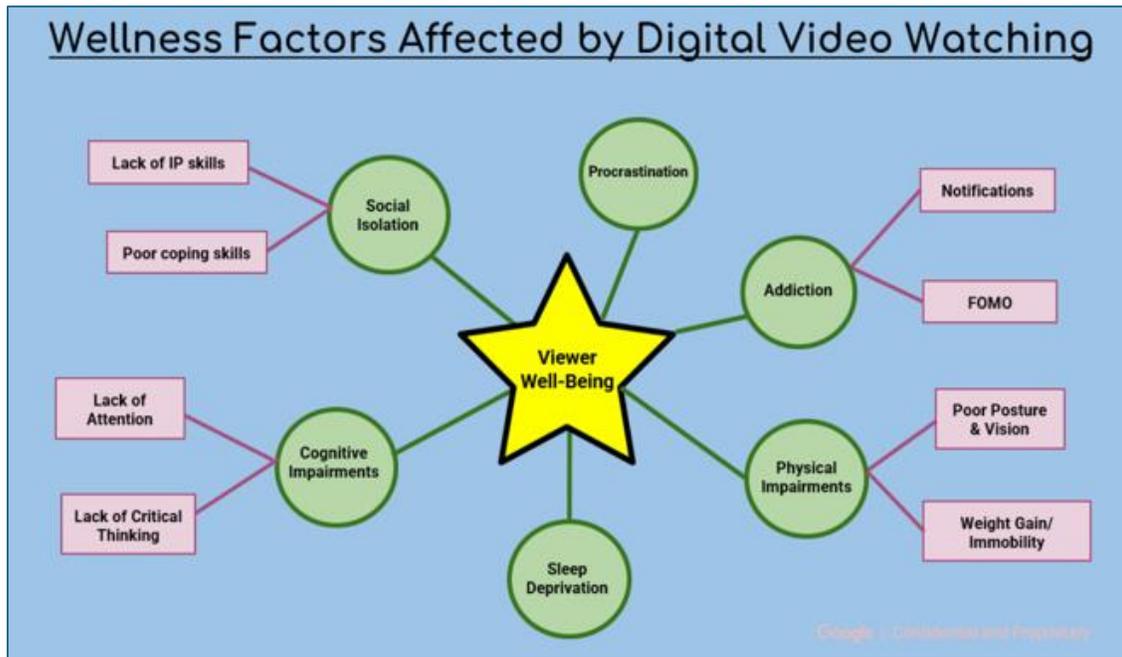


Document 3: Deposition of Alison Lee, PhD, Exhibit 10 at p. 14

68. Although a sizeable percentage of respondents felt that Instagram ameliorated certain outcomes, a sizeable percentage also felt it made them worse—and in the case of body image for teen girls, that represented over 1/3 of respondents.<sup>24</sup> Consistent with differential susceptibilities, the effect of social media is not uniform across all children, and a subset of them are especially vulnerable to it while others may derive benefit. This is why the net effect on a population may be attenuated whereas at an individual level there can be significant harms. Although, these data are subjective self-reports, the questions have face validity and other studies have used validated measures of affect and mood and are addressed in my report.

69. YouTube also has a conceptual framework with some antecedent pathways:

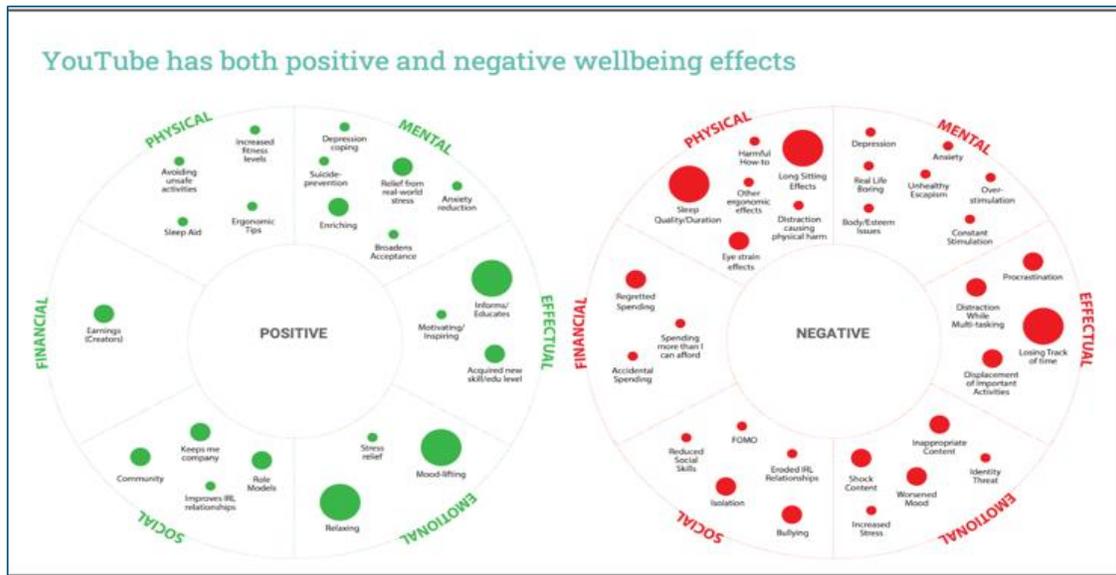
<sup>24</sup> See also META3047MDL-003-00094811 at -4828: “1 in 3 teen girls blame Instagram for making their body image issues and problematic social media use worse.”



Document 4: GOOG-3047MDL-04918852 at Slide 5

Their framework includes pathways to social isolation, addiction, sleep deprivation, and cognitive impairments, among other “wellness factors”.

70. I also reviewed a YouTube framework, below, for what it sees as its “positive” and “negative” effects of a variety of domains. My interpretation is that the size of the circles corresponds with the effect size. Again, many of the “negative well-being effects” they ascribe to YouTube are discussed in this report.



Document 5: GOOG-3047MDL-00898168 at Slide 7

71. Lastly, while I have not located a similar visual diagram for TikTok or Snapchat, I would note that internal TikTok documents reflect the knowledge that “[c]hildren that are most vulnerable off-line are typically those who are most vulnerable online.”<sup>25</sup> Similarly, internal TikTok presentations note that, “Whatever our background, we all inherently understand that children are more vulnerable than adults and that we as adults have a responsibility to keep them safe.”<sup>26</sup> Despite that recognition, made clear to them by outside experts, I found no evidence that TikTok warned its teen users, or their parents, that they were at increased risk of harm.

72. In Snap’s case, the lack of conceptual models seems to be the result of willful blindness. Communications intended for Snap’s Board of Directors acknowledged that “Ages 13 to 17 years are a large Snap demographic and, given their age, are also a vulnerable population. We believe they require a heightened standard of care.” Evan Spiegel, Snap’s CEO, testified that

<sup>25</sup> TIKTOK3047MDL-018-00361108, -1108

<sup>26</sup> TIKTOK3047MDL-044-00859648 at Slide 5

he personally agreed with these statements and that Snap has a moral responsibility to children who use Snapchat.<sup>27</sup> However, Morgan Hammerstrom, Snap's Director of Product Research, testified that she had never been asked to research a user's experience in app as it relates to their mental health.<sup>28</sup> She also testified specifically that she had never researched whether or not users find Snapchat to be addictive, and that such information "wouldn't have anything to do with my job or the role that I have at Snapchat."<sup>29</sup> Similarly, Lauryl Schraedly, Snap's former Global Head of Consumer Insights, testified that the Consumer Insights team was never asked to assess the impact of Snapchat on users' mental health or whether Snapchat was creating addictive behavior in its users.<sup>30</sup>

**B. Other Foundational Concepts**

73. Before delving into each of the relevant outcomes, a few more methodological and psychological constructs relevant to social media effects are worthy of review.

**i) The Psychology of "Flow"**

74. Humans bring innate and acquired skills to the challenges they face. When skills are high and challenges are low, the task at hand is sufficiently easy that it can induce boredom. Conversely, when challenges are high and skills are low, the task is sufficiently hard that it can induce anxiety. The psychologist Mihaly Csikszentmihalyi introduced the concept of "flow" as that mental state where the challenges and skills are sufficiently balanced that the experience is engrossing, engagement is easy, enjoyment is high, and time passes effortlessly.<sup>31</sup> Flow states

---

<sup>27</sup> Evan Spiegel Rough Dep. Tr. at 17:21-20:1

<sup>28</sup> Morgan Hammerstrom Dep. Tr. at 96:22-97:13, 606:3-608:7

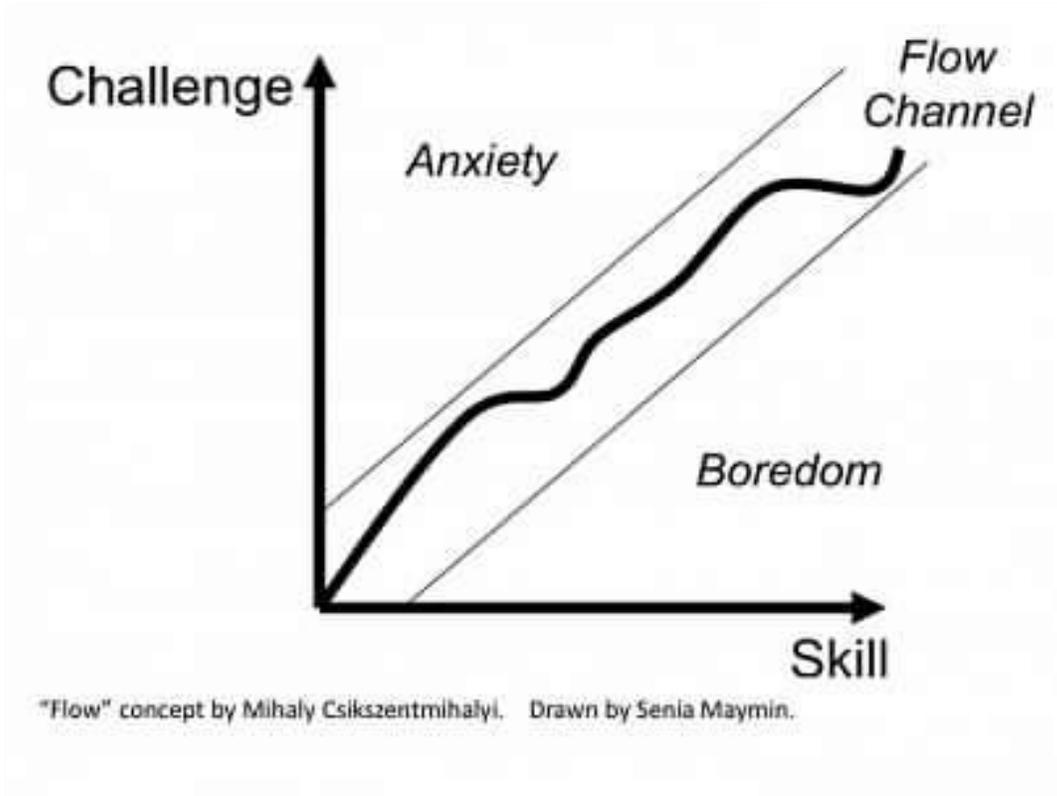
<sup>29</sup> Morgan Hammerstrom Dep. Tr. 148:9-19, 151:3-9.

<sup>30</sup> Lauryl Schraedly Dep. Tr. at 280:18-281:12.

<sup>31</sup> Csikszentmihalyi M. Flow : the psychology of optimal experience. Simon & Schuster; 1994.

result in dopamine release (discussed in Section VII below) and are inherently pleasurable. The flow channel is demonstrated in the figure below.

**Figure 8: Psychology of Optimal Experience<sup>32</sup>**



75. In the “real world” flow can be achieved by such things as well-timed promotions at work so that one feels deployed at the limit—but not beyond—their skill set, or in gaming contexts such as chess clubs by finding players that are worthy opponents. Both of those examples take time and effort and maybe even luck to achieve. Job promotions are rarely “perfectly timed” if they happen at all, and finding the “right” chess player can be challenging. With effort, some people can find flow in art, sports, music, and even work. Although Csikszentmihalyi maintained that “flow” was the key to a happy and fulfilling life, he cautioned that is not inherently or

<sup>32</sup>Maymin, S., *Flow*, THIS EMOTIONAL LIFE (May 24, 2023), available at <https://thisemotionallife.org/blogs/flow/>

universally a “good” thing. He argued it could be misused in business and war and that “mountaineers” and “gamblers” could become obsessed with it to the point of neglecting other aspects of their lives.

76. Structured by algorithms as opposed to real world constraints, the online experience can create flow instantaneously and effortlessly and maintain it indefinitely. Consider the simple game of “Candy Crush,” which launched in 2012 and continues to command the attention of hundreds of millions of players.<sup>33</sup> It requires no skill to play for the first time and so anxiety is low. In fact, reading the rules is unnecessary. The game is intuitive and there is no barrier—no friction—to beginning. Once a player begins playing, the game quickly and seamlessly ratchets up its difficulty in accordance with a player’s skill and engagement: never so hard that it becomes frustrating, never so easy that it becomes boring. Two players starting at the same time play different games, but each is likely to find theirs enjoyable. Keeping people in a flow state is an engagement strategy that many social media sites actively deploy.

77. People in a state of flow are, by definition, deeply engaged in the experience and less mindful of outside distractions or perturbations: it is an “escape.” This explains in part why people with underlying mental health conditions, or particularly disturbing realities (e.g. challenging socioeconomic circumstances), are more susceptible to the allure of a flow state and at greater risk for becoming addicted to what provides it to them—whether this is alcohol, drugs, gambling, or social media. To that end, Qin and colleagues specifically investigated if flow (which

---

<sup>33</sup> Keza MacDonald, *Crushing it: Why millions of people still can't stop playing Candy Crush*, The Guardian (Aug. 1, 2024), <https://www.theguardian.com/games/article/2024/aug/01/crushing-it-why-millions-of-people-still-cant-stop-playing-candy-crush>.

they parameterized as enjoyment, concentration and time distortion) was associated with problematic TikTok use and found a strong correlation.<sup>34</sup>

## ii) Active vs Passive Social Media Use

78. The experience of being on social media can be provisionally and conceptually divided into “active” and “passive” use. Active use entails posting or interacting with content while passive use entails viewing or scrolling. These distinctions are artificial and arbitrary: active users also view content and many people engage in both types of uses in single sessions, making operationalizing usage patterns problematic. Furthermore, they are potentially confounded (see section IV.A) since people who are depressed (or just down or tired at a given moment) might be more passive just as people who are manic (or just joyous or energetic at a given moment) might be more active. Nevertheless, these “distinct” patterns of usage found their way into the scientific literature in part as a means to potentially explain heterogenous findings relating usage to outcomes. For example, might it be that small overall effect sizes or “positive” vs. “negative” studies could be further elucidated by studying the *ways* in which people were using social media?

79. The theoretical basis for this hypothesis was that active users might be garnering social support through their interactions whereas passive users might be more likely to be engaging in social comparisons.<sup>35</sup> This theory implicitly discounts the possibility that active interactions might also be problematic—cyberbullying or asking about one’s appearance is interactive.

---

<sup>34</sup> Qin Y, Musetti A, Omar B. Flow Experience Is a Key Factor in the Likelihood of Adolescents' Problematic TikTok Use: The Moderating Role of Active Parental Mediation. *Int J Environ Res Public Health*. Jan 23 2023;20(3)doi:10.3390/ijerph20032089; Qin Y, Omar B, Musetti A. The addiction behavior of short-form video app TikTok: The information quality and system quality perspective. *Front Psychol*. 2022;13:932805. doi:10.3389/fpsyg.2022.932805.

<sup>35</sup> Godard R, Holtzman S. Are active and passive social media use related to mental health, wellbeing, and social support outcomes? A meta-analysis of 141 studies. *Journal of Computer-Mediated Communication*. 2024;29(1)doi:10.1093/jcmc/zmad055

Similarly, passive scrolling might involve watching neutral, salubrious, or social comparative content.

80. Some individual studies examined usage type and found significant mediation effects (recall that mediation analyses test possible mechanisms, see section IV.A); others did not. These inconsistent results are precisely why systematic reviews and meta-analyses were invented: as a means to synthesize and reconcile different studies with varied results (section IV.A). A comprehensive meta-analysis of 141 studies that examined passive vs. active use was performed in 2024.<sup>36</sup> Collectively, the studies yielded 897 effect sizes (508 active and 134 passive) drawn from over 145,000 participants.

81. The analyses in the studies included some that were between subjects and some that were within subjects. As the names suggest, between subject comparisons involve looking at the effects on one type of user versus another. The problem with this comparison is that it is potentially confounded: underlying differences in users' mental health might be associated both with their type of use, and with its effects. Within subject comparisons, on the other hand, compare the same individuals' usage pattern and their outcomes at different time points and thus explicitly control for differences in individuals. The people are the same; only their usage and affect changes.

82. The results of the within subject analyses are presented in Figure 7. In all cases except wellbeing, the vast majority of studies found negligible effects based on type of usage. The authors conclude "All within subjects effects tested in this meta-analysis were negligible."<sup>37</sup>

---

<sup>36</sup> Godard R, Holtzman S. Are active and passive social media use related to mental health, wellbeing, and social support outcomes? A meta-analysis of 141 studies. *Journal of Computer-Mediated Communication*. 2024;29(1)doi:10.1093/jcmc/zmad055

<sup>37</sup> Godard R, Holtzman S. Are active and passive social media use related to mental health, wellbeing, and social support outcomes? A meta-analysis of 141 studies. *Journal of Computer-Mediated Communication*. 2024;29(1)doi:10.1093/jcmc/zmad055

**Figure 9: Results of Within Subject Analyses – Active vs. Passive Use**

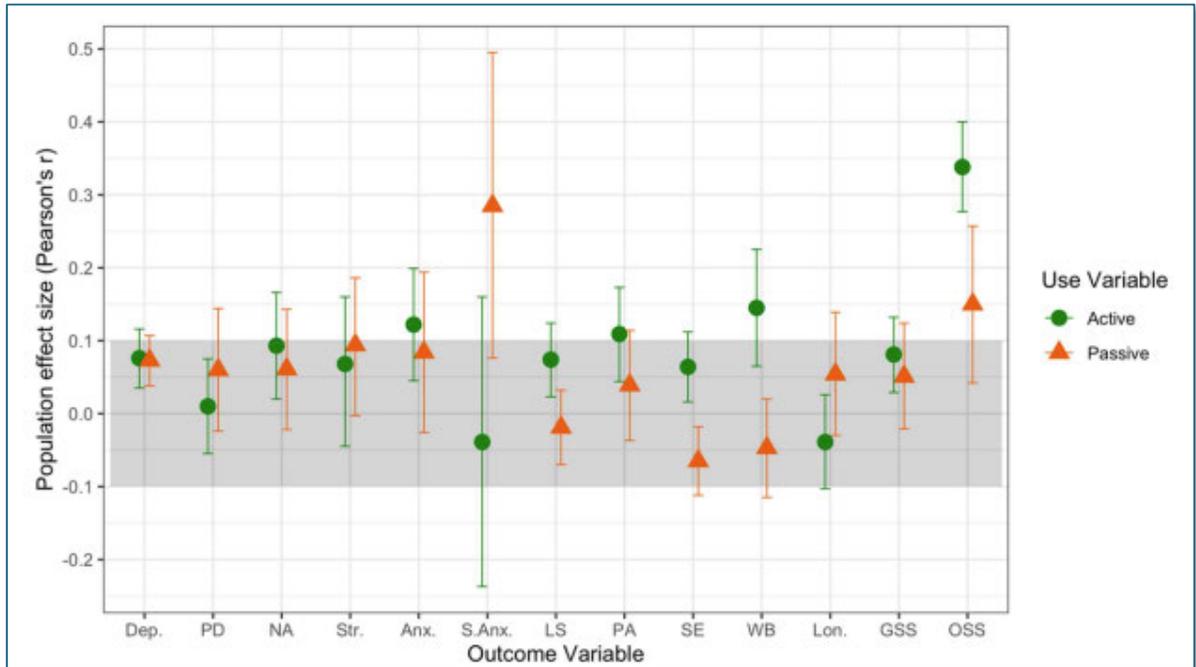
Variables	Associations with wellbeing outcomes		
	<i>Worse</i> (%)	<i>Negligible</i> (%)	<i>Better</i> (%)
<b>Active use</b>			
Illbeing	9 (29)	21 (68)	1 (3)
Wellbeing	1 (6)	12 (71)	4 (24)
Social wellbeing	0 (0)	5 (83)	1 (17)
<b>Passive use</b>			
Illbeing	2 (11)	15 (83)	1 (6)
Wellbeing	10 (43)	11 (48)	2 (9)
Social wellbeing	0 (0)	6 (86)	1 (14)

Note: Worse effects defined as  $r \leq -.10$  for positive outcomes (e.g., life satisfaction) and  $r \geq .10$  for negative outcomes (e.g., depressive symptoms). Negligible effects defined as  $-.10 < r < .10$ . Better effects defined as  $r \geq .10$  for positive outcomes (e.g., life satisfaction) and  $r \leq -.10$  for negative outcomes (e.g., depressive symptoms).

83. Next, the authors summarized the effects across 13 distinct psychological outcomes.

Those results are presented in Figure 10 below.

**Figure 10: Effects of Active vs Passive Use Across Psychological Outcomes**



Dep (depressive symptoms); PD (psychological distress); NA (negative affect); Str (stress); Anx (anxiety); S. Anx (social anxiety); LS (life satisfaction); PA (positive affect); SE (self-esteem); WB (wellbeing); Lon (loneliness); GSS (global social support); OSS (online social support). Gray areas indicate negligible effect sizes.

84. As can be seen, most of the effects are in the gray, negligible effect size range and not statistically different from each other. And in all but two of them—wellbeing and online social support—active vs. passive use are statistically indistinguishable. Let’s discuss each in turn. The difference in effect size for wellbeing is .15 (small). As discussed above, and noted by the authors of the paper, that minimal difference could be accounted for by better wellbeing *causing* more active usage rather than the other way around. The difference in online social support is slightly larger (.20). However, this too is not surprising since part of online social support entails having exchanges (active use) with other individuals, making the findings somewhat tautologic. Having reviewed the paper, I concur with the authors’ conclusions that “the mostly negligible associations

between active and passive social media use and mental health and wellbeing highlight that the public must remain cautious of overly simplistic or enthusiastic statements about the benefits of active or the harms of passive social media use.<sup>38, 39</sup>

### iii) Measuring “Screen Time”

85. For independent scientists, robust estimates of the time teens spend on social media are difficult to attain. Teen self-reporting of social media use is generally accepted as not fully accurate. Without industry collaboration, the most precise estimates of how children spend time on their phones are derived from data acquired through passive sensing—seamlessly and invisibly measuring what sites are visited, what apps are used, and for how long they are deployed during the course of a given day. Common Sense Media did such a study in 2022.<sup>40</sup> They installed a passive sensing technology on the Android phones of 203 children ages 11-17. The results are summarized in the following figures.

---

<sup>38</sup> Mr. Zuckerberg has made such statements on numerous occasions. For instance, on January 11, 2018 he posted on Facebook: “The research shows that when we use social media to connect with people we care about, it can be good for our well-being. We can feel more connected and less lonely, and that correlates with long term measures of happiness and health. On the other hand, passively reading articles or watching videos -- even if they're entertaining or informative - - may not be as good.”

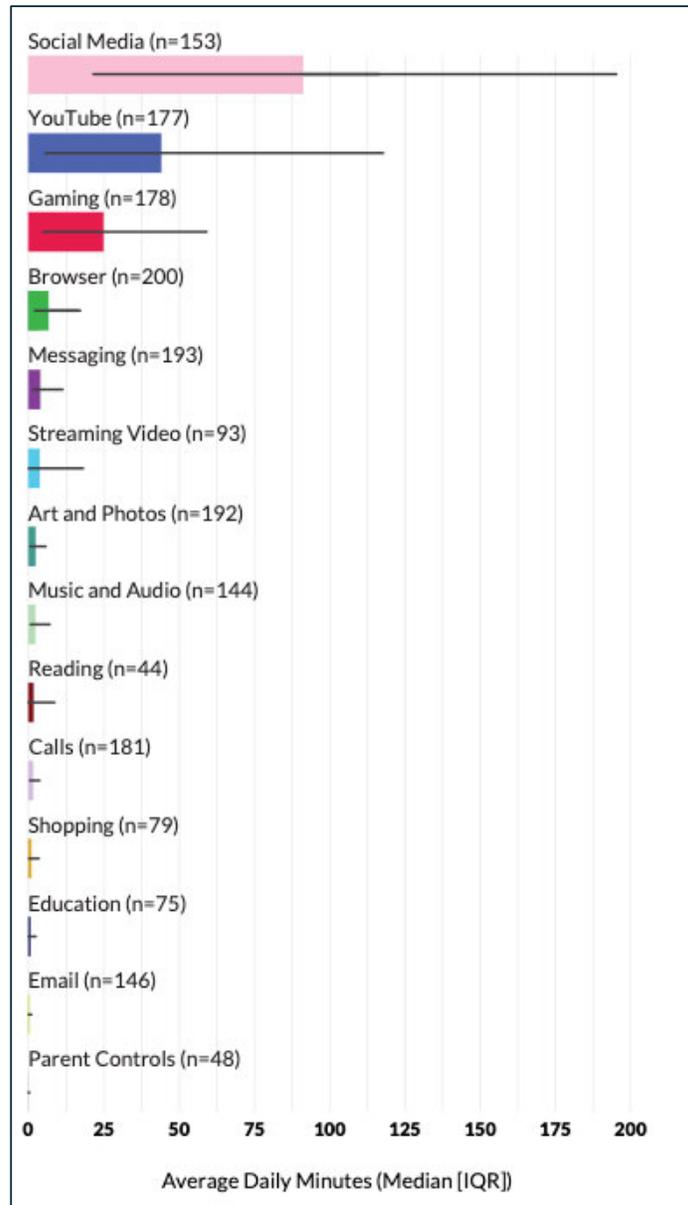
[https://www.facebook.com/zuck/posts/10104413015393571?ref=embed\\_post](https://www.facebook.com/zuck/posts/10104413015393571?ref=embed_post).

Likewise, in testimony before Congress, he said: “What we find in general is that if you're using social media in order to build relationships, right? So you're — you're sharing content with friends, you're interacting, then that is associated with all of the long-term measures of well-being that you'd intuitively think of. Long-term health, long-term happiness, long-term feeling connected, feeling less lonely. But if you're using the Internet and social media primarily to just passively consume content, and you're not engaging with other people, then it doesn't have those positive effects and it could be negative.” Transcript of Mark Zuckerberg’s Senate hearing (Apr. 10, 2018), Washington Post, <https://www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing/>.

<sup>39</sup> Godard R, Holtzman S. Are active and passive social media use related to mental health, wellbeing, and social support outcomes? A meta-analysis of 141 studies. *Journal of Computer-Mediated Communication*. 2024;29(1)doi:10.1093/jcmc/zmad055;

<sup>40</sup> Radesky J, Weeks HM, Schaller A, Robb M, Mann S, Lenhart, Constant Companion: A Week in the Life of a Young Person's Smartphone Use , Common Sense. 2023;

Figure 11: Median and IQR of different app usage



**Figure 12: Popular apps, average daily duration, and % of total smartphone usage on a typical day**

App name	N (%) users	Average daily duration Median [IQR]**	Range (hour:minutes)	Percentage of daily use (median)***
TikTok	102 (50.2%)	1:52 [0:24 - 2:57]	<0:01 - 7:48	38.4%
YouTube	175 (86.2%)	0:40 [0:05 - 1:52]	<0:01 - 10:13	18.2%
Instagram	70 (34.5%)	0:16 [0:03 - 0:52]	<0:01 - 2:56	5.9%
Snapchat	79 (38.9%)	0:10 [0:02 - 0:36]	<0:01 - 3:13	3.6%
Discord	72 (35.5%)	0:07 [0:02 - 0:24]	<0:01 - 12:20	2.5%
Roblox	74 (36.5%)	0:06 [0:01 - 0:40]	<0:01 - 6:25	2.6%
Chrome	191 (94.1%)	0:04 [0:01 - 0:13]	<0:01 - 1:24	1.5%
Netflix	53 (26.1%)	0:03 [0:01 - 0:17]	<0:01 - 7:31	0.8%
Spotify	81 (39.9%)	0:01 [<0:01 - 0:04]	<0:01 - 0:31	0.6%
Facebook	40 (19.7%)	0:01 [<0:01 - 0:04]	<0:01 - 1:34	0.1%
Google quick search box	180 (88.7%)	0:01 [<0:01 - 0:03]	<0:01 - 0:21	0.6%
Amazon	47 (23.2%)	0:01 [<0:01 - 0:03]	<0:01 - 0:20	0.3%
Pinterest	36 (17.7%)	0:01 [<0:01 - 0:03]	<0:01 - 0:48	0.4%

\*Calculated only among participants who used that app.  
\*\*Median is the value that 50% of the users are under and 50% are over. IQR is the Interquartile Range, which is the middle 50% of users, with 25% of users under the first value and 25% of users over the second value.  
\*\*\*Percentage of daily use is calculated among those who use the app and as a percentage of all their smartphone use in a day.

86. Social media (TikTok, Instagram, Facebook, Snapchat, and YouTube) are the most used apps. Over 60% of the median total daily media time is spent on TikTok, YouTube, Instagram, Snapchat, and Facebook. Given that, it is not unreasonable to posit that even studies that focus on the effects of *overall* “screen time” are driven in large part by usage of these apps.

87. My group has also collected more recent data (2024) using passive sensing in a nationally representative sample of 229 13–18-year-old US children. This study included both iOS and android phones. These results are summarized below.

**Figure 13: Android and iOS App usage in US Teens 13-18 years of age<sup>41</sup>**

<b>Label</b>	<b># users</b>	<b>Min</b>	<b>25th%</b>	<b>Mdn</b>	<b>75th%</b>	<b>Max</b>	<b>M</b>	<b>SD</b>	<b>% daily use (Mdn)</b>
TikTok (min)	173	0.08	8.62	70.10	140.31	323.16	85.10	84.62	22.9%
Facebook (min)	170	0.01	0.47	1.92	5.77	210.25	8.27	22.16	0.6%
Instagram (min)	209	0.01	4.71	15.78	53.54	347.88	38.03	54.46	5.2%
YouTube (min)	201	0.00	1.85	13.48	40.83	452.46	46.50	83.39	4.4%
Snapchat (min)	154	0.00	0.85	3.89	15.28	539.53	19.38	58.42	1.3%

88. Given the relatively small sample sizes of both studies, the two-year difference in time that the data were collected, and the fact that one used both Android and iOS whereas the other was limited to Android alone, the inconsistencies in the data are not surprising. The median and interquartile range of TikTok use for example in the CSM sample is 1:52 [24-2:57] whereas in ours it is 1:10 [8.62-2:20]. TikTok’s own data reports a median of 1:20 min per day in children 13-17 years of age.<sup>42</sup> I presume that TikTok’s data adequately represents the truth given the source and the sample size, but their data are limited to their platform and most children are on more than one. All totaled, a sizeable amount of time is spent by the “average” teenager on SM sites and more than enough to profoundly influence mental health and behavior.

## **VII. Pre-Teen and Teen Brain Development**

89. Pre-teens and teens are particularly vulnerable to mental health harms from social media use due to their biological and psychological development. Children are not simply small adults, and the impact of any experience or exposure must be understood in the context of their

<sup>41</sup><https://jamanetwork.com/journals/jamapediatrics/fullarticle/2829879#:~:text=We%20found%20that%20adolescents%20spent,on%20their%20phone%20during%20school>.

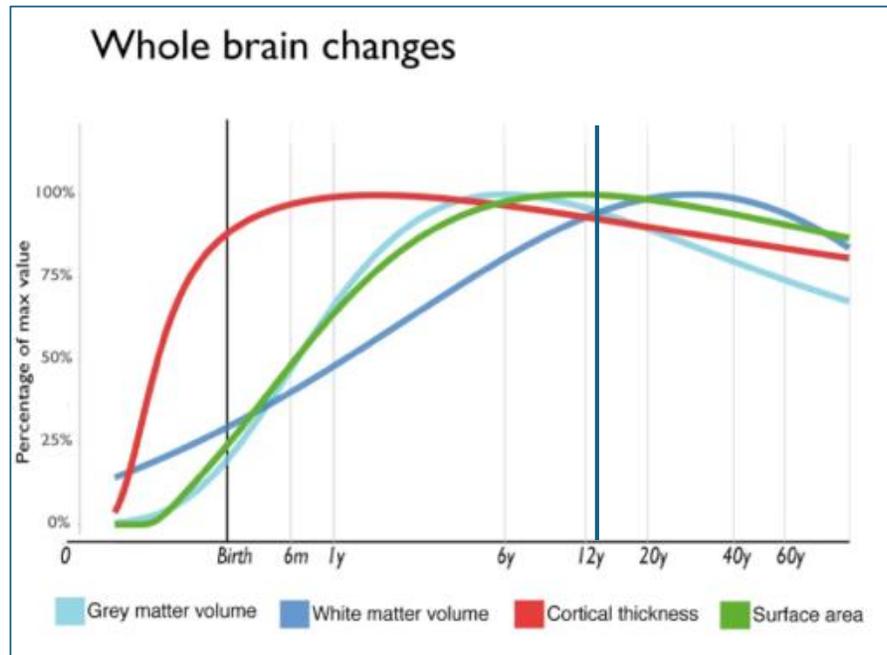
<sup>42</sup> TIKTOK3047MDL-002-00098058, -8060

ongoing biological and psychological development. Because of their brain development, they are particularly vulnerable to experiences such as “FoMo” and social contagion, both of which are discussed later in this report.

90. Human brains develop throughout adolescence into early adulthood. Different regions mature at different rates. At birth, the brainstem and cerebellum are highly developed, supporting vital functions such as breathing, heart rate, and basic motor control. During infancy and early childhood, the limbic system, particularly the amygdala and hippocampus, rapidly develops, facilitating emotional responses and memory formation. Later, the cerebral cortex, responsible for higher cognitive functions, undergoes significant growth, with sensory and motor areas maturing first, followed by language centers, which develop rapidly in early childhood. The prefrontal cortex, often referred to as the “CEO of the brain” because it is essential for decision-making, impulse control, and complex reasoning, is the last to fully mature, and typically completes development at around age 26. During adolescence, synaptic pruning strengthens important neural connections while eliminating less commonly used ones, refining cognitive abilities. Myelination, the process of insulating nerve fibers to improve communication between brain regions, progresses throughout childhood and adolescence, with the prefrontal cortex again being the last region to complete the process. This prolonged development explains why teens and emerging adults may struggle with long-term planning and impulse control compared to fully mature adults, and why social media can have significant impacts on children and adolescents.

91. The figure below from the University of Cambridge shows when different brain regions reach 100% capacity. White matter volume (shown in dark blue) is the part of the brain that plays a crucial role in memory, attention, and decision making. I have added a blue vertical line corresponding to age 13 the current “minimal” age for SM usage set by industry.

Figure 14: Brain Maturation<sup>43</sup>



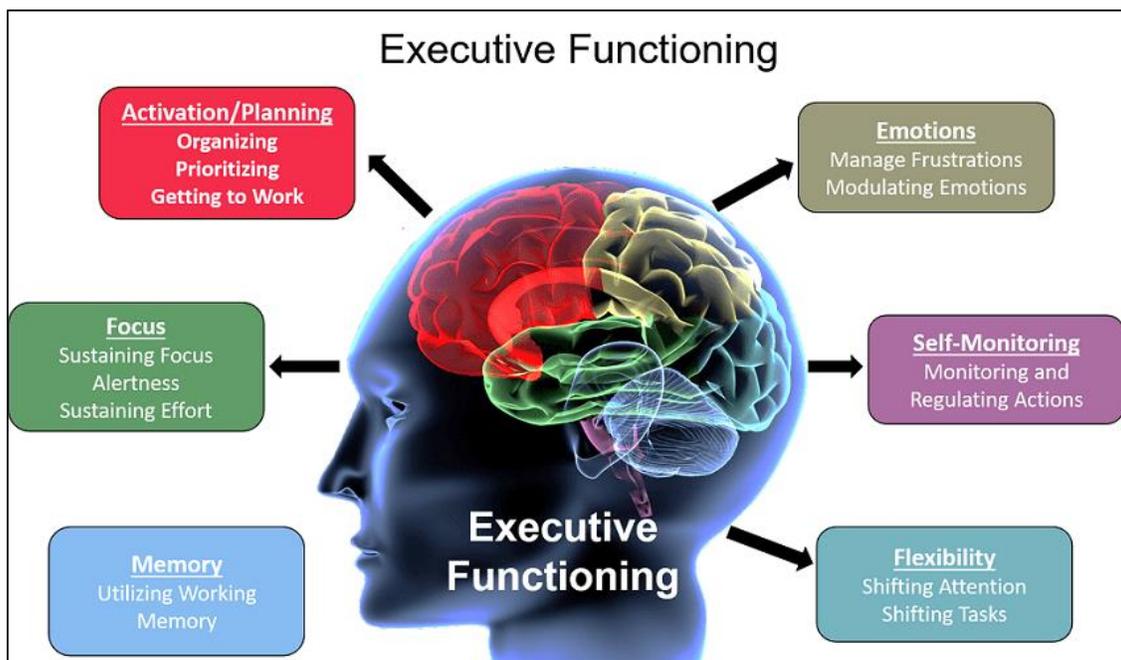
92. “Executive Function” is a foundational construct used by cognitive psychology and neuroscience to describe a set of skills that emerge as the brain develops. It is comprised of several key capacities, as shown in Figure 13 **PRIV** . All these play essential roles in human development and function and underpin both reactions to and effects of environmental stimuli.

93. Children’s developmental trajectories are highly individualized, which is to say there is considerable variability in the age at which these capacities are fully present. When we say that the “typical” 5-year-old can do X it means that as many as ½ to 1/3 cannot do it.... yet. Further, cognitive capacity is modulated by inhibitory control or “self-monitoring” (as shown in Figure 13). What this means is that even if a child “knows” the right thing to choose or to do, their lack of impulse control might make them get it wrong or not act appropriately. Optimal executive

<sup>43</sup> <https://www.cam.ac.uk/stories/BrainCharts>

function involves a titration of latency (delay) and accuracy, and executive function is essential to sound decision making. This lack of fully developed executive functioning makes adolescents particularly vulnerable to the harmful effects of social media. As noted below, this vulnerability is acknowledged in Defendants' internal documents.

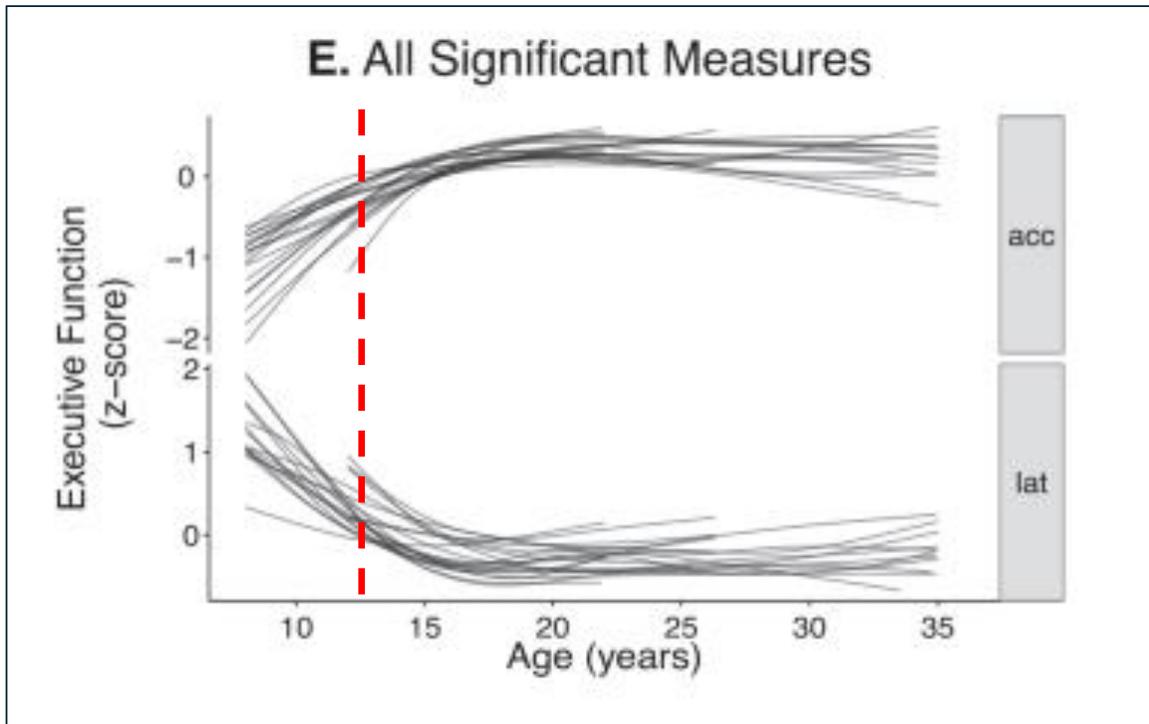
**Figure 15: Executive Functioning**



94. A recent paper used multiple population-based data sets to examine the evolution of executive function from adolescence to early adulthood.<sup>44</sup> A summary graph of all of the deployed measures in all the samples is presented below.

<sup>44</sup> Tervo-Clemmens B, Calabro FJ, Parr AC, Fedor J, Foran W, Luna B. A canonical trajectory of executive function maturation from adolescence to adulthood. *Nature Communications*. 2023/10/30 2023;14(1):6922. doi:10.1038/s41467-023-42540-8

Figure 16: Measures of Executive Function<sup>45</sup>



95. From ages 9 -35 accuracy increases and latency decreases. The trajectories of both are very steep during that period. The red dotted line represents age 13, the current “allowable” age for usage of SMs as set by industry. Again, keeping in mind that the solid lines represent “averages” at any given age, half of all children are *below* that estimate. Note how much accuracy increases and latency decreases after the red line.

96. This is something that the defendant platforms appear to realize. For example, TikTok found that even among those that enabled their screen management tools, they saw no

<sup>45</sup> Tervo-Clemmens B, Calabro FJ, Parr AC, Fedor J, Foran W, Luna B. A canonical trajectory of executive function maturation from adolescence to adulthood. *Nature Communications*. 2023/10/30 2023;14(1):6922. doi:10.1038/s41467-023-42540-8

benefits (i.e. screen time use was not reduced) for their <18-year-old users, which they explained by saying “minors do not have the executive function to control their screen time.”<sup>46</sup>

## VIII. Addiction

97. Addiction is a construct based on directly measurable psychological and physiological attributes related to reliance on and/or withdrawal from a substrate. It is a complex condition characterized by compulsive engagement in rewarding stimuli despite adverse consequences. It often involves substances like drugs or alcohol. Teenagers and young adults who biologically lack higher cortical functioning including impulse control are more vulnerable to potentially addictive substances and behaviors. Epidemiological studies have shown that earlier onset of drug intake is associated with greater likelihood of development of substance use problems.<sup>47</sup> In fact, the majority of problematic substance users (e.g. tobacco and alcohol) begin usage before the age of 21.<sup>48</sup>

98. Addiction can also be due to tolerance and withdrawal from certain behaviors. Not all behavioral addictions are currently recognized in the Diagnostic and Statistical Manual of Mental Disorder, Fifth Edition, (DSM-5) which is the American Psychiatric Association’s (APA) guide to mental and brain-related conditions.<sup>49</sup> However, there is an increasing recognition of the

---

<sup>46</sup> TIKTOK3047MDL-039-LARK-00213033, -3036

<sup>47</sup> Crews F, He J, Hodge C. Adolescent cortical development: a critical period of vulnerability for addiction. *Pharmacol Biochem Behav.* Feb 2007;86(2):189-99. doi:10.1016/j.pbb.2006.12.001

<sup>48</sup> See Goldstein RB, Dawson DA, Grant BF. Antisocial Behavioral Syndromes in Adulthood and Alcohol Use Disorder Treatment over Three-Year Follow-Up: Results from Wave 2 of the National Epidemiologic Survey on Alcohol and Related Conditions. *J Am Psychiatr Nurses Assoc.* Jul 2010;16(4):212-26. doi:10.1177/1078390310375846; King KM, Chassin L. A prospective study of the effects of age of initiation of alcohol and drug use on young adult substance dependence. *J Stud Alcohol Drugs.* Mar 2007;68(2):256-65. doi:10.15288/jsad.2007.68.256.

<sup>49</sup> It should be noted that the APA and DSM typically are exceedingly slow to characterize official diagnoses. For example, consider alcoholism. On September 13, 1957, Jim Willis, a

need for the DSM-5 to do so. To date, only a single behavioral addiction, gambling, is officially recognized by the DSM-5. The DSM-5 considered including “gaming disorder” in its 2013 edition but determined it was “in need of further study.”<sup>50</sup> I am currently a member of a committee that is proposing that gaming disorder be included in a revision (DSM-5-TR). Meta has rightly recognized in an internal document that “medical diagnoses change definitions over time based on new evidence.”<sup>51</sup> In 2022, the World Health Organization (WHO) did recognize the existence of “gaming disorder” as a clinical entity, and it is included as a diagnosis in the International Classifications of Disease 11 (ICD-11) which is the manual that physicians use to diagnose patients.<sup>52</sup>

99. Recently, the American Psychiatric Association has recognized “technology addiction” as “excessive and compulsive use of the internet or online activities [that] can lead to negative consequences in various aspects of an individual’s life.”<sup>53</sup> “Social media addiction” is recognized as its own condition, characterized as “involv[ing] problematic and compulsive use of social media; an obsessive need to check and update social media platforms, often resulting in

---

recovering alcoholic who had achieved sobriety through Alcoholics Anonymous, started Gamblers Anonymous using an analogous 12 step process. It wasn’t until 1980 (23 years later) that the affliction was included in the DSMIII and was then listed as “pathological gambling.” It was officially labeled an “addiction” in the 2013 edition (DSM-V) the same one that said that gaming disorder was “in need of further study.” Although the APA now calls it an “addiction,” members of GA still refer to it as “compulsive gambling” as they have since its inception. In the 56 years it took for the APA to classify and re-classify gambling, millions of people were successfully treated for it using the GA approach all over the world. The nomenclature here—problematic, compulsive, addictive—matters less than the manifest damage the behavior brings.

<sup>50</sup> American Psychiatric Association. *Desk reference to the diagnostic criteria from DSM-5*. American Psychiatric Publishing; 2013:xlvi, p. 395

<sup>51</sup> META3047MDL-014-00359270, -9278

<sup>52</sup> Organization WH. *International Classification of Diseases Eleventh Revision (ICD-11)*. World Health Organization; 2022.

<sup>53</sup> American Psychiatric Association, What is Technology Addiction?, <https://www.psychiatry.org/patients-families/technology-addictions-social-media-and-more/what-is-technology-addiction> (last accessed Apr. 14, 2025).

problems in functioning and disrupted real-world relationships.”<sup>54</sup> The APA further recognizes that “children and adolescents are particularly vulnerable to technological addiction because their brains are still developing” and “excessive problematic use of social media” has the potential to develop into a behavioral addiction for children and adolescents.<sup>55</sup>

100. Because of the increased risks of social media to youth, several national associations and reports have been published with recommendations for actions that can be taken to help reduce the risk of mental health injury. Some of these reports include recommendations made in the textbook I edited, *Handbook of Children and Screens*, as well as the “Social Media and Youth Mental Health” Report by the U.S. Surgeon General in 2023;<sup>56</sup> “Social Media and Adolescent Health” by the National Academies of Sciences, Engineering, and Medicine in 2024<sup>57</sup>; “Health Advisory on Social Media Use in Adolescence” by the American Psychological Association in 2023<sup>58</sup>; and a report by the Jed Foundation in 2024.<sup>59,60</sup>

101. The American Academy of Pediatrics also recognizes problematic use and social media addiction and advises parents that “It’s also important to recognize that it’s not something

---

<sup>54</sup> American Psychiatric Association, What is Technology Addiction?, <https://www.psychiatry.org/patients-families/technology-addictions-social-media-and-more/what-is-technology-addiction> (last accessed Apr. 14, 2025).

<sup>55</sup> American Psychiatric Association, What is Technology Addiction?, <https://www.psychiatry.org/patients-families/technology-addictions-social-media-and-more/what-is-technology-addiction> (last accessed Apr. 14, 2025).

<sup>56</sup> Office of the Surgeon General, *Social Media and Youth Mental Health: The U.S. Surgeon General’s Advisory* (2023), available at <https://pubmed.ncbi.nlm.nih.gov/37721985/>

<sup>57</sup> <https://nap.nationalacademies.org/catalog/27396/social-media-and-adolescent-health>

<sup>58</sup> <https://www.apa.org/topics/social-media-internet/health-advisory-adolescent-social-media-use>

<sup>59</sup> <https://jedfoundation.org/the-jed-foundation-jed-recommendations-for-safeguarding-youth-well-being-on-social-media-platforms/>

<sup>60</sup> Many of these reports readily recognize the deleterious effects of social media on children. I would note that the 2023 NASEM report takes an improbably conservative approach to the literature recognizing harms to children – a body of literature that has only grown since its publication.

wrong with the teen using the platform causing them to feel this way; many interactive technologies are specifically designed to capture and hold a user’s interest. It can be hard for children and teens to overcome those design features.”<sup>61</sup> The National Eating Disorders Association also recognizes that “research is increasingly clear that media does indeed contribute and that exposure to and pressure exerted by media increase body dissatisfaction and disordered eating.”<sup>62</sup> Taken together, these consensus statements leave little doubt that leading professional medical and psychological organizations recognize that social media and its problematic and addictive usage is harming children and teenagers today.

**A. Validated Social Media Addiction Scales**

102. A variety of screening instruments for what has been called “Social Network Use Disorder” have been developed. A systematic review of the scales revealed that two of them have the best validation data to support them: the Social Media Disorder Scale (SMDS) and the Bergen Social Media Addiction Scale Short Form (BSMAS-SF).<sup>63</sup> Both of these measure features of substance abuse disorder including: salience, tolerance, preoccupation, impaired role performance, loss of control, and withdrawal symptoms. These features are consistent with those considered by the DSM-5.

103. Validation of scales includes collecting normative data from a large and diverse sample of people and then developing a clinical cutoff (the details of how that is done are beyond

---

<sup>61</sup> <https://www.aap.org/en/patient-care/media-and-children/center-of-excellence-on-social-media-and-youth-mental-health/qa-portal/qa-portal-library/qa-portal-library-questions/problematic-technology-use/?srsltid=AfmBOorKPQQSzENMf3PnJhedPK39d89jvoL7LLSIH9OsEUa5MZ6624M7>

<sup>62</sup> <https://www.nationaleatingdisorders.org/media-and-eating-disorders/>

<sup>63</sup> Schlossarek S, Schmidt H, Bischof A, et al. Psychometric Properties of Screening Instruments for Social Network Use Disorder in Children and Adolescents: A Systematic Review. *JAMA Pediatr.* Apr 1 2023;177(4):419-426. doi:10.1001/jamapediatrics.2022.5741

the scope of this report). One might rightly ask if this usage pattern constitutes pathology or simply enthusiastic usage. To test this, researchers assess “convergent validity.” Specifically, how does the measured construct correlate with other outcomes we would expect it to predict. For example, we would predict that people with problematic social media usage would have increased risks of other mental health disorders (as is the case with other addictions). To that end, a recent meta-analysis of 18 studies that assessed “problematic social media usage” with anxiety and depressive symptoms found a correlation of .348 (“medium”) and .273 (almost “medium”) respectively.<sup>64</sup> While the defendants’ experts note that one can substitute another construct (e.g. socializing with friends) for time on social media in the BSMAS, and find alarming rates of “addiction” to friends, that finding has not been subjected to convergent validity.

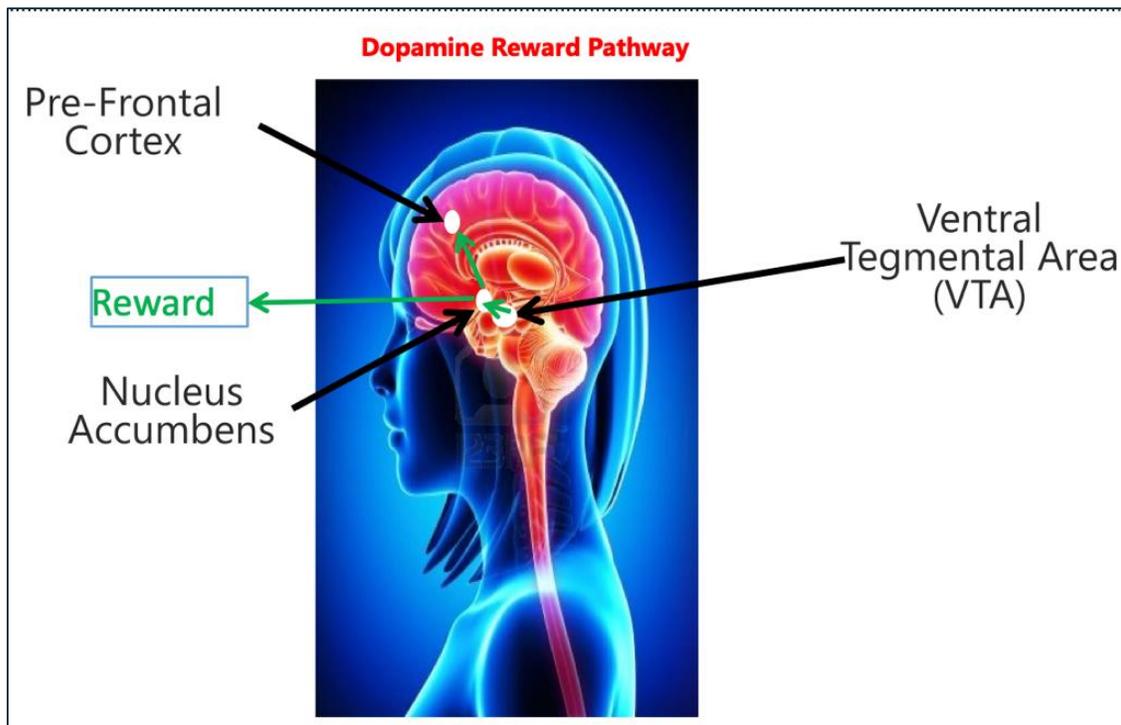
#### **B. The Mechanism of Addiction**

104. Addiction (both behavioral and substance-based) is grounded in the brain's dopamine reward system. Exposure to a stimulus is processed in the Ventral Tegmental Area (VTA) of the mid-brain that releases dopamine. When that stimulus is “favorable,” the VTA signals the nucleus accumbens (the pleasure center of the brain) which in turn signals the Pre-Frontal Cortex (the executive center of the brain as discussed before) effectively saying “I liked that” so “do it again” or “get more of it.”

---

<sup>64</sup> Shannon H, Bush K, Villeneuve P, Hellemans K, Guimond S. Problematic Social Media Use in Adolescents and Young Adults: Systematic Review and Meta-analysis. *JMIR Ment Health*. 2022;9(4).

**Figure 17: The Dopamine Reward Pathway**



105. This is a generic pathway; it is activated for example when parents praise children for behaviors (e.g. saying “thank you”) thereby increasing the probability (hopefully) that they will act that way more often. In pathological circumstances, given prolonged exposure to intensely pleasurable stimuli, the brain’s natural reward pathways can be altered, making it increasingly difficult to experience pleasure from other activities. At a neurobiological level then, behavioral and substance-based addictions have a final common pathway. Functional magnetic resonance imaging (fMRI) studies have demonstrated that social media usage (and Facebook in particular) activates the nucleus accumbens.<sup>65</sup> Indeed, Meta documents acknowledge that Facebook “does

<sup>65</sup> See Meshi D, Morawetz C, Heekeren HR. Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use. Original Research. *Frontiers in Human Neuroscience*. 2013-August-29 2013;7doi:10.3389/fnhum.2013.00439; Wadsley M, Ihssen N. A Systematic Review of Structural and Functional MRI Studies Investigating Social Networking Site Use. *Brain Sciences*. 2023;13(5):787.

activate the brain's reward system."<sup>66</sup> Moreover, a three year longitudinal study of 6<sup>th</sup> and 7<sup>th</sup> grade students found changes in functional activation of the brain based on reported habitual checking of social media sites at baseline.<sup>67</sup> Specifically, habitual checkers' brains demonstrated differential activation of specific regions in anticipation of social cues compared to non-habitual checkers. This suggests that their brains may be being conditioned as a result of repeated activation (a feature of addiction).

106. The causes of addiction are multifaceted, involving a combination of genetic, environmental, and psychological factors. Although not determinative, genetics may predispose individuals to addiction, while environmental factors such as peer pressure, stress, or trauma can trigger problematic substance use or behaviors. Mental health conditions like depression and anxiety are also linked to addiction, as individuals may turn to substances or behaviors as a form of self-medication. Early exposure to addictive substances, particularly during childhood, increases the likelihood of addiction later in life.

### **C. The Formation of Habit—Distinguishing Addiction from Habit**

107. Addiction can be distinguished from habit. Although both involve repeated behaviors and activation of the dopamine reward center, they differ in terms of control, dependence, and consequences. A habit is a routine that is performed regularly and often unconsciously or reflexively (brushing one's teeth for example). People can typically stop habits without experiencing withdrawal. But habits can become addictions. A habitual use of alcohol after work can progress to alcohol misuse or abuse, for example.

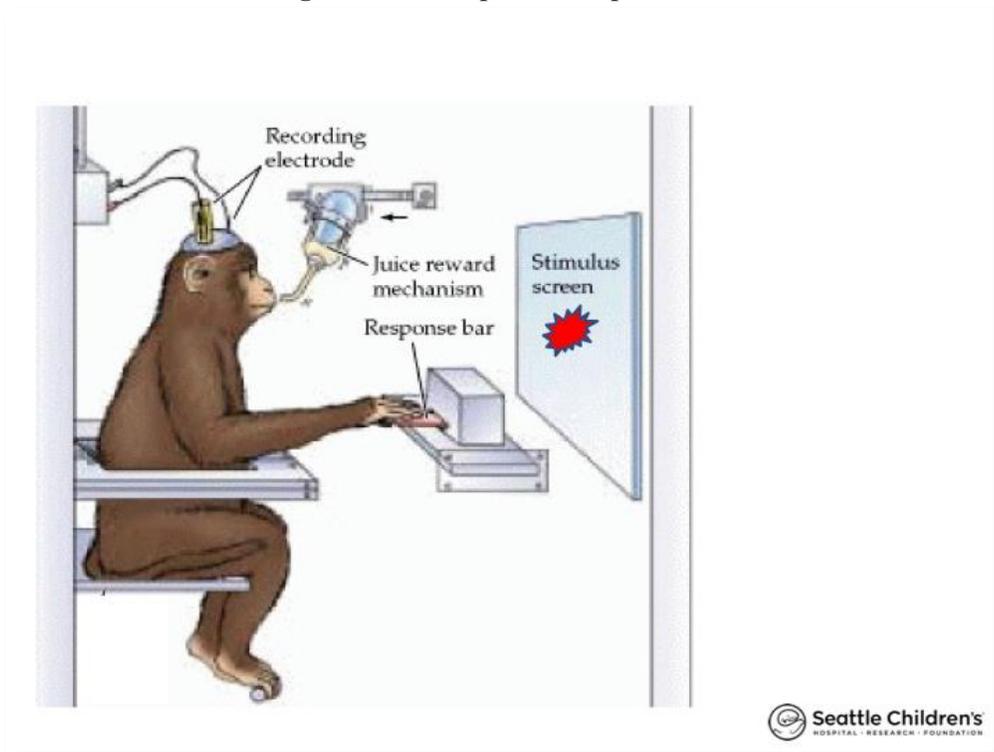
---

<sup>66</sup> META3047MDL-014-00359270, -9288

<sup>67</sup> Maza MT, Fox KA, Kwon S-J, et al. Association of Habitual Checking Behaviors on Social Media With Longitudinal Functional Brain Development. *JAMA Pediatrics*. 2023;177(2):160-167. doi:10.1001/jamapediatrics.2022.4924

108. Much of what we know about the emergence of habit comes from seminal work done by a neuroscientist named Wolfram Schultz with his colleagues and a *Macaque* monkey named Julio. In a typical experiment, Julio was seated in a chamber with a device that recorded activity in his Nucleus Accumbens (the pleasure center as detailed above). He would stare at a blank screen while being given access to a response bar. At random periodic intervals, a shape would appear on the screen and if he pushed the lever when it did, a juice reward dispenser delivered blackberry juice (his favorite libation).<sup>68</sup>

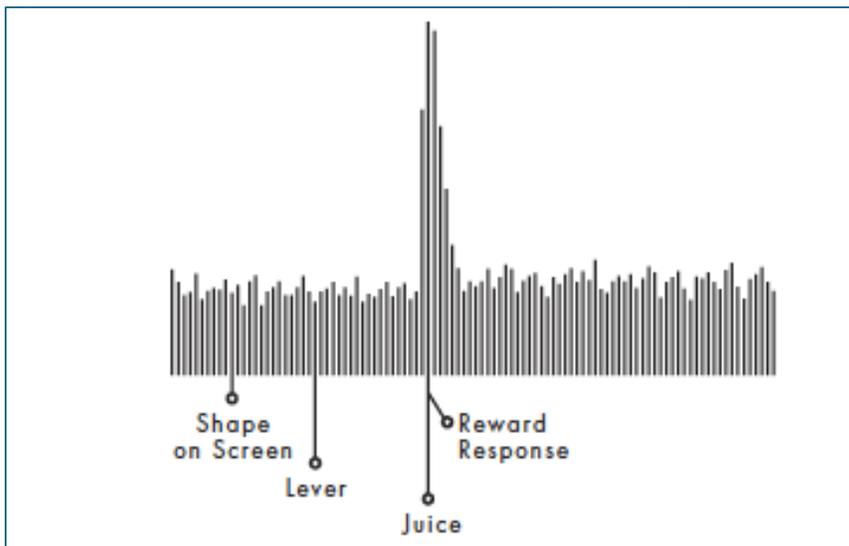
**Figure 18: Picture Simulating Electrode Dopamine Experiment**



<sup>68</sup> Ljungberg T, Apicella P, Schultz W. Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol.* 1992;67(1):145-63.

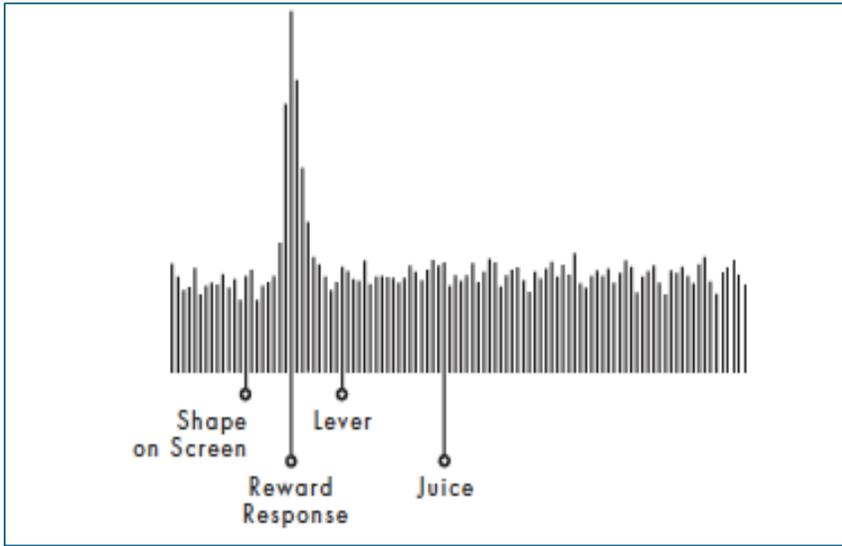
109. The figure below shows the activity recorded in Julio’s nucleus accumbens during the training session during which he was effectively learning that pushing the response lever *when* a shape appeared led to a reward. What can be seen is a spike in activity, a “reward” response, when the juice is delivered, which happens after the shape appears on the screen and the lever is pushed.

**Figure 19: Nucleus Accumbens Activity during “Training”**

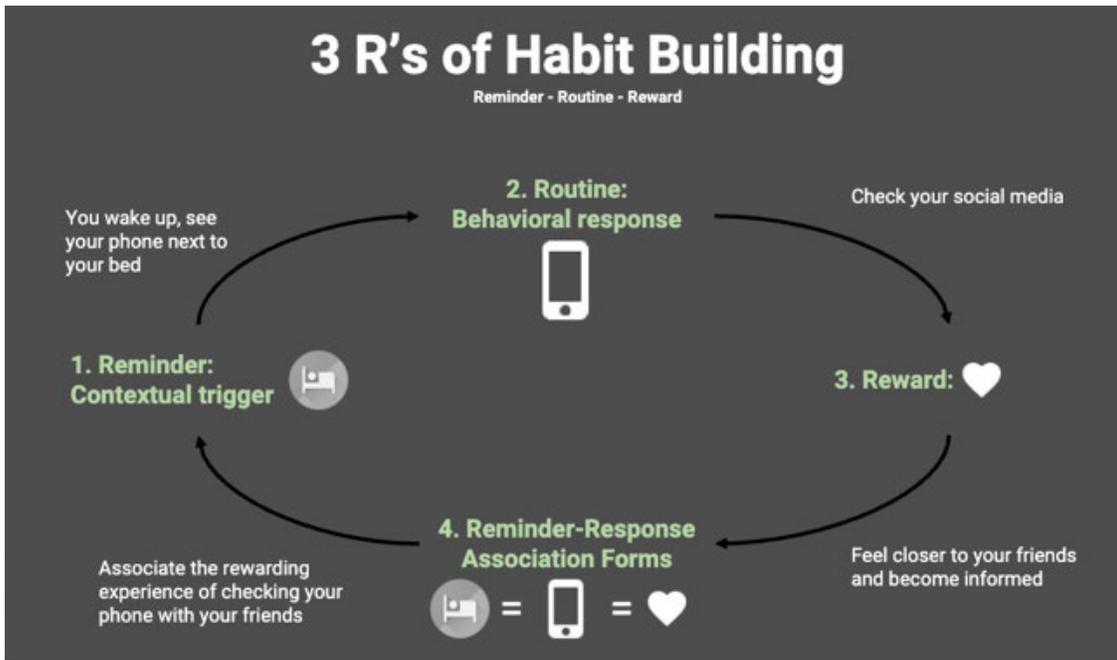


110. The next figure shows the pattern after Julio is “trained.” The reward response occurs the moment he sees the shape appear on the screen and *before* the juice is actually dispensed which is to say he experiences the reward activation simply because he anticipates it because he *associates* the shape with pleasure.

Figure 20: Nucleus Accumbens Activity once "Trained"

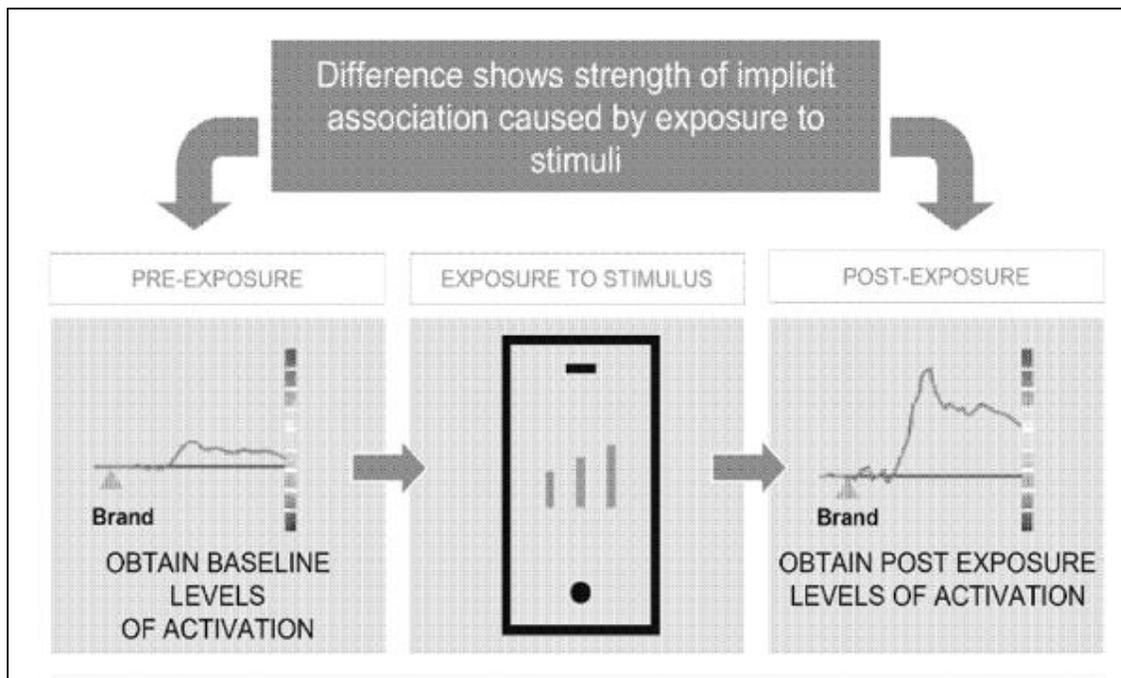


111. YouTube is keenly aware of the science behind habit formation as it appeared to use it to help build its brand. In a December 2024 presentation by [REDACTED] (User Experience Researcher on the growth team) the following slide appears:



Document 6: GOOG-3047MDL-01268284 at Slide 6

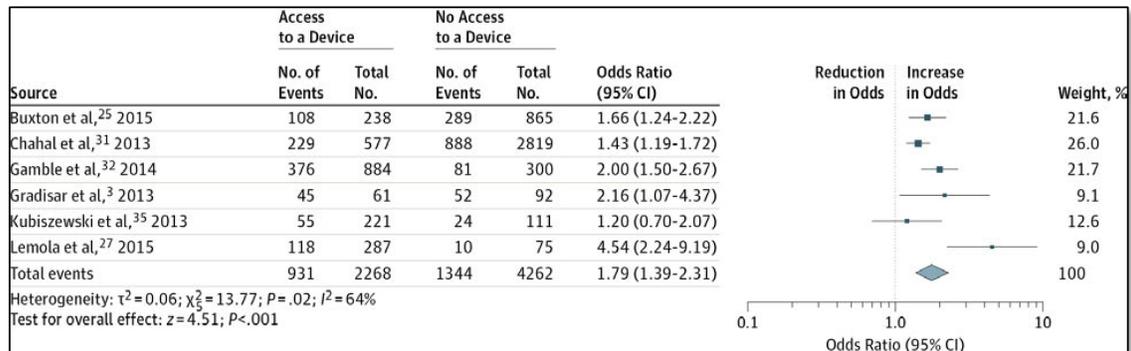
112. TikTok also engaged Nielsen Consumer Neuroscience to explore deploying the habit paradigm to associate their videos with a particular product advertised in tandem with it. In an approach reminiscent of Schulz’s work, except with humans instead of Julio, they monitor brain activation before and after exposure to an advertisement.



Document 7: TIKTOK3047MDL-020-00376995, -7003

113. The formation of habit explains, in part, how many adolescents associate the mere presence of their device with pleasure to such a great enough extent that it can be difficult to resist. This “habitual” reaction has been documented in studies that assess the *presence* of a phone in a child’s bedroom and the occurrence of sleep problems. Figure 19 (below) shows data from a metaanalysis (as explained above, metaanalyses provide a summary estimate of multiple different studies).

**Figure 21: Summary of Metanalysis Assessing Effect of the Presence of a Phone in a Child’s Bedroom and the Occurrence of Sleep Problems<sup>69</sup>**



114. The summary estimate shows that the mere presence (not necessarily the *usage*) of a device in a bedroom is associated with a 79% increased risk of sleep problems. The device is a proxy for the apps it houses (e.g. social media) and there is ample evidence that SM increases sleep problems. One plausible explanation is that the habitual usage of phones is activating the reward reflex – for example, by inducing the teen to think of “likes” on the SM platform – and thereby impeding sleep initiation and quality. This will be discussed in greater detail in the “Sleep” section of this report (Section X.C).

**D. Problematic Social Media Use Can Occur as Habitual, Compulsive, or Addictive**

115. Whether it is called “compulsive,” “addictive,” or “problematic” some patterns of digital media use are untoward outcomes of their own as they impede daily functioning. But they also exacerbate all the other direct effects of social media sites since they drive increased exposure to the platforms.

<sup>69</sup> Carter B, Rees P, Hale L, Bhattacharjee D, Paradkar MS. Association Between Portable Screen-Based Media Device Access or Use and Sleep Outcomes: A Systematic Review and Meta-analysis. *JAMA Pediatr.* Dec 1 2016;170(12):1202-1208. doi:10.1001/jamapediatrics.2016.2341

116. The consequences of addiction extend beyond the individual, affecting their relationships, career, and overall well-being. Psychologically, addiction often results in feelings of isolation, guilt, or shame, further perpetuating its cycle. An essential feature of “addiction” is that it impedes activities of daily living such that one chooses to indulge in the activity or substance rather than work or socialize. Problematic social media use can lead to diminished sleep and missed school. Attempts at reducing usage can cause anxiety, depression, and irritability and conflict when attempted by exogenous agents e.g. parents.<sup>70</sup>

117. Addiction lies at one extreme of a usage and behavior continuum as depicted below:

**Figure 22: Social Network Site Usage Continuum**



118. While addiction is a clinical diagnosis and typically based on established and accepted expert criteria derived from the scientific literature, anything to the “right” of casual usage in the schema above increases the likelihood of untoward mental and physical health effects. One needn’t meet criteria for a clinical diagnosis of addiction *per se* to be negatively impacted by a behavior. In that sense, addiction is an identifiable harm resulting from the larger umbrella of problematic usage, which also causes anxiety, depression, and other mental health harms. For example, one need not be diagnosed with “alcoholism” *per se* to suffer negative health effects of

---

<sup>70</sup> Radesky J, Weeks HM, Schaller A, Robb M, Mann S, Lenhart, Constant Companion: A Week in the Life of a Young Person's Smartphone Use (2023), available at [https://www.common sense media.org/sites/default/files/research/report/2023-cs-smartphone-research-report\\_final-for-web.pdf](https://www.common sense media.org/sites/default/files/research/report/2023-cs-smartphone-research-report_final-for-web.pdf)

“excessive” drinking. Increases in all-cause mortality can be seen even at “medium” level drinkers.<sup>71</sup>

119. In the case of substance abuse in particular, decades of research established a taxonomy that is more nuanced than simply alcohol “abuse” vs “use.” The National Institute of Alcohol Abuse and Alcoholism (NIAAA) has established different usage patterns that have been studied both for their independent effects on functioning and how each might ultimately lead to the clinical entity of Alcohol Use Disorder (Table 1). Any behavior beyond “moderate drinking” i.e. below the thick black line is viewed as problematic and can lead to untoward health effects including but not limited to Alcohol Use Disorder itself. In a recent viewpoint in *JAMA*, I proposed developing an analogous taxonomy for “Media Use Disorder.”

---

<sup>71</sup> Zhao J, Stockwell T, Naimi T, Churchill S, Clay J, Sherk A. Association Between Daily Alcohol Intake and Risk of All-Cause Mortality: A Systematic Review and Meta-analyses. *JAMA Network Open*. 2023;6(3):e236185-e236185. doi:10.1001/jamanetworkopen.2023.6185

**Table 4: Comparison of NIAAA Alcohol Drinking Pattern Definitions with Provisional Cutpoints from SBU MEDiA Study Patterns**

<b>NIAAA Alcohol Drinking Patterns</b>	<b>SBU MEDIA Study Usage Patterns</b>
<b>Moderate Drinking</b>	<b>Moderate Media Use</b>
Two drinks or less in a day for men and one drink or less in a day for women, when alcohol is consumed. Drinking less is better for health than drinking more.	Less than 5 hours per day (<50 <sup>th</sup> percentile)
<b>Binge Drinking</b>	<b>Binge Media Use</b>
Five or more drinks (male) or four or more drinks (female), in about two hours.	More than 4 hours in a continuous session*
<b>High-Intensity Drinking</b>	<b>High-Intensity Media Use</b>
10 or more standard drinks (or alcohol drink equivalents) for males and eight or more for females.	12 hours or more in one day (95 <sup>th</sup> percentile)
<b>Heavy Drinking</b>	<b>Heavy Media Use</b>
<ul style="list-style-type: none"> <li>• Consuming five or more drinks on any day or 15 or more per week (male)</li> <li>• Consuming four or more on any day or eight or more drinks per week (female)</li> </ul>	9 hours or more per day (85 <sup>th</sup> percentile) or 60 hours per week (85 <sup>th</sup> percentile)
<b>Alcohol Misuse</b>	<b>Media Misuse</b>
Alcohol misuse refers to drinking in a manner, situation, amount, or frequency that could cause harm to the person who drinks or to those around them. Alcohol misuse includes binge drinking and heavy alcohol use.	Media misuse entails binge, high-intensity, and heavy viewing that include inappropriate timing (e.g., school or sleeping hours), inappropriate content (e.g., cyberbullying, pro-eating disorder content), or viewing during dangerous situations (e.g., distracted driving).
<b>Alcohol Use Disorder</b>	<b>Media Use Disorder</b>
AUD is a medical condition characterized by an impaired ability to stop or control alcohol use despite adverse social, occupational, or health consequences. It encompasses the conditions that some people refer to as alcohol abuse, alcohol dependence, alcohol addiction, and the colloquial term, alcoholism.	MUD is characterized by an impaired ability to stop or control media use despite adverse social, occupational/school-related, or mental health consequences. Various validated measures exist related to social media use.

\*The definition of binge smartphone use should be empirically derived. This provisional cut point of 4 hours of continuous viewing is based on a systematic review of definitions of binge-watching overall media use.

120. The point of this taxonomy is to emphasize that time spent on the platforms can be sufficiently excessive that it can impede opportunities for other essential activities of daily living (e.g. in person socializing, playing, exercising). What we call such usage is less important than that simple reality. Given that teenagers should sleep for 8-10 hours per day and attend school for 6-7 hours per day, those that spend 9 or more hours per day on their device must be forgoing sleep or using it during school (or both) and still be on it every free minute of the day. All of the defendants in this case were explicitly told by outside experts and parents that their products induced problematic usage (their own data confirmed this; their own leaders acknowledged it) and yet I found no evidence that they warned teens or their parents of that risk. When asked if Meta warned the public, kids, or parents of the increased risk of problematic use of Instagram, Bejar said plainly, “They did not.”<sup>72</sup>

#### **E. Social Contagion via Social Media Platforms**

121. Social contagion refers to the phenomenon where ideas, emotions, behaviors, or attitudes spread through social networks, much like a virus. It highlights the interconnected nature of human societies, where individuals unconsciously or consciously influence one another. For example, laughter in a group can quickly become contagious, even among those who do not know why others are laughing. Similarly, trends, such as fashion styles or internet challenges, often spread rapidly because individuals adopt behaviors observed in others to feel connected or accepted within a group.

122. The mechanisms behind social contagion are rooted in psychological and sociological principles. Mirror neurons in the brain play a role by enabling individuals to mimic others’ emotions or actions, fostering empathy and shared experiences. Additionally, conformity

---

<sup>72</sup> Arturo Bejar Dep. Tr. at 143:15-20

and peer pressure amplify the spread of behaviors, particularly in tight-knit social circles. While social contagion can have positive effects, such as the adoption of healthy habits or the rapid dissemination of valuable information, it can also perpetuate harmful behaviors like panic during crises or the spread of misinformation. Understanding social contagion is crucial for designing interventions in areas such as public health, education, and online media, where both positive and negative influences can scale rapidly.

123. A 2014 study tested the social contagion phenomenon on social media. It assessed the effects of rainfall on an individual's social media posts and found (unsurprisingly) that rainfall increased the probability of negative posts. But the study also found that it increased the probability of negative posts of rainy-day city people's friends who lived in cities where it was *not* raining on a given day.<sup>73</sup> While this study was observational, it deployed what is known as an instrumental variable design. Specifically, the rainfall in "City A" can cause negative affect in that city but it cannot directly cause negative affect in "City B" where it is *not* raining. Therefore, rain cannot be a confounder and the negativity of an individual in City B can be attributed to their connection with an individual in City A whose mood is drearier because of the precipitation. "Instrumental variable" approaches such as this are among the strongest observational designs because of the way they circumvent potential confounding.

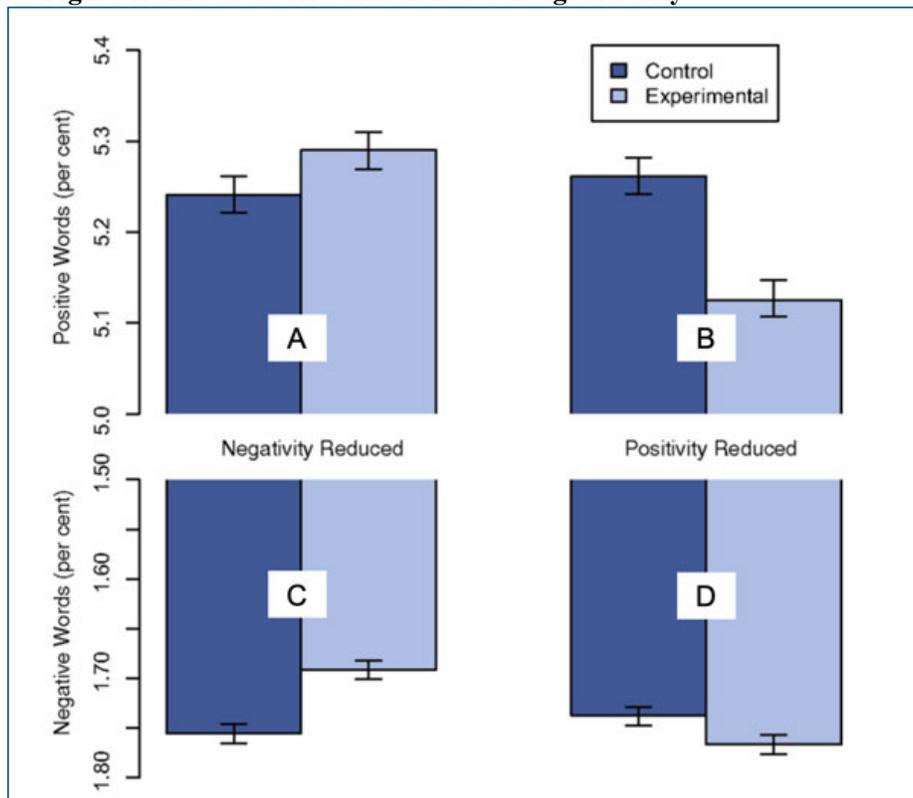
124. Most prominently – in a yet to be repeated publicly available study the likes of which only the social media industry itself could do – Facebook tested social contagion theory on their site using an experimental design in which 689,003 people's news feeds were randomized to

---

<sup>73</sup> Coviello L, Sohn Y, Kramer AD, Marlow C, Franceschetti M, Christakis NA, Fowler JH. Detecting emotional contagion in massive social networks. PLoS One. 2014 Mar 12;9(3):e90315. doi: 10.1371/journal.pone.0090315. PMID: 24621792; PMCID: PMC3951248.

highlight fewer positively and fewer negatively worded posts.<sup>74</sup> The outcome of interest was the valence of the content they subsequently posted. In other words, they tested the hypothesis that seeing more positive content induced one to post more positive content and seeing less negative content induced one to post less negative content.

**Figure 23: Figure from Facebook’s Emotional Contagion Study<sup>75</sup>**



As Panel C to the left shows, reducing the frequency of negative posts on someone’s feeds results in fewer negative words posted by the “experimental” group compared to the “control” group.

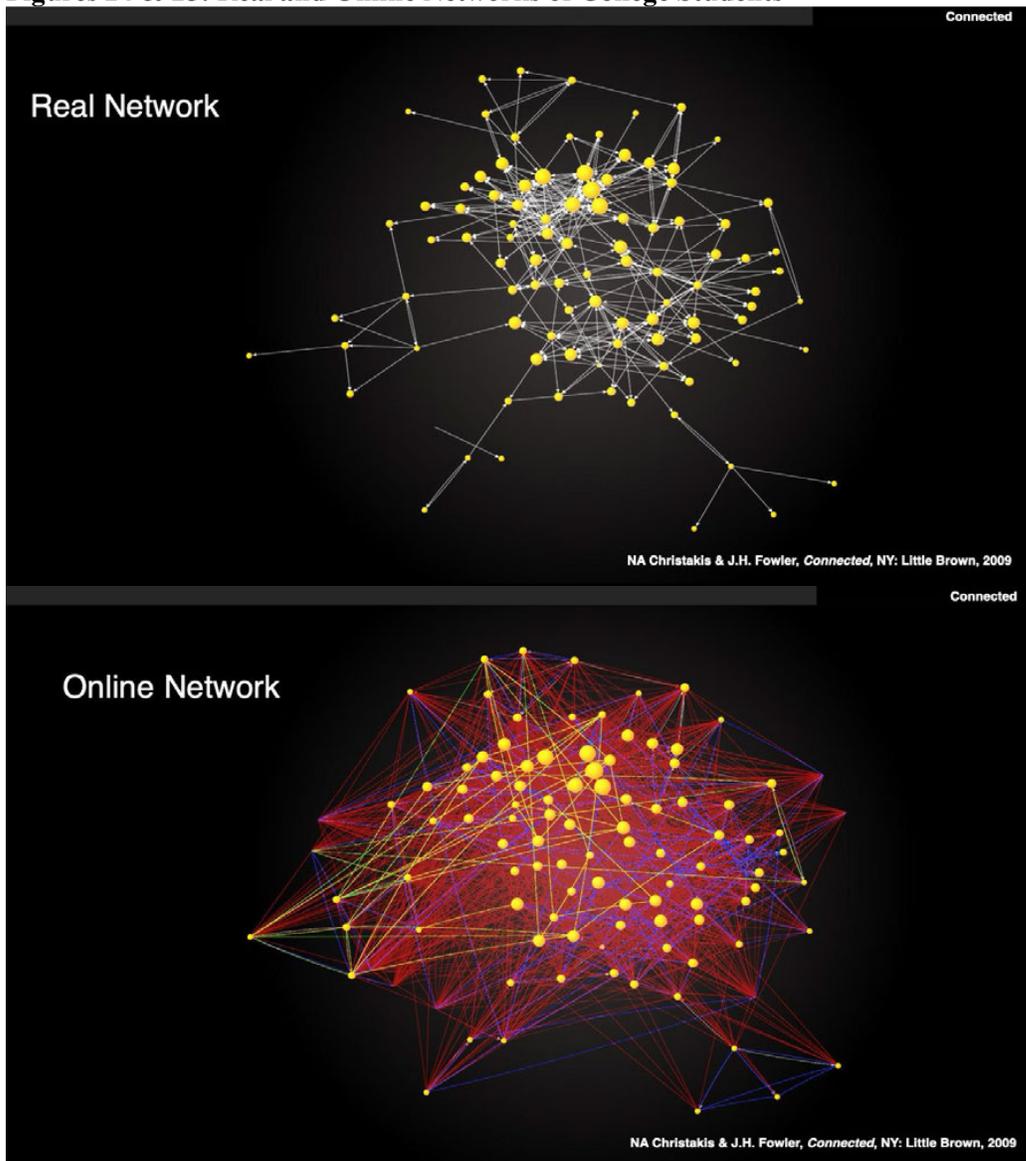
<sup>74</sup> Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. 2014;111(24):8788-8790. doi:doi:10.1073/pnas.1320040111

<sup>75</sup> Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. 2014;111(24):8788-8790. doi:doi:10.1073/pnas.1320040111

Conversely, as Panel B shows, reducing positive posts results in fewer positive words posted (again experimental vs control). But as Panels A & D demonstrate, reduced negative words results in more positive ones and reduced positive posts results in more negative ones.

125. Social media platforms did not create the phenomenon of social contagion, but they provide an extraordinary mechanism to amplify it. In the real world, the negative affect of individuals spreads via the people they have direct contact with and then to the people that those individuals have direct contact with. The online world is decidedly different. The figures below illustrate the social networks of college students in the real-world vs those on Facebook:

Figures 24 & 25: Real and Online Networks of College Students<sup>76</sup>



126. Each yellow dot represents a student, and each line represents a connection between that student and a “friend.” Yellow lines represent “real world” connections and red lines represent online ones. The average number of “friends” students listed in their real-world network is 6.6

---

<sup>76</sup> Nickolas A. Christakis and James H. Fowler (2009), *Connected: The Surprising Power of our Social Networks and How they Shape our Lives*, Little, Brown, New York, NY. 353 pages.

whereas in their online network it is 110. Those differences in network sizes are enabled and facilitated by the platforms' algorithms which constantly suggest additional friends to add to one's network and teenagers in particular have a propensity to maximize theirs. While it is true that a real-world interaction has more impact on another individual than an online one on *average*, the sheer number of interactions is so much greater online that the potential for social contagion effect from a societal perspective is considerably larger.

#### **F. Fear of Missing Out (FoMo)**

127. "Fear of Missing Out" (FoMo) is a relatively recently defined construct (circa 2004) whereby affected individuals are apprehensive that not checking social media sites will result in missing opportunities afforded to one's "friends" online. It has been defined as "the uneasy and sometimes all-consuming feeling that you're missing out – that your peers are doing, in the know about, or in possession of more or something better than you."<sup>77</sup> FoMo can result in a compulsive need to maintain connection to social media to mitigate it although paradoxically frequent checking of one's social media profile can confirm that, in fact, others are having fun that does not include the affected individual. Several design aspects of social media—including notifications, likes, infinite scroll, and friends' maps or friends you may know features—take advantage of this psychological phenomenon to drive usage. FoMo-driven social media usage can at once provide reassurance (or even opportunities for engagement) and distress. Accordingly, and not surprisingly, higher levels of FoMo are associated with both more positive and negative attitudes about social

---

<sup>77</sup> A. Przybylski, et. al., *Motivational, Emotional, And Behavioral Correlates Of Fear Of Missing Out*, 29 COMPUTERS IN HUMAN BEHAVIOR 1841-1848 (2013), located at <https://www.sciencedirect.com/science/article/abs/pii/S0747563213000800>

media use.<sup>78</sup> Although it clearly is contiguous with and has features in common with addiction or habitual usage, I specifically call it out as its own construct as it supports the complicated and heterogenous relationship between individuals and social media use.<sup>79</sup>

128. The original and most widely used scale for FoMo was developed in 2010 by Przybylski and it is presented below:

**Figure 26: The Final 10-Item of the Fear of Missing Out Scale**

<b>Table 3: The final 10-item version of the Fear of Missing Out scale (FoMoFoMoS)</b>				
Below is a collection of statements about your everyday experience. Using the scale provided please indicate how true each statement is of your general experiences. Please answer according to what really reflects your experiences rather than what you think your experiences should be. Please treat each item separately from every other item.				
<b>Not at all true of me</b>	<b>Slightly true of me</b>	<b>Moderately true of me</b>	<b>Very true of me</b>	<b>Extremely true of me</b>
1	2	3	4	5
1. I fear others have more rewarding experiences than me.				
2. I fear my friends have more rewarding experiences than me.				
3. I get worried when I find out my friends are having fun without me.				
4. I get anxious when I don't know what my friends are up to.				
5. It is important that I understand my friends "in jokes".				
6. Sometimes, I wonder if I spend too much time keeping up with what is going on.				
7. It bothers me when I miss an opportunity to meet up with friends.				
8. When I have a good time, it is important for me to share the details online (e.g. updating status).				
9. When I miss out on a planned get-together it bothers me.				
10. When I go on vacation, I continue to keep tabs on what my friends are doing.				

<sup>78</sup> Przybylski AK, Murayama K, DeHaan CR, Gladwell V. Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*. 2013/07/01/ 2013;29(4):1841-1848. doi:<https://doi.org/10.1016/j.chb.2013.02.014>

<sup>79</sup> Akbari M, Seydavi M, Palmieri S, Mansueto G, Caselli G, Spada MM. Fear of missing out (FoMoFoMo) and internet use: A comprehensive systematic review and meta-analysis. *Journal of Behavioral Addictions*. 31 Dec. 2021;10(4):879-900.

129. It is notable that only a single item from the scale, (#8), makes specific reference to being “online.” In fact, as a construct, FoMo could very well have existed before or despite the internet. To that end, some have argued that FoMo should be viewed as being both a “trait” (something that is stable and enduring) and a “state” (something that is temporary and situational).<sup>80</sup> Humans have likely, for example, worried that their friends “were having fun without them” for decades (if not millennia) but there was no ready way to either deliberately or inadvertently check if any or all of them were. What is more, design features of apps (e.g. friend locations and alerts) do something that was previously impossible: passively and seamlessly track your friends so as to alert you when they may have congregated without you at a fun location near you. So, while SM’s may not have created FoMo, the use of SM can greatly exacerbate it, increase its prevalence, or induce it at least temporarily.

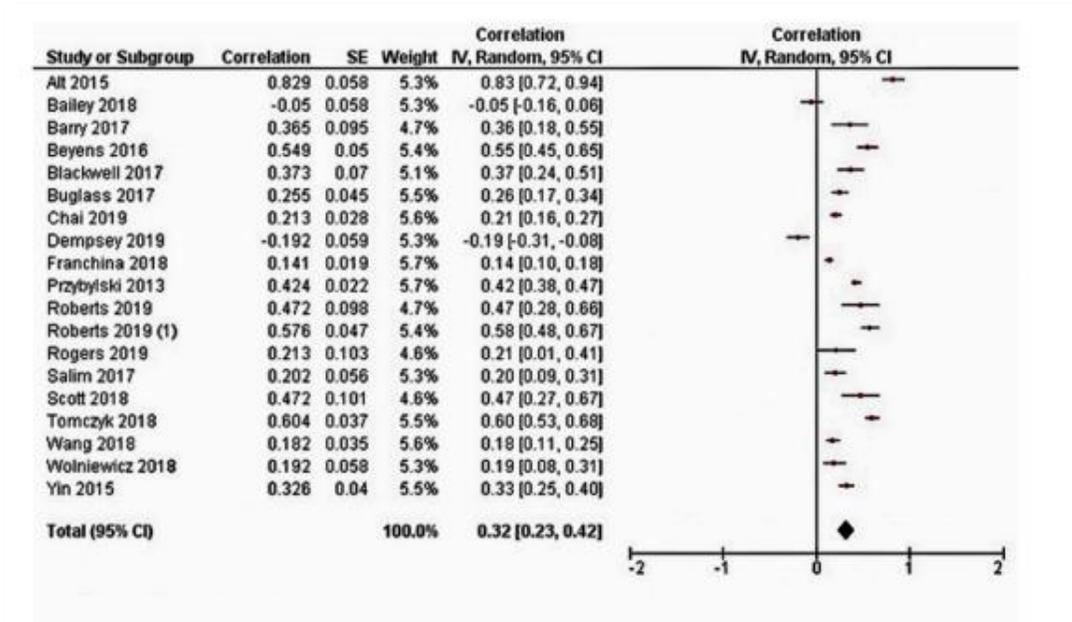
130. Conceptually, FoMo and social media use are mutually reinforcing. That is, the drive to compulsively check SM to reduce FoMo leads to problematic SM use and in turn additional FoMo. In support of this, a 2020 metanalysis of 33 studies included 13 which analyzed the relationship between FoMo and social media use, 14 which analyzed the relationship between FoMo and problematic social media use and the remaining six studies which examined both relationships.<sup>81</sup> The metanalytic results are summarized in the figures below. FoMo had a moderate correlation with SM use ( $r=.32$ ) and a moderate/strong correlation with problematic SM use ( $r=.49$ ).

---

<sup>80</sup> Wegmann E, Oberst U, Stodt B, Brand M. Online-specific fear of missing out and Internet-use expectancies contribute to symptoms of Internet-communication disorder. *Addictive Behaviors Reports*. 2017;5:33-42. doi:<https://doi.org/10.1016/j.abrep.2017.04.001>

<sup>81</sup> Fioravanti G, Casale S, Benucci SB, et al. Fear of missing out and social networking sites use and abuse: A meta-analysis. *Computers in Human Behavior*. 2021/09/01/ 2021;122:106839. doi:<https://doi.org/10.1016/j.chb.2021.106839>

Figure 27: FoMo and Problematic Social Media Use



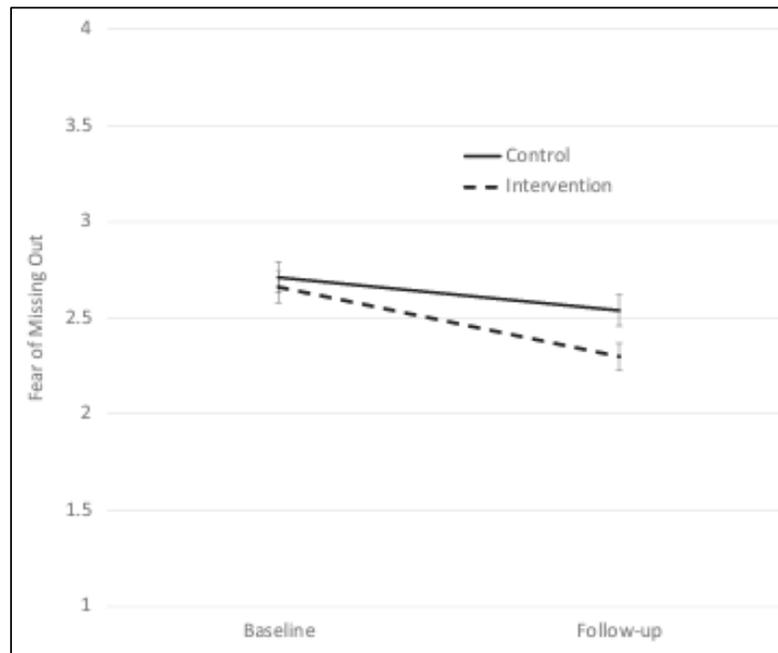
The stronger effect size for problematic usage is consistent with FoMo being a driver of excessive social media use.

131. A recent (2025) randomized controlled trial of social media reduction (too recent to be included in any metaanalyses to date) conducted in 220 college students with at least two symptoms of anxiety or depression asked participants in the intervention arm to reduce social media use to no more than 1 hour per day (the control arm could continue as usual).<sup>82</sup> The researchers used daily screen shots of phone screentime reports to measure social media usage and assessed FoMo, depression, anxiety, and sleep as outcomes after 3 weeks of treatment. FoMo

<sup>82</sup> Davis CG, Goldfield GS. Limiting social media use decreases depression, anxiety, and fear of missing out in youth with emotional distress: A randomized controlled trial. *Psychology of Popular Media*. 2025;14(1):1-11. doi:10.1037/ppm0000536

results using the Przybylski measure are presented below; depression, anxiety and sleep outcomes are presented in their respective sections later in this report.<sup>83</sup>

**Figure 28: FoMo Results Using Przybylski measure**



Students in the control arm used social media 188.76 min per day and students in the intervention arm used social media an average of 78.25 min per day (a reduction of approximately 50%).<sup>84</sup> As can be seen in the figure, FoMo was significantly reduced in the intervention compared to the control. The experimental design of this study presents a very strong causal argument that social media use plays a causal role in FoMo for people with underlying depression or anxiety.

<sup>83</sup> Davis CG, Goldfield GS. Limiting social media use decreases depression, anxiety, and fear of missing out in youth with emotional distress: A randomized controlled trial. *Psychology of Popular Media*. 2025;14(1):1-11. doi:10.1037/ppm0000536

<sup>84</sup> Davis CG, Goldfield GS. Limiting social media use decreases depression, anxiety, and fear of missing out in youth with emotional distress: A randomized controlled trial. *Psychology of Popular Media*. 2025;14(1):1-11. doi:10.1037/ppm0000536

132. An internal presentation by Shruti Bhutada, wellbeing lead at Meta, illustrated what Meta’s research showed regarding what teens perceived to be the sources and causes of FoMo.

what are people worried they'll miss?	possible contributing factors
<ul style="list-style-type: none"><li>• an <b>action</b> (e.g., a conversation, a commenting back and forth, or missing some promotion)</li><li>• on <b>information</b> (e.g., find out about an event, etc.).</li><li>• chances to <b>nourish relationships</b> (e.g. birthday reminders, business promotions, personal updates from friends)</li></ul>	<ul style="list-style-type: none"><li>• <b>Signing up for more notifications</b> to avoid missing posts</li><li>• <b>Limited news feed controls:</b> Feel they have limited control over what will be in their news feed and they might make them miss posts</li><li>• <b>Playing "catch up":</b> More content generated if they don't log in for a while.</li></ul>

Document 8: META3047MDL-019-00106590, -6593

133. The above conclusions were based on Meta’s own survey research of ~2,500 Instagram users (evenly split between US and UK) ages 13-17.<sup>85</sup> Ultimately, Bhutada noted in the presentation that: “Young people are acutely aware that Instagram can be bad for their mental health yet are compelled to spend time on the app for fear of missing out on cultural and social trends.”<sup>86</sup> Similarly, internal Snapchat documents acknowledge that FoMo is a “negative” that “people report about Social Media”: “It’s easy to feel left out” and there is “Pressure to be always reachable.”<sup>87</sup> Snapchat’s qualitative research into parental perceptions revealed their concerns about the Snap Map feature in particular (which identifies for users where their friends are located):

<sup>85</sup> META3047MDL-003-00000029, -0031

<sup>86</sup> META3047MDL-003-00000029, -0053

<sup>87</sup> SNAP1924968 at -5012

“parents cited anecdotal evidence of their teens seeing that their friends were gathering together without them and feeling upset or experiencing FoMo (fear of missing out).”<sup>88</sup>

### **G. Internal Documents from Defendant Platforms Recognizing Problematic and Addictive Usage Among Users**

#### **i) Meta**

134. As early as 2017, Meta was being asked by external reporters about the problem of “internet addiction” as the excerpt from the email below indicates:

**From:** Josh Constine <[joshc@techcrunch.com](mailto:joshc@techcrunch.com)>  
**Date:** Monday, August 21, 2017 at 12:10 PM  
**To:** Lindsey Held <[lindseyh@fb.com](mailto:lindseyh@fb.com)>  
**Subject:** Re: Checking in

Hey Lindsey,

Thanks for getting this set up. Early to mid september should work. Off the record works to collect background, but would be good if we could try to isolate some things that could be used as quotes or that could later be clarified in quote form.

My main questions are:

- What research is Facebook doing about Internet addiction?
- What are the conclusions of this research in general, and in regards to healthy Facebook usage?
- Are there particular time quantities of usage or usage behaviors that indicate someone is using Facebook in ways related to Internet addiction?
- Has Facebook explored any product design changes or features to promote healthy usage and deter Internet addiction?
- What is the perspective of Facebook's leadership on Internet addiction: Is it a serious problem? Is it something Facebook exacerbates? Is it Facebook's responsibility to deter it or point people towards help with dealing with it?

*Document 9: META3047MDL-040-00317525 at -00317526*

135. The above questions asked by an external reporter are especially apt given that they were posed shortly after its inception when Meta could have taken actions to mitigate the growing problem of addiction, problematic use, and ensuing harms. Meta, however, took a different path

---

<sup>88</sup> SNAP0019128 at -9140

aimed at minimizing any concerns including those raised by items proposed by academic researchers as recognized in the document below:

Message

---

**From:** Lauren Scissors [/O=THEFACEBOOK/OU=EXTERNAL (FYDIBOHF25SPDLT)/CN=RECIPIENTS/CN=D8517B9F4F494001AD228E59940069A2]  
**Sent:** 8/30/2017 12:49:24 AM  
**To:** Lindsey Held [lindseyh@fb.com]; David Ginsberg [ginsberg@fb.com]  
**CC:** Daniel Harrison [dlh@fb.com]  
**Subject:** Re: Checking in

Re: internal research on addiction, the only recent work I know about was done by an intern on the data science team:

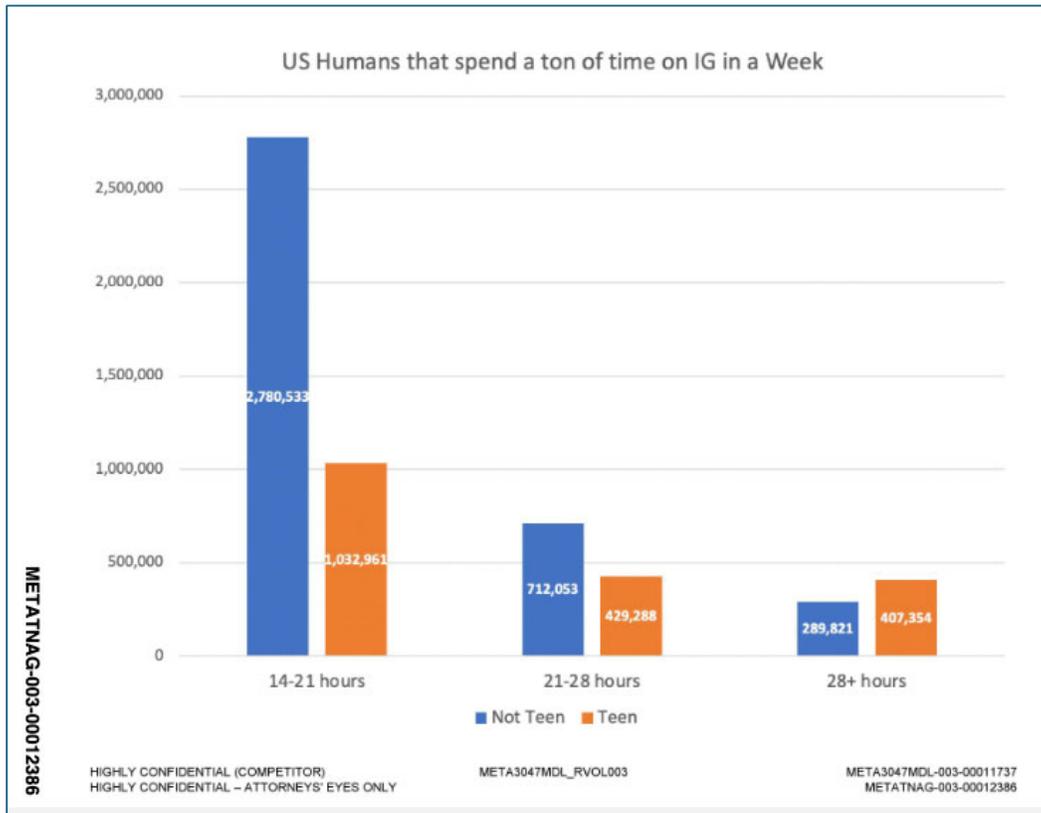
<https://www.facebook.com/notes/will-essilfie/have-we-made-people-addicted-to-facebook/1766984813316275/>

I would argue that this work picks up on addictive tendencies, not full-blown addiction, so huge caveat on the title. Comms asked us not to use survey items (that were drawn from external academic research) that asked about the more severe symptoms of addiction.

The other place to check would be PAC. I will reach out to Jennifer Guadagno to see if 1) there's any past internal research on this, 2) there's any ongoing projects on this, 3) we have existing relationships with academics who are studying this.

*Document 10: META3047MD-040-00317525*

136. Meta's internal documents reveal the company's awareness of a very large amount of time spent by certain users of its platforms. Internal studies of time on Instagram confirmed a "ton of time" being spent per user:



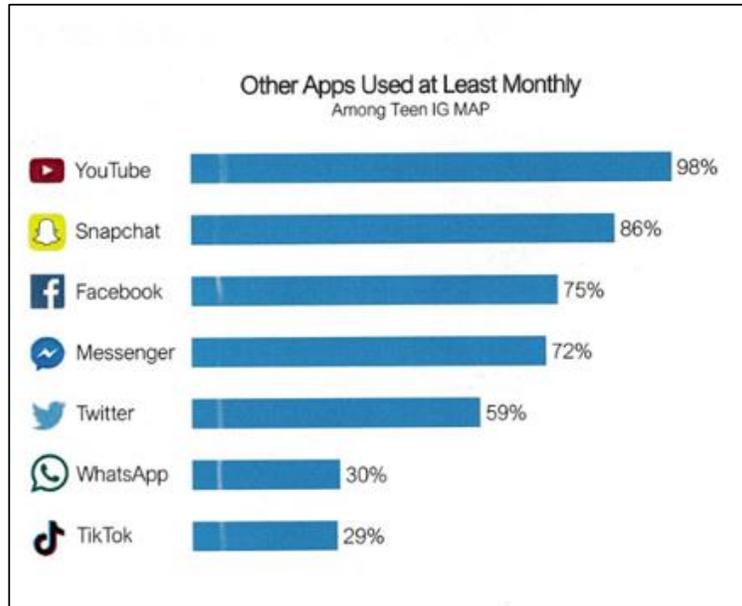
Document 11: META3047MDL-003-00011737

Another internal 2021 Meta document reported similar numbers: “On a given day, 0.1% of DAP spend >6.83 hours .... Within a given month, 0.1% of MAP spend >127.43 total hours, i.e. they average >4.55 hours per day for the entire month.”<sup>89</sup>

137. These data, directly measured by Meta, suggests that there are over 407,000 teens in the US who are spending more than 4 hours per day on Instagram alone. Keeping in mind that they should be awake for no more than 16 hours per day to have adequate sleep, they are either spending more than 25% of their waking hours on Instagram (including in many cases during school) or forgoing sleep to do so.

<sup>89</sup> META3047MDL-031-00048769, -8771

138. And of course, this is only time spent on Instagram and does not include additional time on other social media platforms. Meta’s commissioned study (~1000 users) from 2019 queried teens specifically about non-Instagram site usage and revealed that the vast majority regularly use other ones as well.



*Document 12: Wendy Gross Deposition Exhibit 6 at Slide 3*

139. Understandably, given these usage statistics, Meta was concerned about “problematic usage.” Internal documents reflect employees discussing “creating a world of addicted monsters”<sup>90</sup> and “making people’s mental health deteriorate slowly over time.”<sup>91</sup> In a 2017 internal Meta document, Matt Killingsworth states that “1-10% of college students exhibit a high-degree of Facebook addiction.”<sup>92</sup> While other Meta internal documents argue that “addiction” to Facebook has not yet been established, they concede that “there are parts of the addictive process

<sup>90</sup> META3047MDL-003-00011718, -1718

<sup>91</sup> META3047MDL-072-00304285 at -4288.

<sup>92</sup> META3047MDL-005-00000001, -0001

that may be at play and contributing to common issues for people.”<sup>93</sup> Relevant here, an internal Meta document concluded that “[a] large fraction of users struggle with their Facebook/Instagram use....A significant minority report serious difficulties.”<sup>94</sup>

140. One of the more insightful studies of “problematic usage” from within Meta was conducted in 2019 by Moira Burke, PhD, a Meta UX Research scientist, and her colleagues. Dr. Burke conducted a survey study of 20,000 US users and linked the responses to actual platform use. (Again, this is the kind of study that requires proprietary access to data and hence can only be done by or in collaboration with industry). They defined “problematic use” based on their review of the literature as follows:

We define problematic use as person who has:

- one or more **serious problems in life** they attribute to Facebook:
  - Facebook hurts their relationships “very much”,
  - they “very often” or “always” get less sleep because of Facebook,
  - Facebook hurts their work or school performance “greatly”, or
  - Facebook has a “very negative” impact on their lives
- AND one or more **concern about how they use FB**:
  - “very little or no control” over the time they spend on Facebook, or
  - are “very” or “extremely” concerned about missing posts from not logging in frequently enough (FOMO).

*Document 13: META3047MDL-020-00588361, -8363*

---

<sup>93</sup> META3047MDL-014-00359270, -9284

<sup>94</sup> Haugen\_00010114, -0120

141. Meta researchers also developed a taxonomy of different types of “Problematic use.”

Types of Problematic Use	
•	Sleep disruption - affecting people’s ability to get proper sleep
•	Relationships - affecting both online and offline relationships
•	Productivity - affecting productivity or one’s ability to be successful at work or school
•	Fear of missing out (FOMO) - affecting one’s mental state via fear, anxiety, or, intrusive thoughts about missing out if not checking Facebook enough
•	Control over time spent - aspects of the user experience, social norms of one’s network, or compulsive habits that make it difficult to regulate one’s duration of use or quality of use
•	Global negative life impact - an overall feeling that Facebook, on average, has a negative impact on one’s life

Document 14: META3047MDL-019-00106590, -6591

142. Based on this research, Meta estimated the prevalence and severity of problematic use. The results of two such studies are below:

	Sub-Area	Description	Prevalence
1	Problematic Use	Serious, negative impact on sleep, relationships, work, or lives, combined with lack of control over FB use.	3.1% severe*, 55% mild
2	Social Comparison	Evaluating yourself negatively to others whom you perceive as better off. Happens online and offline, though amplified on FB.	5% chronic, 40% mild
3	Loneliness	Difference between quality of friendships person has and wants. People who are already feeling lonely turn to FB to feel better, and reducing FB use can improve loneliness.	7% chronic, 36% mild

\* Represents US only, other stats are global FB MAP.

Document 15: Jennifer Guadagno Deposition Exhibit 29 at Slide 9

	Sub-Area	General Prevalence (extreme/moderate)	FB Associated Prevalence (extreme/moderate)	Priority
1	Problematic Use	na / 55%	12%/being estimated	H
2	Social Comparison	5%/ 40%	1.8%/4.3%	H
3	Loneliness	7%/37%	1.2%/6.8%	H
4	Measurement work (overall well-being)	n/a		H
5	Other areas (e.g. Mental Health, Conflict, MLEs, Social Support, etc.)	n/a		

*Document 16: Jennifer Guadagno Deposition Exhibit 32 at Slide 10*

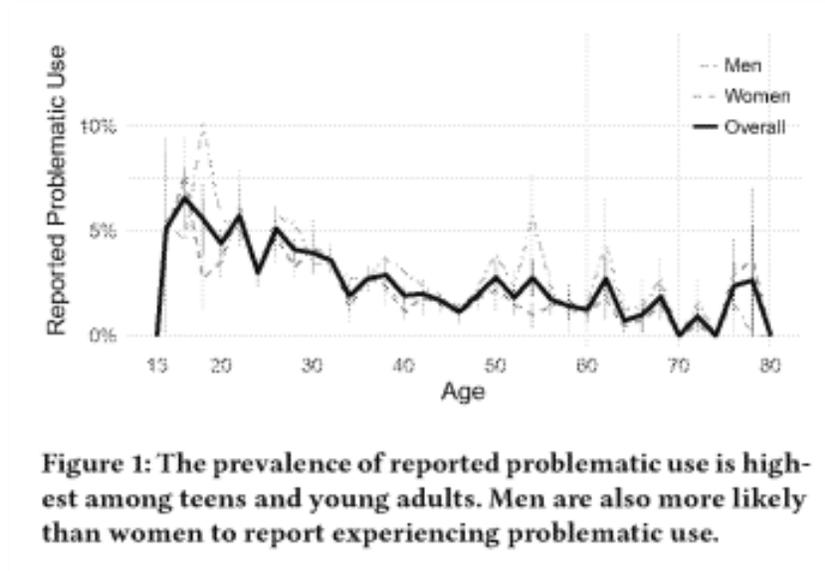
143. It is not clear from what I have seen how these prevalences were estimated and why they are divergent. It seems Document 15 is “general prevalence” whereas Document 16 includes Facebook prevalence. In her deposition, when queried about this discrepancy, Burke responded “This study presents a larger statistic because it’s a very different set of survey questions and it’s a different set of people that were asked. So this looks like it’s an international survey.”<sup>95</sup> When asked if this “other,” larger prevalence statistic was ever publicly disclosed by Meta, she responded “No.”

144. In either case, Meta’s studies found that between 3-12% of its users have “severe” or “extreme” “problematic use,” and that 55% have “mild” or “moderate” problematic use. This is consistent with the metanalytic global estimated range of 5-13%<sup>96</sup>

<sup>95</sup> Moira Burke Dep. Tr. at 175:16-21

<sup>96</sup> Cheng C, Lau Y, Chan L, Luk JW. Prevalence of social media addiction across 32 nations: Meta-analysis with subgroup analysis of classification schemes and cultural values. *Addictive Behaviors*. 2021/06/01/ 2021;117:106845. doi:<https://doi.org/10.1016/j.addbeh.2021.106845>

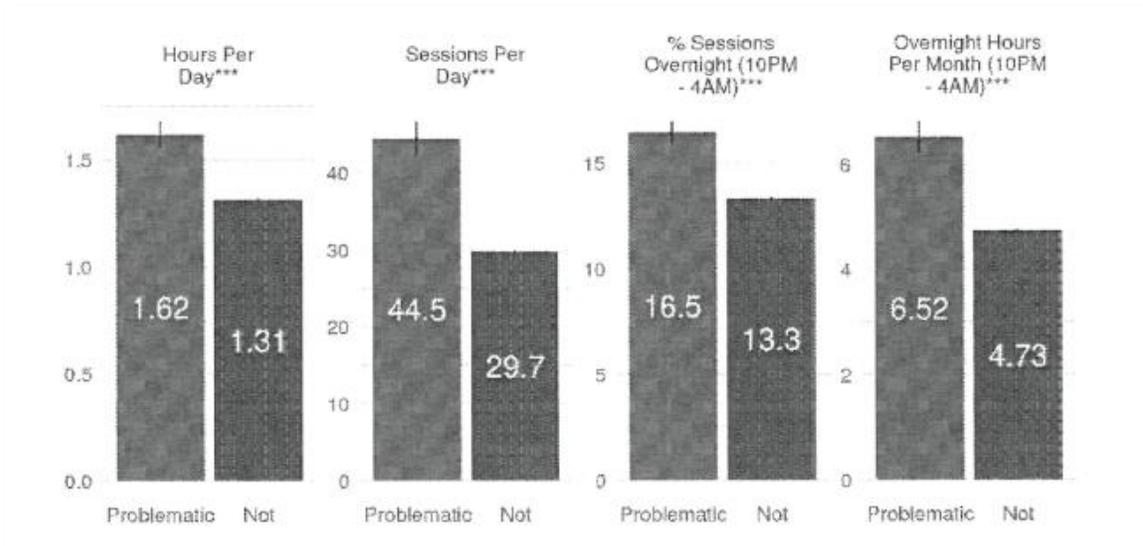
145. Importantly, Meta understood that the prevalence of “problematic use” per their data varies by age as shown in the below graph:



Document 17: META3047MDL-014-00046411, -6416

Consistent with the “increased vulnerability” of younger people, the prevalence is “highest among teens and young adults.”

146. The same study does what can be construed as some validation by examining differences in usage among those that have problematic usage versus those that do not. Those results are presented below:



*Document 18: META3047MDL-020-00588361, -8366*

Consistent with what one would expect, people with “problematic usage” have different usage patterns all favoring increased use. Most striking is the number of sessions per day (a difference of 15 sessions per day), and overnight use (a difference of almost 2 hours per month). During his deposition, former Meta data scientist George Volichenko credited the idea that late night usage constituted problematic use “More than four hours after midnight, I -- I feel like it's hard to argue that that's a problematic amount, right?”<sup>97</sup>

147. Unfortunately, documents I have reviewed indicate that Meta failed to adequately resource teams to address problematic usage and related wellness concerns, at least compared to the resources devoted to new “product” or “growth” objectives. In April 2017, then-head of Instagram Kevin Systrom asked for 13 additional engineering headcounts to make good on his “public commitment to making Instagram a place where people feel safe to be themselves, without

<sup>97</sup> George Volichenko Deposition Transcript at 81:1-4

criticism or harassment” and to address “critical areas for safety on IG.”<sup>98</sup> In response, Mr. Zuckerberg noted that he would add Instagram to a “mix” of other teams seeking access to a pool of unallocated engineers—but due to “more extreme issues on FB right now” “probably can’t get you 13 engineers in the near term.”<sup>99</sup> In an Instagram quarterly review several months later, Mr. Zuckerberg was told (perhaps not surprisingly given his decision) that “Instagram PAC [Protect and Care] is far behind FB PAC and we could become a major liability for FB Inc.”<sup>100</sup> Specifically: “We’re continuing to see an increase in high intensity abuse....However, the PAC team is only 22 engineers. We aren’t staffed to both integrate with FB and do IG specific work to stay ahead of all potential PAC related issues.”<sup>101</sup>

148. A couple of years later, in April 2019, David Ginsberg sent an email to Mr. Zuckerberg “requesting additional headcount to fund the ‘Well-Being 10x’ initiative;” specifically, 17 heads for Facebook and 7 heads for Instagram.<sup>102</sup> These resources were needed, Mr. Ginsberg explained, to “move quickly in the areas we have confidence in our understanding,” specifically “problematic use [g]iven its prevalence and our confidence around product interventions.” However, Mr. Ginsberg was informed by Meta’s CFO that the request “was not funded,” with Instagram’s CEO Adam Mosseri remarking, “I don’t see us funding this from Instagram any time soon.”<sup>103</sup> Separate documents confirm that the “skeleton crew” in wellbeing did not obtain the resources needed to tackle problematic use successfully: “We asked Mark for incremental HC to

---

<sup>98</sup> META3047MDL-014-00378084, -8085.

<sup>99</sup> META3047MDL-014-00378084, -8085.

<sup>100</sup> META3047MDL-050-00331333, -1334.

<sup>101</sup> META3047MDL-050-00331333, -1334; *see also* META3047MDL-050-00331327 (transmission of document to Mr. Zuckerberg)

<sup>102</sup> META3047MDL-003-00145472, -5472

<sup>103</sup> META3047MDL-003-00145472, -5472

fund it with more research/DS [data scientists] and a few eng [engineers] to tackle problematic use a few months ago but didn't get it.”<sup>104</sup>

149. Two more years after that, even Mr. Mosseri had to concede that the lack of resources to well-being efforts had become a problem. He acknowledged in a private message from October 2021 to another executive, “I’m really worried about this... We’ve been talking about this for a long time but have made little progress.”<sup>105</sup> And yet Mr. Zuckerberg still continued to deny resource requests. On November 10, 2021, Nick Clegg (a Meta executive) wrote Mr. Zuckerberg to “circl[e] back” on an earlier email seeking “investment needed to strengthen Meta’s position on well-being,” specifically an additional 25 cross-functional head count to form a “central well-being product pod” and noting “the increased urgency of all this.” Naomi Gleit responded to the chain, informing Mr. Zuckerberg, “Mark FWIW this is my #1 ‘below the line’ project to fund on Social Impact.”<sup>106</sup> Once again, the funding did not materialize.<sup>107</sup>

150. Consistent with all of these episodes, Dr. Allison Lee in her deposition acknowledged that there was only one team dedicated to integrity and 15 or more dedicated to Reels.<sup>108</sup>

151. In certain documents I have reviewed, Meta employees have been candid that safety tools rolled out by the company were built principally to address public relations problems, and that their efficacy was limited by the company’s desire not to curtail growth of their user base. As recently as 2018, problematic usage was viewed internally by some leaders at Meta not as a public

---

<sup>104</sup> META3047MDL-046-00477173 -7175

<sup>105</sup> META3047MDL-003-00171401, -1403

<sup>106</sup> META3047MDL-003-00190950, -0950

<sup>107</sup> Kang-Xing Jin Dep. Tr. at 471:23-473:10 (“no incremental headcount would be forthcoming for the work” and “the lack of headcount certainly impeded progress.”).

<sup>108</sup> Alison Lee Dep. Tr. at 163:16-22.

health issue but rather as a “public relations issue” for which a “pushback” strategy needed to be developed.<sup>109</sup> In a July 2018 email, ██████████ (Global Head of Brand and Product Marketing) acknowledged to other executives that safety tools were part of this “pushback”—and, as such, calibrated in a matter to prevent impact on the company’s bottom line: “We’re building the tools/controls to benefit policy or reputation outcomes... but in implementation we’re optimizing for business / growth outcomes. That’s a fine decision and I understand very clearly why we’re making it. That said we shouldn’t still expect to get sustained reputation ‘credit’ for these developments, which is what I do think Mark [Zuckerberg] and Chris [Cox] expect.”<sup>110</sup> An internal product audit from 2022 stated all this plainly: “[t]he stance we have historically taken is to give people control, but not in a way that hurts metrics.”<sup>111</sup>

152. In his deposition, George Volichenko recalled only “one” safety feature Meta initiated for teen problematic use while he was at the company—the “Take a Break” feature which, when turned on, would prompt users to step away from social media after a prespecified period of time (e.g. 10, 20, 30 min, etc.). Consistent with Meta optimizing these features to ensure minimal or zero impact on user engagement, Volichenko testified that Meta’s goal for adoption of this feature was reduced in the first half of 2022 from 0.6% to 0.25% of teen users activating the feature.<sup>112</sup> The actual number achieved was 0.18%. In other words, despite touting “Take a Break” as a way for users to set limits on the length and frequency of their sessions, Meta knew that 99.82% of teen users simply didn’t turn it on. Per the deposition, the reasons for this low usage rate appear to be because the feature was difficult to find and easy to ignore. In light of that, there

---

<sup>109</sup> See META3047MDL-003-00082165

<sup>110</sup> META3047MDL-040-00317980, -7980.

<sup>111</sup> META3047MDL-047-01167629, -7644

<sup>112</sup> George Volichenko Deposition Transcript at 98:2-102:4; *See also* George Volichenko Dep. Exhibit 3 at 1.

was consideration given to making it “opt out,” rather than “opt in,” meaning that the default would be to prompt teens to take a break at some pre-determined interval—but Volichenko testified that this was rejected because it would have affected core metrics negatively.<sup>113</sup>

153. Meta’s orientation towards problematic use as a public relations problem ultimately resulted in an aggressive effort by the company to push back on the *Wall Street Journal* after it leaked internal company documents indicating, among other findings, that “Facebook researchers have found that 1 in 8 of its users report engaging in compulsive use of social media.”<sup>114</sup> Notably, internal researchers did not share this same orientation, with Dr. Guadagno for one acknowledging the *Journal* presented an accurate and balanced story.

Jennifer L Guadagno (11/05/2021 12:51:41 PDT):  
>Hi Funda, wanted to pass along some feedback I'm feeling and getting from the team related to problematic use WSJ article and company response today.  
>  
>General feeling is that the WSJ article actually felt quite fair, and the response from us felt worse. The tone was defensive and felt like it overpowered the opportunity to really land the message that this is an issue we take seriously (and is an issue across tech spaces). The response also was confusing in places - specifically being clumsy with how addiction and problematic use were mentioned that adds more confusion, and part about "why would we do research that would potentially show higher levels of problematic use" section.

*Document 19: META3047MDL-040-00533279, -3279*

**ii) TikTok**

154. For its part, TikTok repeatedly asserts in its documents that their product induces compulsive, if not addictive, usage. As one missive succinctly states: “In sum, compulsive usage on TikTok is rampant and our users need better tools to understand their usage, manage it effectively, and ensure being on TikTok is time well spent.”<sup>115</sup> TikTok also appears to have understood that their especially young user base was at increased risk of experiencing compulsive

<sup>113</sup> George Volichenko Deposition Transcript at 114:5-116:5

<sup>114</sup> Jennifer Guadagno Deposition Exhibit 33 at 1

<sup>115</sup> TIKTOK3047MDL-002-00091634, -1636

usage.<sup>116</sup> “TikTok is particularly popular with younger users, who are particularly sensitive to reinforcement in the form of social reward and have minimal ability to self-regulate effectively.”<sup>117</sup> “Adolescents [are] more easily persuaded, there is a large population of them on TikTok (~30% DAU), and they likely don’t understand [the risks] of unhealthy usage as well as older users.”<sup>118</sup> A May 2022 internal TikTok presentation included survey data showing 35% of respondents thought TikTok LIVE was addictive.<sup>119</sup> A highlighted quote said, “It’s also very addictive and harmful to everyone, but especially seems to harm teenagers.”<sup>120</sup> Despite this, Maher testified at deposition that they could not recall ever informing users or their parents that TikTok was addictive and potentially harmful to pre-teens and teenagers.<sup>121</sup>

155. Given the draw of their platform and the demographics of their users, it is not surprising that people spend an inordinate amount of time on the app. TikTok’s internal data provide a level of granularity that would be the envy of any independent scientist who has labored to estimate the amount of time teens spent on any app.

---

<sup>116</sup> TIKTOK3047MDL-004-00318045, -8046 (“[Our expert partners tell us that [minors] are more likely to engage in and imitate DA because developmentally the part of the brain that is able to discern and make sensible decisions develops more slowly than the part that chases adrenaline haha.”).

<sup>117</sup> TIKTOK3047MDL-002-00091634, -1639.

<sup>118</sup> TIKTOK3047MDL-002-00091634, 1640.

<sup>119</sup> Maher dep. at 107:10–107:13; Exhibit 7

<sup>120</sup> Maher dep. at 108:25–109:3; Exhibits 10 and 11

<sup>121</sup> Maher dep. at 109:21–110:25; 175:17–176:1; 114:22–115; t 109:21-111:6; 111:1-12

Average time spent - Age groups [Age Gate Data]						
Age level	Avg Daily Active Users	Avg Daily Accumulated Duration (minutes)	50th percentile	66th percentile	75th percentile	90th percentile
L1 (13-15)	27,928,801	106.36	79.58	123.98	156.43	243.80
L2 (16-17)	41,622,675	106.78	81.25	124.17	155.63	241.10
<b>L12 (13-17)</b>	<b>69,551,476</b>	<b>106.61</b>	<b>80.60</b>	<b>124.10</b>	<b>155.95</b>	<b>242.18</b>
L3 (18-24)	205,994,552	96.12	69.00	109.17	139.50	224.17

Document 20:TIKTOK3047MDL-002-00098058, -8060

156. The median time 13–17-year-olds spend on TikTok is a little more than 1.3 hours per day. The 90<sup>th</sup> percentile for the same age cohort is more than 4 hours per day or more than 25% of their waking hours. Perhaps because of this considerable usage, TikTok conducted a survey in 2023 that revealed that 59% of teens feel they need a screen management tool.<sup>122</sup> In light of that, TikTok developed one that activates after 60 minutes of daily usage for teens 13-17 years of age. At that point, a prompt informs them that they have reached that “limit” and asks them if they want to continue. To do so, they must enter the “PIN,” which is preset to “1234” and universal and unalterable.<sup>123</sup> The logic behind the “PIN” is to provide some but not too much “friction” to bypassing the prompt. In fact, TikTok deliberately changed to this simple “1234” default setting

<sup>122</sup> TIKTOK3047MDL-010-00329585, -9594

<sup>123</sup> TIKTOK3047MDL-010-00329585, -9599

from a “custom” PIN so as to avoid the “frustration” or friction that comes from the “memory” problem of people forgetting their PIN and being blocked from the app until they reset it.<sup>124</sup>

157. TikTok internally acknowledges that their product is “addictive,” or that “compulsive use” is “rampant” as evinced below:

We have learned from *Project Who* that our users' biggest usage deterrent is that they think the platform is addictive. We also see many app store reviews that echo the following sentiment, "Do not download this app unless you're able to spent at least two hours a day on it. It's addicting!" (Appendix D).

This issue is further supported by external research and reveals other more concerning effects as well. According to a study of 1600 8-18 year olds, 8-12 year olds use almost 5 hours of entertainment screen media per day and teens use just over 7 hours per day, with 62% over 4 hours and 29% over 8 hours (*Common Sense Media*). This compulsive usage correlates with a slew of negative mental effects like loss of analytical skills, memory formation, contextual thinking, conversational depth, empathy, and increased anxiety (*Pew 2018*). Various similar studies to Pew's also conclude that compulsive usage interferes with essential personal responsibilities like sufficient sleep, work/school responsibilities, and connecting with loved ones (*Europe PMC*).

In sum, compulsive usage on TikTok is rampant and our users need better tools to understand their usage, manage it effectively, and ensure being on TikTok is time well spent (Interviews).

*Document 21:TIKTOK3047MDL-002-00091634, -1636*

Their own commissioned focus group study of teens fleshed out further qualitative details:

---

<sup>124</sup> TIKTOK3047MDL-010-00329585, -9600

TikTok is really engaging and they spontaneously declare they spend a lot of time on it, this usually involves scrolling videos from the For You Feed. The majority voices a sense of discomfort for being on TT too much, often with no awareness of the time they spent. This becomes a sort of estrangement from reality, falling into another dimension, well expressed in some teens' words (losing the dimension of time, addiction, being like "suspended").

Most of the teens say that "*too much TikTok*" puts them in a state of mental and physical stress. They report feeling stupid, guilty, lost, also isolated and sad. Sometimes they physically feel some drawbacks (eyes ache, head ache).

*Document 22: TIKTOK3047MDL-099-LARK-04759856, -9857*

158. With respect to time teens spent on Tik Tok, on December 15, 2022, Jordan Furlong (Digital Wellbeing Group Product Manager) queried in a group chat about the implications of setting a hard cap on minors' daily time on the app of 60, 90, or 120 minutes.<sup>125</sup> Isha Sha (Senior Data Scientist) immediately asked, "Can you please give more context as to why we want to aggressively curb minor's screen time?"<sup>126</sup> Josh Stickler (Director of Product Management) replied, "There is intensifying criticism at the highest levels of US and EU politicians about addiction-related harm among teens on TikTok."<sup>127</sup> After some back and forth to refine the query, data were shared to the group chat and are as follows:

---

<sup>125</sup> TIKTOK3047MDL-015-00342728

<sup>126</sup> TIKTOK3047MDL-015-00342728 at -2729.

<sup>127</sup> TIKTOK3047MDL-015-00342728 at -2729.

- On weekdays, the average daily duration for Minors is 106.7 minutes, 30% longer than adults
- 63% of minors stay longer than 60 mins, 35% longer than 120 mins
- On weekends, the average daily duration for Minors is 127 minutes, 40% longer than adults
- 68% of minors stay longer than 60 mins, 43% stay longer than 120 mins
- In the RoW areas:
- On weekdays, the average daily duration for Minors is 93 minutes, 10% longer than adults
- 53% of Minors stay longer than 60 mins, 29% stay longer than 120 mins
- On weekends, the average daily duration for Minors is 105 minutes, 15% longer than adults
- 56% of Minors stay longer than 60 mins, 34% stay longer than 120 mins

*Document 23:TIKTOK3047MDL-015-00342728, -2736*

Later in the chat, Stickler reports:

Josh Stickler 2022-12-23 01:24:54

Thanks!! Making sure I understand data and am applying it properly: if we instituted a 90 minute "hard" screen time cap in Europe for Minors (here defined as users with declared/self-reported 13-17 ages), average stay duration across the entire European user population would decrease from 83.71 -> 80.51 minutes, i.e. a 3.8% decrease

*Document 24:TIKTOK3047MDL-015-00342728, -2736*

159. Several things are notable in this exchange. First, average daily weekday usage is considerable (93-107 minutes). Further 29-43% of teens spend more than 120 minutes per weekday on TikTok, exceeding the 2016 *total* recreational daily screen time limit set by the American Academy of Pediatrics (guidelines I helped author).<sup>128</sup> Second, the motivation for exploring the deployment of the cap appears to be entirely reactive and intended to provide cover

---

<sup>128</sup> Reid Chassiakos YL, Radesky J, Christakis D, Moreno MA, Cross C. Children and Adolescents and Digital Media. *American Academy of Pediatrics*. Nov 2016;138(5)doi:10.1542/peds.2016-2593

for external political concerns about “addiction.” The initial push back from Shah confirms that there is no inherent interest (at least on her part) to curb teen usage. Third, the implications are framed entirely in terms of how much time spent on the platform would be reduced (3.8%) and hence how much ad revenue might decrease.

160. In the end, when TikTok did explore their “Screen Limit Management” tools, they set a “maximum 5% drop-in stay time” cap as shown below.

**How to define success?**

 The objective of this feature is to help users who want support managing their time spent on the platform. This is accomplished by helping users stay more aware of their time spent on the app in a single session.

We expect that screen time management features may reduce core metrics for cohorts that excessively use the platform or who are self-disciplined, but will increase long term retention since wellbeing features make TikTok a healthier and more sustainable experience for our users.

Our guardrails include (1) a maximum 5% drop in stay time for target user groups like minors and excessive users (2) retention.

*Document 25: TIKTOK3047MDL-001-00004654, -4659*

After discussing these potential tradeoffs with Wenjia, he proposed that we can accept a 5% drop in stay time for Screen Time Management features for special user groups like minors and excessive users. This should however not come at the expense of retention. That said, we don't expect significant impact to stay time with this feature since it is only improving awareness and is not an intervention.

*Document 26: TIKTOK3047MDL-078-LARK-01711316, -1322*

Consistent with their corporate strategic approach, the new screen time management system was subjected to a rigorous A/B test protocol.

**iii) Snap**

161. Snap documents reflect that many users found Snapchat to be addictive and harmful to their mental health. In 2013, just two years after Snap was founded, an email to Evan Spiegel—Snap’s founder and CEO—observed that kids “that have the snapchat addiction have no room for anything else. Snaps dominate their life.”<sup>129</sup> The email goes on to say, “Lucky for Snapchat that England is the home of Europe’s best boarding schools. Kids from the rest of the world here [sic] about snapchat from a friend that is in an English boarding school.”<sup>130</sup> In this exchange, the “Snapchat addiction” is a product selling point.

162. Subsequent user interviews conducted by Snap bore out this early observation about addiction, showing that some children were opening the app hundreds of times a day.<sup>131</sup> But to Snap, this behavior remained a positive, with Snap characterizing those in the 90<sup>th</sup> percentile of time spent on Snapchat as “elite.”<sup>132</sup> “Elite” users were also disproportionately young and female.<sup>133</sup> Among this group, the median amount of time spent was 100 minutes per day.<sup>134</sup> Over the course of a month, that adds up to 50 hours, the equivalent of these users spending a fulltime workweek every month just on Snapchat.

---

<sup>129</sup> SNAP2324154, -4154.

<sup>130</sup> SNAP2324154, -4154.

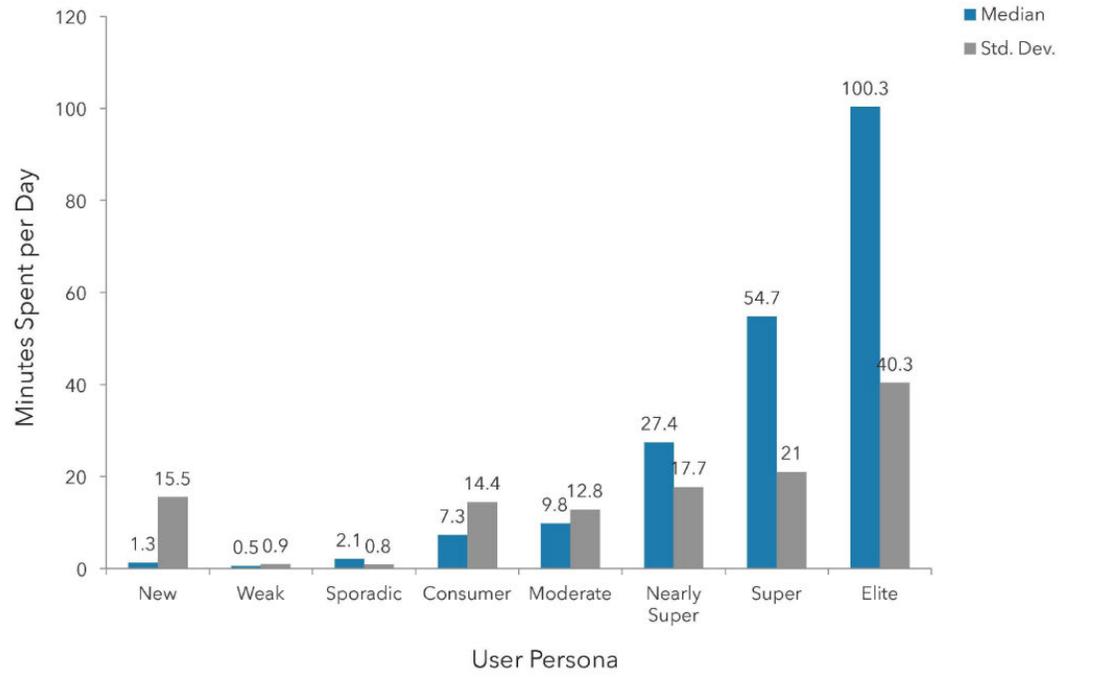
<sup>131</sup> SNAP2372970 at -2971.

<sup>132</sup> SNAP3121196, -1205.

<sup>133</sup> SNAP3121196, -1197.

<sup>134</sup> SNAP3121196, -1221

## Elite Snappers spend a median of 1.6 Hours per day in-app



Document 27: SNAP3121196, -1221

163. Snap also recognized compulsive use in the context of Snapstreaks. Snapstreaks occur when friends “Snap back and forth with each other at least once a day, every day,” at which point the platform rewards the users with various trophies such as a fire emoji on their Chat screen.<sup>135</sup> In 2018, Evan Spiegel, Snap’s CEO, referred to streaks as “toxic behavior” that Snap shouldn’t reinforce.<sup>136</sup> Indeed, Snap employees often recognized the addictive nature of Streaks

<sup>135</sup> Snapchat Support, *How do Snapstreaks work and when do they expire?*, <https://help.snapchat.com/hc/en-us/articles/7012394193684-How-do-Snapstreaks-work-and-when-do-they-expire> (last accessed Apr. 14, 2025)

<sup>136</sup> SNAP7140925, -0925

in the same breath that emphasized how important streaks are to Snap’s business model, such as this email from Josh Siegel, a senior Snap product manager:

The general product stance on Streaks is that we don't love them (it was an accidentally addictive, somewhat unhealthy feature that gamifies friendships in a weird way), but they're too delicate to touch right now. 50M+ users have streaks, a few million probably only use the app for streaks. We definitely don't want to invest in or exacerbate the problem right now, but I'm hoping it's something we can get to in 2019 via other ways to encourage daily behaviors.

*Document 28: SNAP4389271, -9271*

164. Snap’s Spotlight feature is another attribute of the platform that contributes to addictive use. Spotlight showcases viral videos to a broad audience (even to individuals who are not connected to the poster). “Spotlight” was in addition to their “Discover” tab which allowed people to search for specific content of interest to them. Nona Yadegar (Director of Public Policy) communicated with Evan Geary (Head of Platform Policy) about the introduction of Spotlight as follows:

**From:** Nona Farahnik Yadegar [REDACTED]  
**Sent:** 10/13/2020 7:48:01 PM  
**To:** [REDACTED]  
**Subject:** Re: Spotlight Q

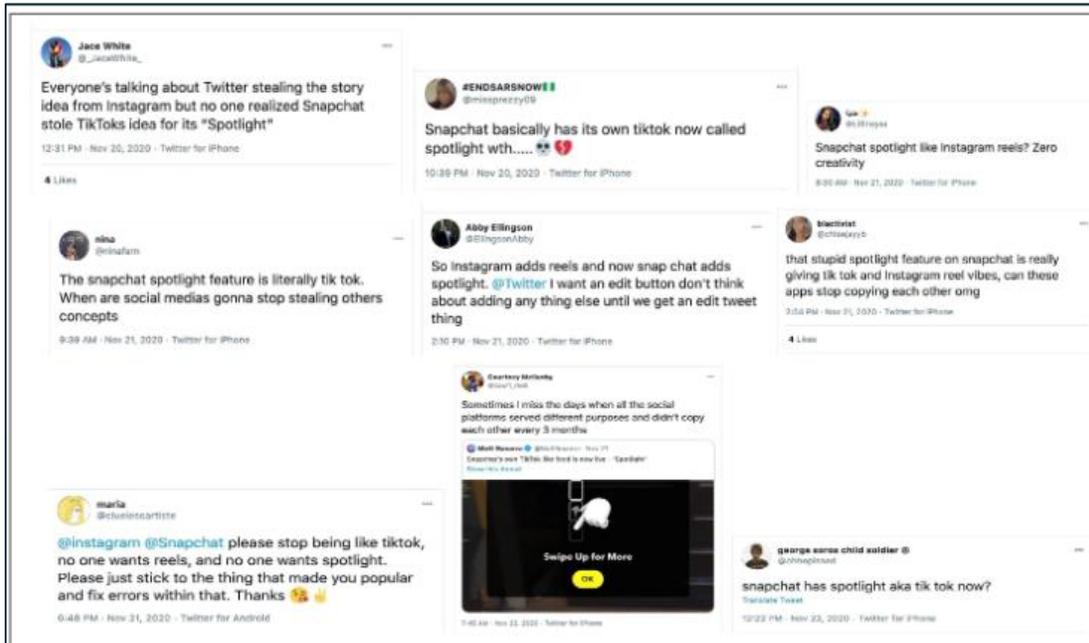
lol ur emails are so funny

On Tue, Oct 13, 2020 at 5:17 PM [REDACTED] wrote:  
Worth noting that all Spotlight content will be pre-reviewed by humans (TBD on how long that lasts).  
Not sure what to say about addictive endless scrolling. We already have an endless scroll design in Discover and I think we wish it was more ~~addictive~~ compelling.

*Document 29: SNAP1393050, -3050*

165. Making light of a serious concern, [REDACTED] presents Spotlight’s endless scroll design as inspired by Discover (an earlier feature), noting for humorous or perhaps ironic effect that the company wishes that design were more “compelling” (striking out “addictive”). But once Spotlight debuted, many saw it for what it was, a blatant knockoff of the highly addictive “TikTok

For You” feed (and the very similar, itself-derivative Instagram Reels feature).<sup>137</sup> And there was considerable “blow back” presented within Snap documents.



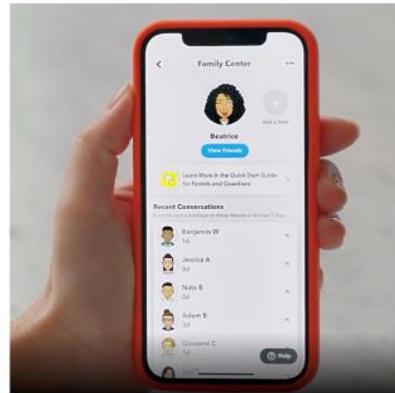
Document 30: SNAP0188592, -8613

166. Recognizing how addictive its product could be, Snap explored and implemented some remediation strategies. In 2022, later than some of its competitors, Snap launched its “Family Center,” designed to give parents more control over their teen’s usage of the platform. But their focus group research on this proved equivocal at best:

<sup>137</sup> SNAP0188592, -8605-06 (“A lot of chatter about all platforms copying each other, mainly TikTok...”)

**We showed Snapchat's Family Center to two parent focus groups. Overall, they view it as "a step in the right direction" but one that "feels like the bare minimum"**

- Most had not heard of Family Center
- The biggest dislike is they wouldn't be able to see the actual content of the conversations – "don't really know how helpful this would be to me"
- They do like the alert to teens that parents can see the account and the mirrored view of what parents see
- Parents feel kids will figure out a way around Family Center to hide things from their parents (i.e., create fake or private account parents can't see)



HIGHLY CONFIDENTIAL (COMPETITOR)

SNAP0404304

*Document 31: SNAP0404262, -4304*

Overall, the response was similar to that of other social media parental controls: Cumbersome to find (or use), limited information, and too easy for teens to circumvent.

167. Another remediation strategy that Snap considered but did not deploy was turning off notifications during certain times such as school hours. At the time of my report, this option is not available to users.

For teens, Snapchat is addictive and they are on it "all the time"	➔	Allow users to turn off notifications during school hours or when they should be working or studying. Let users set their own time limits within the app
--	---	--

*Document 32: SNAP0404262, -4288*

It is particularly notable that this fix was not implemented given that it was identified as a solution to the problem that "Snapchat is addictive and [teens] are on it 'all the time.'"

iv) **YouTube**

168. YouTube documents reflect a similar continuum of problematic usage including addiction. A 2019 Google report found that 45% of survey respondents "unintentionally stay on

YT longer than they want,” a core feature of “addiction.” The “insight” accompanying that finding was that “YouTube is designed around increasing users’ engagement, not maintaining user’s intention.”<sup>138</sup> Promoting “engagement” at the expense of “intention” is a very effective strategy to build an addictive product. That same 2019 report estimated that 5% of 13–24-year-olds watch 3 hour per day “habitually” and 1-2% watch 4 or more.<sup>139</sup>

169. Another survey found that 60% of survey respondents “go down the rabbit hole” unintentionally on YT, using it longer than desired.<sup>140</sup> The same report noted that “20% of teens and 30% of YA who watch weekday evenings don’t leave YT until after midnight.”<sup>141</sup>

### Excessive video watching is related to addiction

- **Watching short videos results in a “quick fix” of dopamine**
  - Dopamine is related to feelings of reward
  - Similar to feelings of reward when using drugs or other addictive substances
- **Researchers feel that YT is built with the intention of being addictive**
  - Designed with tricks to encourage binge-watching (i.e. autoplay, recommendations, etc).
  - These “tricks” have become routine
  - Technology & well-being need to meet

```

graph TD
    Guilt --> EmotionalTrigger[Emotional Trigger]
    EmotionalTrigger --> Craving
    Craving --> Ritual
    Ritual --> Using
    Using --> Guilt
    
```

Addiction Cycle

(Howard, 2012; Gunantillake, 2017)

*Document 33: GOOG-3047MDL-04918852 at Slide 12*

<sup>138</sup> GOOG-3047MDL-00236723 at Slide 11

<sup>139</sup> GOOG-3047MDL-00236723 at Slide 41

<sup>140</sup> Ex. 37 at 16 to Cristos Goodrow Deposition (2-20-25)

<sup>141</sup> Ex. 37 at 12 of Cristos Goodrow Deposition (2-20-25)

170. As Cristos Goodrow explained in 2011, “all other things being equal, our goal is to increase (video) watch time.... essentially ignoring the initial intention the user had when coming to YouTube if it helps to increase entertainment (measured via watch time).”<sup>142</sup> “The objective of every view should be to drive the most long-term engagement with minimal user effort.”<sup>143</sup>

171. At that time, Google estimated that there were 69 million teens on its platform meaning that over 4 million teens were watching more than 3 hours per day. That amount of video consumption per day should be viewed as “problematic” to say the least and YouTube acknowledged as much, proposing “targeted alerts and active education depending on the pattern of use.” In 2018, a YouTube presentation stated that “gaming content is popular on their platform” and that if “DSM criteria were applied to watching gaming videos, 1 in 5 teens would be diagnosed with addiction.”<sup>144</sup>

172. YouTube has a “digital wellbeing” group that among other things offered “take a break” and “bedtime” reminders. YouTube first launched an opt-in “take-a-break” reminder and “time watched profile” in 2018, which were described as “needed quick wins” that garnered “positive press”.<sup>145</sup> YouTube then launched an opt-in “bedtime reminder” in 2020.<sup>146</sup>

173. It was not until 2021 that YouTube turned these “take-a-break” and bedtime reminders “on by default,” for declared teen accounts.<sup>147</sup>

174. YouTube employee Reid Watson, a Group Product Manager, testified that he didn’t believe the default setting was applied to teen users who access YouTube in a “signed out” state or

---

<sup>142</sup> GOOG-3047MDL-02185032, -5098

<sup>143</sup> GOOG-3047MDL-02001804 at -012001809

<sup>144</sup> GOOG-3047MDL-00874191 at Slide 27; GOOG-3047MDL-00937887 at 7914

<sup>145</sup> GOOG-3047MDL-00937887 at 7914; James Beser 30(b)(6) Dep. Exhibit 1.

<sup>146</sup> James Beser 30(b)(6) Dep. Exhibit 1.

<sup>147</sup> James Beser 30(b)(6) Dep. Exhibit 1.

who misrepresented their age when creating an account.<sup>148</sup> In addition, these features, which YouTube framed as “digital wellbeing tools” did not function on YouTube Shorts until August or September of 2022.<sup>149</sup>

175. Although YouTube touted the importance of its digital wellbeing tools, I have not seen evidence that YouTube measured the effectiveness of these tools to determine whether it was having a meaningful impact on addiction and overuse.

176. On January 30, 2023, Erin Turner, YouTube Group Product Manager queried what “the success of those [reminders]” were, and the responsive data did not identify any meaningful metric for *effectiveness* and, tellingly, honed in on whether the features would “break anything.” The actual numbers are not provided in the document, rather a list of links to view them is:

- # of users who see a take a break reminder: [http://shortn/\\_2pMkNjFDgP](http://shortn/_2pMkNjFDgP)
- # of users who see a bedtime reminder: [http://shortn/\\_7jZd1GJYTw](http://shortn/_7jZd1GJYTw)
- # of users who check their time watched stats: [http://shortn/\\_kaPh1a1Y4h](http://shortn/_kaPh1a1Y4h)
- We also check topline YTT metrics when rolling out new features to make sure we didn't break anything. They're expected to be neutral though.

*Document 34: GOOG-3047MDL-02486605, 6605*

177. Rather than track whether these tools were effective at changing a user’s behavior, YouTube decided to track only the number of users who received a take a break or bedtime reminder or who checked their time watched statistics, what I would deem process measures, not outcomes.<sup>150</sup>

178. Not that tracking outcome data would be difficult: YouTube had “all the data in the logs to determine if people actually stop[ped] watching [YouTube]” in response to its take-a-break

---

<sup>148</sup> Reid Watson Dep. Tr. at 24, 250-251

<sup>149</sup> Reid Watson Dep. Exhibit 21.

<sup>150</sup> Cristos Goodrow Dep. Exhibit 37 at 31; Cristos Goodrow Dep. Tr. at 457:25-458:23; Reid Watson Dep. Tr. at 156:3-20; 254:3-255:34; Reid Watson Dep. Exhibit 23

reminder.<sup>151</sup> In fact, in monitoring the impact of the implemented safeguards, YouTube anticipated that their effects would be “expected to be neutral” on overall viewing time. That is a bit counterintuitive. An intervention designed to induce breaks, or set nighttime stop times, should, if effective in any meaningful measure, *reduce* total time on the platform—suggesting that, like their competitors’ similar time-limiting features, these were engineered to be *ineffective* in practice.

179. A 2019 YouTube presentation discussing the success of autoplay also acknowledged “policymaker scrutiny on digital addiction and well-being” and recommendations to ban “autoplay.”<sup>152</sup> The same 2019 presentation noted that “autoplay may create digital wellbeing issues for some users.”<sup>153</sup> Despite this recognition of harm in 2019, and as noted earlier in this report, YouTube waited until 2021 to turn Autoplay “off by default” for teens.<sup>154</sup>

180. Despite this “off by default” change, a 2022 presentation recognized that autoplay contributed to binge use of YouTube.<sup>155</sup> In 2023, YouTube conducted interviews with “professionals, academics, & industry experts.”<sup>156</sup> One of the expert recommendations included “removing autoplay or introducing stopping cues” which would “go a long way to helping people.”<sup>157</sup>

181. Indeed, as the foregoing shows, while internally recognizing and conceding that their product had addictive design features driving some percentage of its clients to develop problematic usage patterns, each of the four platforms took minimal if any steps to mitigate those

---

<sup>151</sup> Reid Watson Dep. Exhibit 25

<sup>152</sup> GOOG-3047MDL-04652560 at -04652570

<sup>153</sup> GOOG-3047MDL-04652560 at -04652575

<sup>154</sup> James Beser Dep. Exhibit 1

<sup>155</sup> GOOG-3047MDL-00874191 at \*7

<sup>156</sup> GOOG-3047MDL-04805860

<sup>157</sup> GOOG-3047MDL-04805860 at Slide 15

features. And even insofar as they did deploy mitigation strategies, they were always evaluated in the context of how they might affect the company's bottom line currently and in the future.

#### **IX. Design Features of Social Media that Drive Usage and Addictive Behavior**

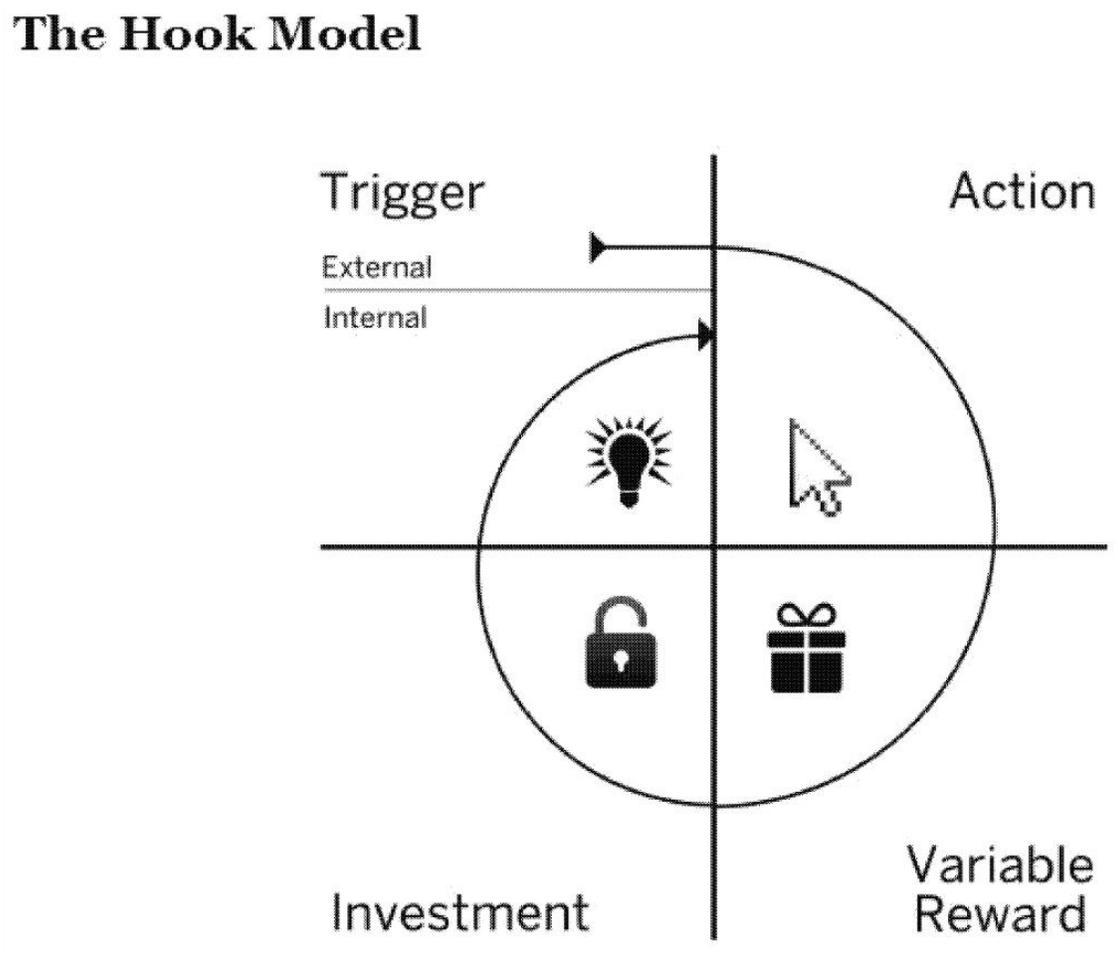
182. Social Media platforms embody numerous design features that promote addiction, problematic use, and attendant harms. This includes design features that exploit intermittent reward, social comparison metrics, and "flow" state. I will discuss each of these in turn, but they all work in tandem to keep users hooked and have their foundations in behavioral psychology, the "father" of which is BF Skinner.

183. Skinner (1904–1990) was an American psychologist, and the Edgar Pierce Professor of Psychology at Harvard University until 1974. He conducted foundational research related to how behaviors are reinforced through the "Skinner" box he invented. Briefly, rats were placed in a box that had a light, a loudspeaker, a response lever, a food dispenser and an electrified grid. The lights and the loudspeaker provided stimuli. These conditioned the rat to pay attention and press the lever when they were activated. Pressing the response lever would in turn dispense (or not dispense) food. The electrified grid could be used to "punish" failure to press the response lever when a stimulus had been delivered.

184. Skinner found that rewarding reaction to the stimulus, by dispensing food if the rat pushed the lever when the light flashed, made the rat more attentive to the light. Conversely, "punishing" the rat for failing to respond to the stimulus by delivering a low-level shock via the grid also made it more attentive. Neither of these results are especially surprising (at least not now) to us. But the most interesting of Skinner's findings were that intermittent unpredictable rewards were more effective at inducing the rat to stay focused on the light than predictable ones. In other words, the rat became more attentive to a stimulus if not every reaction yielded a reward.

This unexpected finding was long ago incorporated into such things as slot machines, where gamblers cannot predict which pull will result in payouts, and as they play they are constantly hearing others around them win, reminding them that their next pull might be a jackpot. All three of these addictive design features are “built in” to social media sites and are reflected in the “hook model” studied by Meta’s researchers (*see below*).

## The Hook Model



*Document 35: META3047MDL-020-00342155, -2155*

185. This figure, which is taken from a book by Nir Eyal titled *Hooked: How to Build Habit Forming Products*<sup>158</sup> also appears in Snap’s documents.<sup>159</sup> The figure is annotated as follows:

**Internal Trigger:** users need a reason to come back to Snapchat (vs going to another social app). If we can get things like the actions/UI and rewards right, the internal triggers become things like FOMO, wanting to get the news, and staying in touch with all friends. Additionally, internal triggers should play on what users are already doing - things like community meetups (e.g. school network), affinity groups (e.g. run club), and family engagement (e.g. sharing photo albums).

Document 36: SNAP5486213, SNAP5486214

186. The *Handbook of Children and Screens*, of which I was the editor, examined several features that promote problematic usage or addictive usage of the platforms. Below, I will summarize some of our findings.

**A. Likes, Comments, and other Metrics**

187. Scientists have found that receiving “likes” on social media platforms are very similar to the “rewards” that researchers have associated with addiction research for decades.<sup>160</sup> Likes can include the like button on Facebook but can also include “hearts” on Instagram and TikTok. Similar metrics include the number of shares on a post, number of comments, and number of followers or friends for a user. While seemingly innocuous, users’ quest for these publicly-visible “rewards” has been linked to a number of secondary harms such as reduced sleep efficiency

---

<sup>158</sup> Eyal N. *Hooked: how to build habit-forming products*. Portfolio/Penguin; 2014:242 pages.

<sup>159</sup> SNAP5486213, -6214

<sup>160</sup> Handbook of Children and Screens at pdf p. 59, 153; See also Sherman LE, Payton AA, Hernandez LM, Greenfeld PM, Dapretto M, *The power of the like in adolescence: effects of peer influence on neural and behavioral responses to social Media*, Psychol Sci. 2016;27(7):1027–35. <https://doi.org/10.1177/0956797616645673>; Sherman LE, Payton AA, Hernandez LM, Greenfeld PM, Dapretto M, *The power of the like in adolescence: effects of peer influence on neural and behavioral responses to social media*, Psychol Sci. 2016;27(7):1027–35, <https://doi.org/10.1177/0956797616645673>.

and duration due to “routine check[ing]” behaviors during the night.<sup>161</sup> They have also been associated with reinforcing addictive behaviors in order to encourage users to spend more time on these platforms.<sup>162</sup> They have even been linked to increased feelings of depression, anxiety, and negative social comparison since receiving “fewer likes” is viewed as a form of “negative peer feedback.”<sup>163</sup>

188. Comments are similar to likes in that they provide users with “quantifiable (and qualitative) feedback” about their experiences on the platform.<sup>164</sup> For teens, this means that they are able to quantify the “success” of their posts with many teens reporting that they post “self-oriented images on social media with the goal of obtaining likes and other forms of feedback such as comments.”<sup>165</sup> Other studies have found that comments are often seen as a mechanism for “gaining [social] status” and reflect a degree of “digital social approval.”<sup>166</sup> As a result, other

---

<sup>161</sup> Handbook of Children and Screens at pdf p. 134; *See also* Rod NH, Dissing AS, Clark A, Gerds TA, Lund R, *Overnight smartphone use: a new public health challenge? A novel study design based on high-resolution smartphone data*, PLoS One. 2018;13(10):e0204811, <https://doi.org/10.1371/journal.pone.0204811>.

<sup>162</sup> Handbook of Children and Screens at pdf p. 210; *See also* Brand M, *Can internet use become addictive?*, Science. 2022;376:798–9, <https://doi.org/10.1126/science.abn4189>; Starcevic V, Aboujaoude E., *Internet addiction: reappraisal of an increasingly inadequate concept*, CNS Spectr. 2017;22(1):7–13.

<sup>163</sup> Handbook of Children and Screens at pdf p. 248; *See also* Lee HY, Jamieson JP, Reis HT, et al, *Getting fewer “likes” than others on social media elicits emotional distress among victimized adolescents*, Child Dev. 2020;91(6):2141–59, <https://doi.org/10.1111/cdev.13422>.

<sup>164</sup> Handbook of Children and Screens at pdf p. 445.

<sup>165</sup> Handbook of Children and Screens at pdf p. 447; *See also* Chua THH, Chang L, *Follow me and like my beautiful selves: Singapore teenage girls’ engagement in self-presentation and peer comparison on social media*, Comput Hum Behav. 2016;55:190–7, <https://doi.org/10.1016/j.chb.2015.09.011.23>; Yau JC, Reich SM, *“It’s just a lot of work”: adolescents’ self-presentation norms and practices on Facebook and Instagram*, J Res Adolesc. 2019;29(1):196–209. <https://doi.org/10.1111/jora.12376>.

<sup>166</sup> Handbook of Children and Screens at pdf p. 448; *See also* Chua THH, Chang L, *Follow me and like my beautiful selves: Singapore teenage girls’ engagement in self-presentation and peer comparison on social media*, Comput Hum Behav. 2016;55:190–7, <https://doi.org/10.1016/j.chb.2015.09.011.23>.

studies have linked comments to negative social comparison since “not receiving enough likes on one’s pictures can negatively affect appearance esteem and prompt delet[ion] of a post.”<sup>167</sup>

## **B. Algorithmic Recommendations**

189. At the heart of every social media platform’s engagement strategy is their proprietary recommendation algorithms. These algorithms are optimized to maximize time on the platform rather than healthy interactions with a person’s social network.<sup>168</sup> They accomplish this in a variety of ways. One way they drive time spent on social media is by inferring the interests of the user (which may or may not be expressed by the user in any direct way) and feeding them an aggregation of posts that, while most likely to keep them engaged, may lead the user down problematic rabbit holes.<sup>169</sup>

---

<sup>167</sup> Handbook of Children and Screens at pdf p. 448; *See also* Baker N, Ferszt G, Breines JG, *A qualitative study exploring female college students’ Instagram use and body image*, *Cyberpsychol Behav Soc Netw.* 2019;22(4):277–82, <https://doi.org/10.1089/cyber.2018.0420>.

<sup>168</sup> Tom Cunningham Dep. Tr. at 50:11-19 (“[T]here’s now in the public domain two or three well-run experiments which show that comparing between chronological ranking where you rank items by the point at which they were posted versus ranking by engagement, the rank by engagement tends to increase a lot of metrics of user retention and time spent and almost by definition viewport views.”); Josh Simons Rough Dep. Tr. at 67:12-21 (“[G]iven that the whole News Feed system is aimed at engagement, that’s a set of repeated patterns of behavior – clicking, liking, sharing and on so – the best way to increase engagement is addiction in some way, you know, is to get people doing those forms of engagement. And so the type of content that the News Feed system made viral was the kind of content that would make you feel, make users feel whatever they needed to feel to produce that type of behavior.”).

<sup>169</sup> Tom Cunningham Dep. Tr. at 29:17-30:4 (“Q. What other types of engagement did – to the extent you can recall – did feed take account of? . . . A. Reshares. Exactly. And video views. And what’s the right noun? Lingers. The time that someone was lingering or watching a – or examining a post.”); Joshua Simons Dep. Tr. at 81:19-25 (“[W]hat the engineers building [News Feed] reported to me is that by defining value in terms of those proxies – clicks, likes, shares, plays – you ended up incentivizing repeated patterns of behavior on the tool that were in the end undesirable, and sometimes actively harmful for the users of the tool themselves.”); Joshua Simons Dep. Tr. at 92:1-6 (“The fact that Facebook’s models are all trained to predict a proxy for what we really care about was understood by engineers inside the company to be one of the likely drivers for things like divisiveness and filter bubbles that were being actively researched inside the company at the time.”).

190. In regard to TikTok, these problematic rabbit holes were also known as “negative filter bubbles.”<sup>170</sup> A 2021 TikTok presentation entitled “Negative Filter Bubble Project Sharing” noted that “Algorithms lead[] people into a feed filled with continuous similar negative videos that are hard to escape.”<sup>171</sup> What’s more, another internal document notes that “**minor creators [fall] into higher density of negative comment filter bubbles.**”<sup>172</sup> To their credit, TikTok did set as a 2023 safety goal to “be more transparent to the public and trusted by teen users and parents,” in addition to building “industrial leading teens safety and well-being features.”<sup>173</sup> As discussed below, these aspirations of transparency and improved safety features that were noted approximately six years after TikTok launched in the United States were too little and too late for many children and teens.

191. Another way recommendation algorithms drive time spent on social media is by utilizing the intermittent variable reward mechanism that Skinner discovered (*e.g.*, likes, notifications, comments that appear unpredictably) which contribute to addiction to the platforms themselves.<sup>174</sup>

192. The addictive power of an engagement-optimized algorithm is recognized within the medical and academic literature. In the chapter regarding “Problematic Internet Use,” the authors’ consensus was that “Attention focused designs intended to generate, or possibly exploit,

---

<sup>170</sup> TIKTOK3047MDL-004-00294514.

<sup>171</sup> TIKTOK3047MDL-004-00294514 at -9518.

<sup>172</sup> TIKTOK3047MDL-060-01143638 at -3640 (bold in original).

<sup>173</sup> TIKTOK3047MDL-060-01143638 at -3645.

<sup>174</sup> Social Media and Youth Mental Health, The U.S. Surgeon General’s Office, available at <https://www.hhs.gov/surgeongeneral/reports-and-publications/youth-mental-health/social-media/index.html> (2023) (“According to one recent model, nearly a third (31%) of social media use may be attributable to self-control challenges magnified by habit formation.”); *See also* Allcott, H., Gentzkow, M., & Song, L. (2022). *Digital Addiction*, AMERICAN ECONOMIC REVIEW, 112 (7): 2424-63. <https://doi.org/10.1257/aer.20210867>

potentially addictive features (e.g., “likes”) and conditioned responses (e.g., notifications) alongside powerful algorithm-based technologies may lead youth to stay online longer than either intended or recommended.”<sup>175</sup>

### C. Auto-Play

193. Academics recognize that autoplay is a feature that is designed to prolong engagement at the cost of displacing “important developmental opportunities for young children” and is even associated with greater child behavioral difficulties.”<sup>176</sup> It has also been linked to difficulties controlling device use more broadly.<sup>177</sup> The latter phenomenon has even been observed in Defendants’ own research. For example, Meta’s researchers learned that clinicians regarded autoplay as not having any “beneficial” role while “detract[ing] from patients’ ability to control amount of time spent” using their platforms.<sup>178</sup>

### D. Infinite Scroll

194. Like autoplay, the Infinite or Endless Scroll feature has also been linked to prolonging engagement at the cost of displacing children’s developmental opportunities.<sup>179</sup> This feature is also linked to behaviors that create a user-sided “time distortion” that results in users

---

<sup>175</sup> Handbook on children and screens p.182

<sup>176</sup> Handbook of Children and Screens at pdf p. 54; *See also* Munzer T, Torres C, Domoff SE, et al. *Child media use during COVID-19: associations with contextual and social-emotional factors*, J Dev Behav Pediatr. 2022; 43(9):e573, <https://doi.org/10.1097/DBP.0000000000001125>.

<sup>177</sup> Handbook of Children and Screens at pdf p. 423; *See also* Vanden Abeele MMP, *Digital wellbeing as a dynamic construct*, Com Theory, 2021;31(4):932–55, <https://doi.org/10.1093/ct/qtaa024>.

<sup>178</sup> META3047MDL-072-00318089, Slide 38

<sup>179</sup> Handbook of Children and Screens at pdf p. 54; *See also* Munzer T, Torres C, Domoff SE, et al. *Child media use during COVID-19: associations with contextual and social-emotional factors*, J Dev Behav Pediatr. 2022; 43(9):e573, <https://doi.org/10.1097/DBP.0000000000001125>.

spending more time on the Defendants’ platforms than they originally intended.<sup>180</sup> Some studies cited (and commissioned) by the Defendants have found that these effects are mitigated by “active” use of their platforms—e.g., generating content or posting content. However, this hypothesis has not been widely accepted,<sup>181</sup> particularly since there is evidence that children are more likely to “watch, play, or scroll through content created by others than they are to use their devices to produce their own content.”<sup>182</sup>

#### E. Beauty Filters

195. Despite being a relatively new social media feature, Augmented Reality (“AR”) filters—commonly referred to as “Beauty Filters”—have been thoroughly studied due to the outsized negative impact they could have on users by exacerbating a socio-psychological phenomenon known as “social comparison.”<sup>183</sup> Prior to the advent of these Beauty Filters, researchers had already identified that social media may exacerbate social comparison and lead to

---

<sup>180</sup> Handbook of Children and Screens at pdf p. 483; *See also* Flayelle M, Brevers D, King DL, Maurage P, Perales JC, Billieux J, *A taxonomy of technology design features that promote potentially addictive online behaviours*, *Nat Rev Psychol.* 2023;2(3):136–50, <https://doi.org/10.1038/s44159-023-00153-4>.

<sup>181</sup> Handbook of Children and Screens at pdf p. 152; *See also* Valkenburg PM. Social media use and well-being: What we know and what we need to know. *Curr Opin Psychol.* 2022;45:101294. <https://doi.org/10.1016/j.copsyc.2021.12.006>; Handbook of Children and Screens at pdf p. 187 (“[S]ome scholars suggest that this dichotomy of passive use being negative and active use being positive has too many exceptions to truly understand youths’ experiences online and outcomes for well-being.”).

<sup>182</sup> Handbook of Children and Screens at pdf p. 541; *See also* Rideout V, Peebles A, Mann S, Robb MB. *Common sense census: media use by tweens and teens*, 2021. Common Sense. 2022; Accessed 24 Mar 2023. [https://www.common sense media.org/sites/default/files/research/report/8-18-census-integrated-reportfnal-web\\_0.pdf](https://www.common sense media.org/sites/default/files/research/report/8-18-census-integrated-reportfnal-web_0.pdf)

<sup>183</sup> Handbook of Children and Screens at pdf p. 178; *See also* Thompson JK, Heinberg LJ, Altabe M, TantleffDunn S, *Exacting beauty: theory, assessment, and treatment of body image disturbance*, American Psychological Association; 1999. <https://doi.org/10.1037/10312-000>.

increased body dissatisfaction and/or disordered eating.<sup>184</sup> However, with the introduction of Beauty Filters—many of which were developed by the Defendants—users are now exposed to “manipulated” photos that depict unrealistic (and in some cases impossible) body image standards.<sup>185</sup>

#### F. Safer Alternative Design

196. I have been asked to consider what recommendations would improve the safety of social media platforms for use by teenagers. Based upon the literature and my own research, decreasing the number of addictive design features would reduce harms substantively both because it would reduce problematic use and all of the other attendant untoward events it leads to (e.g. sleep disturbances). This includes removing design features that foster negative social comparison (such as filters), reducing notifications that distract, and removing metrics, such as snap streaks, which are highly addictive.

---

<sup>184</sup> Handbook of Children and Screens at pdf p. 178; *See also* Holland G, Tiggemann M. *A systematic review of the impact of the use of social networking sites on body image and disordered eating outcomes*, *Body Image*. 2016;17:100–10, <https://doi.org/10.1016/j.bodyim.2016.02.008>; Roberts SR, Maheux AJ, Ladd BA, Choukas-Bradley S, *The role of digital media in adolescents’ body image and disordered eating*, In: Nesi J, Telzer EH, Prinstein MJ, editors. HANDBOOK OF ADOLESCENT DIGITAL MEDIA USE AND MENTAL HEALTH, 1ST ED. Cambridge University Press; 2022. p. 242–63, <https://doi.org/10.1017/9781108976237.014>; Rodgers RF, *The relationship between body image concerns, eating disorders and internet use, Part II: An integrated theoretical model*, *Adolesc Res Rev*. 2016;1(2):121–37, <https://doi.org/10.1007/s40894-015-0017-5>; Perloff RM, *Social media effects on young women’s body image concerns: theoretical perspectives and an agenda for research*, *Sex Roles*. 2014;71(11-12):363–77, <https://doi.org/10.1007/s11199-014-0384-6>.

<sup>185</sup> Kleemans, Daalmans, Carbaat, & Anschütz (2018). Picture Perfect: The Direct Effect of Manipulated Instagram Photos on Body Image in Adolescent Girls. *Media Psychology*; Spitzer, Crosby, & Witte (2022). Looking through a filtered lens: Negative social comparison on social media and suicidal ideation among young adults. *Psychology of Popular Media*; *See also* Meta Research Summary: <https://docs.google.com/document/d/1w-HOfseF2wF9YIpXwUUtP65-olnkPyWcgF5BiAtBEy0/edit?pli=1&tab=t.0#heading=h.sh24qmab6i4m>

197. In addition, platforms that fully disclose the risk of harms to parents and children would increase safety. Based upon my work as a pediatrician, academic, and my knowledge of public health, users do not expect social media to be as addictive and harmful as the literature supports and neither parents nor society treat it as such (viz age restricting access to adults as we do for alcohol or tobacco).

**X. Internal Documents Connecting Features to Harm**

198. Notably, many of the features recognized in the literature as addictive and/or harmful, were identified in research conducted by Meta and presented in 2021. A Mixed Methods Clinicians study identified product features and pathways impacting mental health, including the following:

<sup>3</sup>Based on primary clinician-identified product features and pathways impacting mental health, key potential opportunity areas emerge (continued on next slide)

Product Affordance	Opportunities
The ability to give and receive <b>quick reactions</b> , or feedback on, people's posts, comments, videos, and pictures (e.g., click "like" or "angry face" on a picture someone posts)	<ul style="list-style-type: none"> <li>• Allow users to designate posts (or incorporate as a mode) that mutes reactions</li> <li>• Nudge people to take breaks from use</li> <li>• Educate young users on interpreting social media vs. real life, especially for FOMO inducing post (e.g., banner with educational content linked)</li> <li>• Support users who are being bullied by rapidly removing reported content</li> <li>• Keep encouraging connections, especially with designated close friends and family</li> </ul>
Ability to <b>reference</b> others/be referenced (tagging)	<ul style="list-style-type: none"> <li>• Support users who are being bullied by rapidly removing reported content that they have been tagged in</li> <li>• Educate young users on interpreting social media vs. real life, especially for FOMO inducing post (e.g., banner with educational content linked)</li> <li>• Obtain consent for each instance of tagging</li> <li>• Assess user patterns and restrict potential stalking behavior and flag undesignated users that are repeatedly checking a tag</li> <li>• Notify users of unusual interaction with content</li> <li>• Keep encouraging connections, especially with designated close friends and family</li> </ul>
Ability to share and circulate others' content ( <b>reposting</b> )	<ul style="list-style-type: none"> <li>• Support users who are being bullied by rapidly removing reported content and assist in removing negative circulated content</li> <li>• Enable users to block other users from sharing and circulating their content</li> <li>• Educate young users on interpreting social media vs. real life, especially for FOMO inducing post (e.g., banner with educational content linked)</li> <li>• Keep encouraging connections, especially with designated close friends, family, or those with shared experiences</li> </ul>

Document 37: META3047MDL-040-00049387 at Slide 18

Product Affordance	Opportunities
The ability <b>play automatically</b> without pressing a play button	<ul style="list-style-type: none"> <li>• Nudge people to take breaks from use</li> <li>• Nudge people to reflect on intentions for amount of time they're spending and whether they are meeting that intention</li> <li>• Educate young users on interpreting social media vs. real life, especially for FOMO inducing stories</li> </ul>
The ability to <b>easily click on ads to buy things</b> (e.g., links that take you to a site to buy a recommended product)	<ul style="list-style-type: none"> <li>• Enable users to proactively designate modes that mute easily clickable ads</li> <li>• Flag and prevent users engaging in compulsive shopping (i.e. concentrated and accelerated rates of ad engagement)</li> <li>• Identify when people are engaging problematically with ads and add friction</li> </ul>
The fact that people see posts, videos, images, and ads based on what they have <b>previously done on social media</b> (i.e., what you've liked, what you've searched for, time you've spent watching certain videos, etc.)	<ul style="list-style-type: none"> <li>• Continue to encourage positive and educational content</li> </ul>
	<ul style="list-style-type: none"> <li>• Nudge people to take breaks from use</li> <li>• Nudge people to reflect on intentions for amount of time they're spending and whether they are meeting that intention</li> <li>• Ensure ranking/recommendations do not promote mental health-related misinformation</li> <li>• Ensure ranking/recommendations do not promote problematic mental health-related content (e.g., "thinspo")</li> </ul>
	<ul style="list-style-type: none"> <li>• Continue to encourage positive and educational content</li> </ul>

*Document 38: META3047MDL-040-00049387 at Slide 20*

Of these features, several were identified as “primarily negative,” including Video/Photo filters, location sharing, automatically playing videos, and pop-up notifications.<sup>186</sup>

199. Internally, Defendants documents recognize that they could increase engagement by changing the design of the social media. As early as 2016, Meta (then Facebook) was exploring ways to keep teens on its site and posting content. An exemplar document reflecting the company findings would include the following:

<sup>186</sup> META3047MDL-072-00318089 at Slide 87

#### Summary of Insights and Opportunities

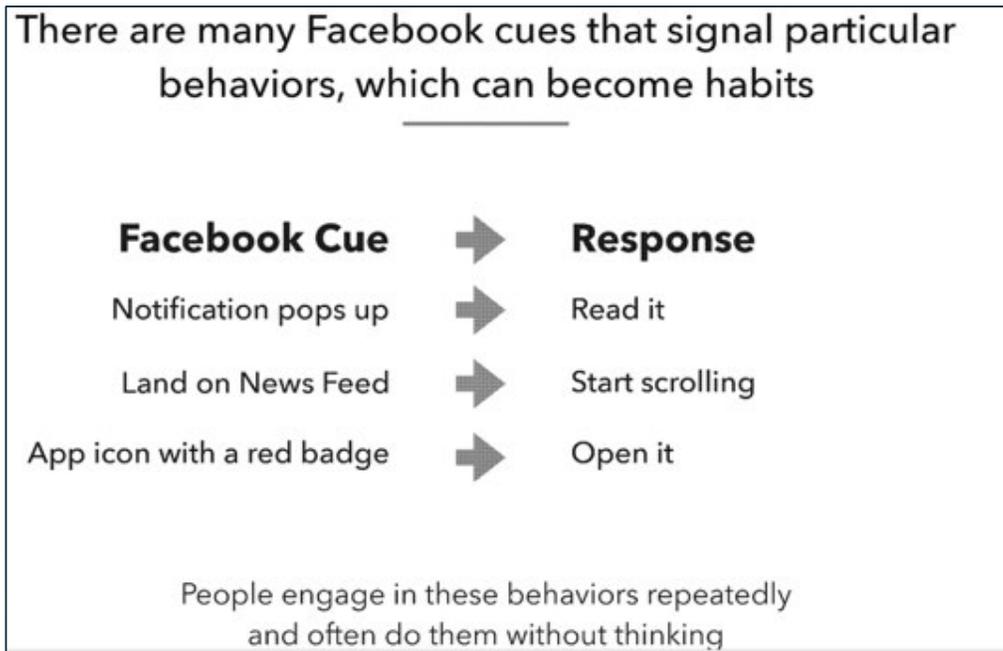
- Insight: Teens posting behaviors reflect that of their friends; both in how much they post and the type of content they share.
  - Opportunity: Ensure teens see lots of content from their teen friends
- Insight: Comments and likes greatly affect teens' likelihood to post and how frequently they post.
  - Opportunity: Create a lower bar for feedback (e.g. views instead of likes, easier commenting— emojis!)
- Insight: Feeling as though they have no photos to share is a barrier to sharing and true especially for Decreasers.
  - Opportunity: Make it easier to take and share photos (e.g. camera first).
- Insight: Fear of losing followers is the number one reason teens don't share to Instagram.
  - Opportunity: Relieve the posting pressures of losing followers by creating a light option for "unseeing" content, instead of "unfollowing" an account.
- Insight: Teens worry about editing (taking too long or doing it perfectly) and it deters them from posting.
  - Opportunity: Make content easier to edit.
- Insight: Tenure on Instagram might discourage posting. This could possibly be due to getting used to posting infrequently per Instagram expectations or not wanting to add too much content to an already content-full profile.
  - Opportunity: Create new ways to share outside of profile (this data was collected pre stories!).
- Insight: Having no photos to share perpetuates a "no-sharing" cycle.
  - Opportunity: Find a way to encourage low posters to share on Instagram.

*Document 39: META3047MDL-031-00096208, -6209*

200. Meta also studied notifications and the ability for this feature to induce habitual or addictive behaviors. For example, one internal document supports the basic addiction principle that “experiencing a reward (or reinforcement) can increase learning and motivation. This contributes to repeated, potentially habitual behaviors.”<sup>187</sup> The document included the following figure:

---

<sup>187</sup> META3047MDL-014-00359270, -9296



*Document 40: META3047MDL-014-00359270, -9302*

201. It is not surprising then, that among the many metrics Meta tracks, the “success rate” of their notifications, defined as increasing daily usage, is a key one. Darius Kilstein, a Director in Data Science at Instagram, reports:

**37% (13.3B daily) of generated IG notifications (~36B, excluding Direct) are received by the user’s device while 63% fail. While DM teens have a higher delivery rate to Push Infra (52%) compared to non-teens (40%), both groups behave similarly in the Push infra funnel**

*Document 41: Darius Kilstein Deposition Exhibit 13 at Slide 47*

202. This is to say that Meta is monitoring (and presumably modulating) the intermittent reward mechanism to ensure that engagement is maximized. As Max Eulenstein, VP of Product, says in a Meta email on Jan 26, 2021, “No one wakes up thinking they want to maximize the number of times they open Instagram that day. But that’s exactly what our product teams are trying

to do.”<sup>188</sup> One mechanism that appears to be especially effective at engaging teens (and adults) is the use of “reels” or short videos that repeat. Meta adopted reels from TikTok after seeing how effective they were at promoting usage.<sup>189</sup> Below are Instagram’s metrics on the viewing of such reels by teens.



Document 42: Darius Kilstein Deposition Exhibit 14 at Slide 20

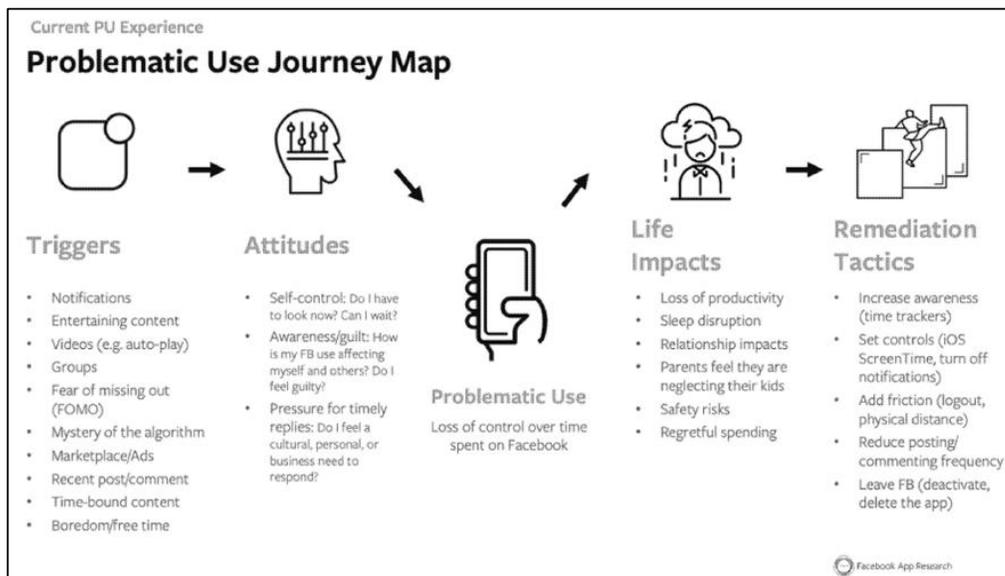
203. The data demonstrated that active daily US teen Instagram users view on average almost 105 reels per day on the platform for a total of 22.5 minutes per day or an average of 20 seconds per reel per day.<sup>190</sup> Briefly watching a short snippet algorithmically curated to one’s interest is gratifying via the dopamine reward pathway discussed in section X.

<sup>188</sup> META3047MDL-014-00352250, -2251

<sup>189</sup> Darius Kilstein Deposition Exhibit 14 at p. 20

<sup>190</sup> Darius Kilstein Deposition Exhibit 14 at p. 20

204. Further, their internal research identified the following “Triggers for Problematic Use on Facebook” which were later included as part of a “Problematic Use Journey Map” as early as 2020 below.<sup>191</sup>



Document 43: META3047MDL-079-00000177, -0200

205. Each “opportunity” for a reward represents a potential algorithm tweak, and many --if not all-- of these were eventually incorporated in some form into the site. For example, a 2018 Facebook presentation has the following two slides:

<sup>191</sup> META3047MDL-079-00000177, -0201

Several aspects of Facebook are consistent with things that research has shown trigger dopamine release and reward-related brain activation (e.g., in the ventral striatum, or VS)



Social approval



Unpredictable rewards

Document 44: META3047MDL-044-00091392 at Slide 24

## Example neuroscience findings: Facebook and reward

**1. Instagram pictures with more vs. fewer "Likes" activate the ventral striatum (VS)**

This means Like counts provide meaningful information for whether something is important.

By "Facebook symbols" does that mean logos? like the blue f or like sign? could include that.

**2. Facebook symbols activate the VS more if you are a frequent user**

If you use FB often, presumably you find it rewarding, and also those images will be more familiar to you and therefore more salient.

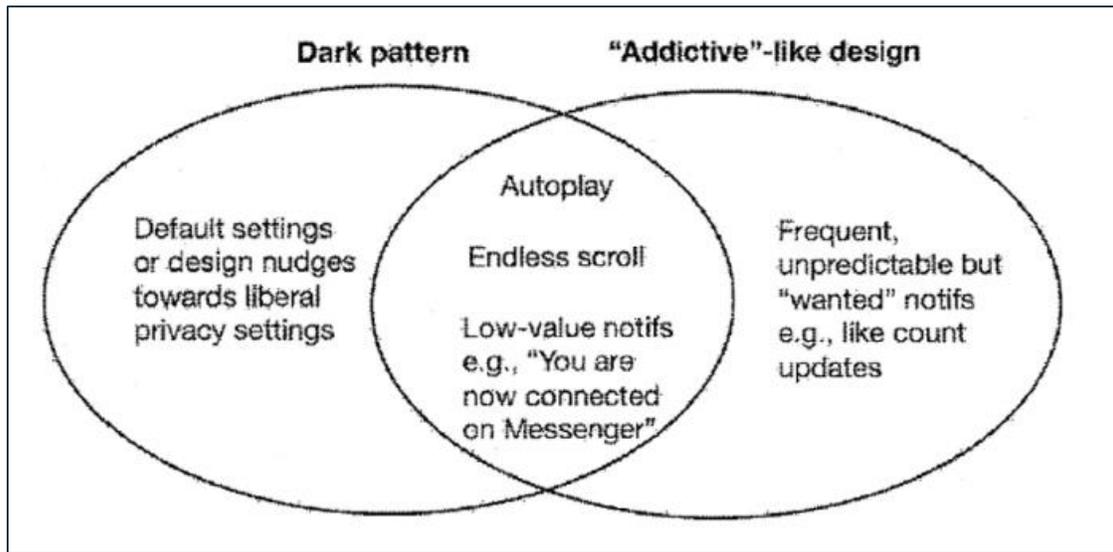
**3. People who have stronger VS responses to reputation gains as compared to monetary gains are also more likely to use FB more**

People who are particularly sensitive to social rewards might find FB use more rewarding and therefore use it more.

1. Sherman et al., 2016, 2017    2. Turel et al., 2014    3. Meshi et al., 2013

Document 45: META3047MDL-044-00091392 at Slide 25

206. Meta's more recent documents present an interesting Venn diagram portraying one way to conceptualize the usage experience from a design perspective (see below).



*Document 46: META3047MDL-044-00108564, -8566*

“Dark Pattern is a user experience term referring to interactions that are deceptive, or that trick you into doing something you didn’t want to do.”<sup>192</sup> Again, this diagram acknowledges that they have “addictive” design features and that some of them are set to “default.” Meta documents from Project “Plato” that was intended to study and mitigate “dark patterns” state that “Some [Facebook] UX patterns rob users of their agency.”<sup>193</sup> Included among them are:

<sup>192</sup> META3047MDL-044-00108564, -8566

<sup>193</sup> META3047MDL-047-01030786, -0786

1. **Bait & Switch:** You set out to do one thing, but a different, undesirable thing happens instead.
2. **Misdirection:** The design purposefully focuses your attention on one thing in order to distract your attention from another.
3. **Privacy betrayal:** You are tricked into publicly sharing more content or information than you really intended to.
4. **Roach motel:** The design makes it very easy for you to get into a certain situation, but then makes it hard for you to get out of it.
5. **Unpredictable, low-value rewards:** The design provides a cue that typically represents a valuable or gratifying experience, but it is generally rated as low value (e.g., certain notifications). In addition, cues for rewards come at unpredictable times and drive repeated checking or updating behaviors. (from "addiction" work)

AL

META3047MDL-047-01030786

*Document 47: META3047MDL-047-01030786*

207. Meta’s researchers also knew that being too critical of these design choices could lead to criticism from internal stakeholders. For example, Jennifer Guadagno noted that her wellbeing team’s efforts to study Facebook addiction as part of Project Plato could “get heavy pushback internally” and added that she was not “sure if it’ll even be allowed to happen” because “if we now know all these things that are potentially bad and then we don’t do anything to fix them” it could lead to problems for the company externally.<sup>194</sup> Bejar in his deposition flatly attests that Meta did not do enough to warn parents or curtail problematic use of their products and that the word addictive is ‘radioactive.’<sup>195</sup>

Jennifer L Guadagno (8/17/2018 14:24:58 PDT):  
 >The one thing that happened recently that's making me more nervous about it all is that we may get heavy pushback internally for doing our side of the work. I'm not sure if it'll even be allowed to happen (main reason being that what if we now know all these things that are potentially bad and then we don't do anything to fix them)

Jennifer L Guadagno (8/17/2018 14:25:59 PDT):  
 >So my concern in combining them would be that our "addiction"/problematic use side is too risky so it's easier to drop that and just do dark patterns. So we'd lose the leverage to push on the "addiction" side specifically

*Document 48: META3047MDL-040-00593848, -3848*

208. Despite learning of these problems as early as 2018, Meta had not made meaningful changes to the platform nearly two years later. A Meta presentation from 2020 arrived at nearly

<sup>194</sup> META3047MDL-040-00593848, -3848

<sup>195</sup> Arturo Bejar Dep. Tr. 136:17-19.

identical conclusions that the Project Plato researchers reached, including the connection between design features and problematic usage. One slide from that 2020 presentation stated that research participants reported “10+ triggers contributing to [problematic use] habits” including:

**Current PU Experience**

## We heard about 10+ triggers contributing to PU habits.

- **Notifications** – Getting too many minor/irrelevant notifications. Try to only look at important ones but get sucked into longer sessions.
- **Entertaining content** - Many said they'd open FB with a clear intent (like checking the news, a specific group, or work-related posts), then get distracted by something entertaining
- **Videos** – Easy to get immersed (especially before bedtime). Auto-play exacerbates the issue.
- **Groups** – Get exponentially more notifications, engaged in chat threads
- **Fear of missing out (FOMO)** – Worry about missing important world news or updates in their social circles.
- **Mystery of the algorithm** – Uncertainty over if they will see posts from those they want; if they can find a post again later.
- **Marketplace / Ads** – Vigilance of buyers/sellers, sales, lower resistance to purchasing at night
- **Recently post/comment** – Higher curiosity to see responses
- **Ephemeral content** (e.g. Stories, birthdays) - Catch it before it's gone
- **Boredom/free time** – Desire to fill downtime or “time pass”

**PII**

*Red dots are toxic on the home screen.*  
P4, 25-34 (m) US

*People liking things can be addictive. I feel compelled to see who liked it. I think it's a bad habit because [I'm] always checking.*  
P2, 45-54 (m) US

*The algorithm doesn't always know what I want to see. I have to do the work to find what I want to see.*  
P7, 35-44 (f) US

*What bothers me the most is getting so entertained; I lose track of time. Especially with the videos. Wow, I spend a lot of time on the videos because they start automatically and when I realize it, I'm already watching.*  
P1, 29 (f) Brazil

Document 49: META3047MDL-079-00000177, -0201

209. Meta was not the only social media company to reach these conclusions about their platforms’ features. For example, YouTube describes “finding a video on YouTube search” as a “predictable reward” compared to “unpredictable rewards” such as “finding a new favorite song while in a mix.”<sup>196</sup> This slide summarizes the latter as “disproportionately [more] delightful compared to predictable rewards because they’re unexpected or exceed expectations.”<sup>197</sup>

YouTube’s researchers describing the reason for this dichotomous design have stated:

When thinking about habit building around YouTube, it is important that we reliably fulfill their goal pursuits (reliable reward) while also providing surprising nuggets of reward (variable reward). Ultimately you're giving people more reasons

<sup>196</sup> GOOG-3047MDL-01268284 at Slide 10

<sup>197</sup> GOOG-3047MDL-01268284 at Slide 10

to come back until they can't even remember why they did. When is the last time you had a goal in mind when you went to Facebook?"<sup>198</sup>

210. YouTube also recognized the presence of these features in their competitor's platforms as well. For example, in the same presentation, YouTube recognized that "repetition reinforces behavior" and identified that Facebook had created "context chains" in which users would reply to a comment, check their News Feed, post a comment, and then repeat the cycle.<sup>199</sup> They also identified a similar behavioral feedback loop for sending Snaps on Snapchat.<sup>200</sup>

211. And later in the same presentation, the role that repetition and reward play, as well as the design features deployed by both YouTube and its competitors is reviewed.

**PRINCIPLE 5**  
**Repetition reinforces behavior**

**W**hen a rewarding action is performed, an association is created between the situation and action. Repetition reinforces and establishes this association, making alternative actions less accessible in that situation.

To form a habit, a behaviour must repeatedly occur as a response to the same stimulus or in the presence of the same context.

Once a link is formed through repetition, you can create "context chains" that habitually link one behavior to another.

**Facebook:**  
Reply to a comment → Check your feed → Post a comment → etc

**Steam:**  
Open app → Check 12 hour Flash Sales → Chat with friends → etc

**YouTube:**  
Receive notification → Check video → Check other videos → etc

 UX Research



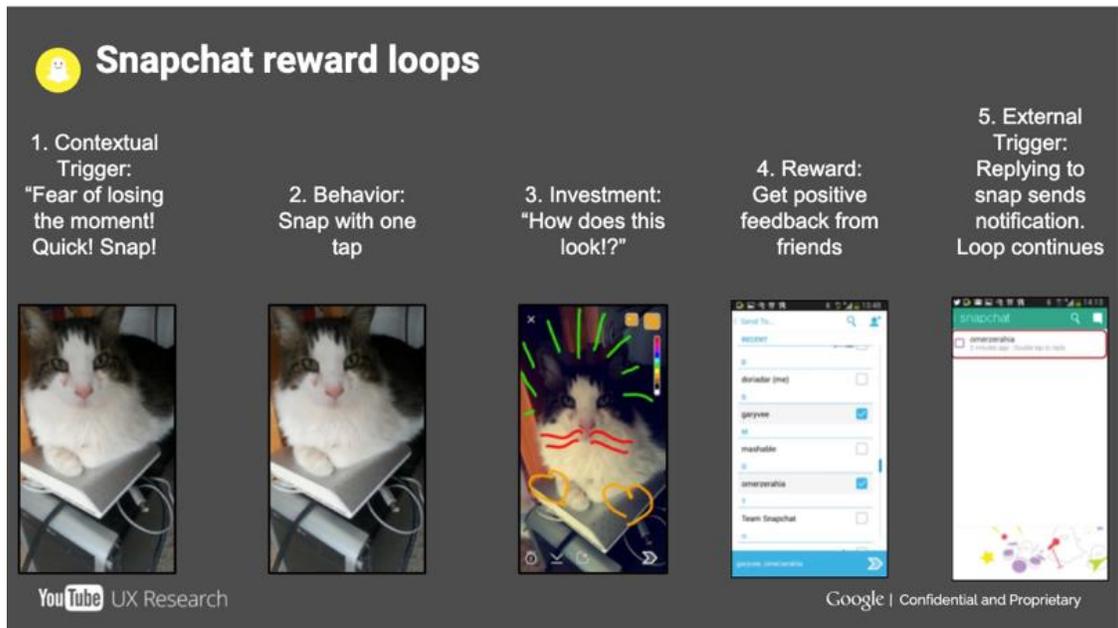
**Example**  
*Clash of Clans, World of Warcraft, and other successful games use "Daily Rewards" as a way to train users to open the app every day. Rewards become progressively larger the longer the "habit chain" lasts.*

Document 50: GOOG-3047MDL-01268284 at Slide 11

<sup>198</sup> GOOG-3047MDL-01268284 at Slide 11

<sup>199</sup> GOOG-3047MDL-01268284 at Slide 11

<sup>200</sup> GOOG-3047MDL-01268284 at Slide 34



*Document 51: GOOG-3047MDL-01268284 at Slide 34*

212. Internal documents also reveal that Instagram is tracking usage/engagement by teen users. An email from Darius Kilstein, on February 5, 2022, states that Meta “looked into the long-term decline of feed imp[ressions] for Teens” and discovered that “teens aren’t relying on Feed for interest consumption as much as they used to” and hypothesized that this shift might explain “why teens now consume fewer Reels than adults.”<sup>201</sup>

- **We looked into the long-term decline of feed imps for Teens and what we found was fascinating:**
  - There may be a generational shift away from Feed. 13yo’s average half as many Feed imps/DAU as 17yo’s & each new cohort of teens consumes less on Feed than the previous.
  - Among teens, US Home Feed Imps/DAU declined -30% between July 2020 and January 2022, compared to -8% for adults. This finding also explains why teens now consume fewer Reels than adults (they under-consume on RIFU (Reels in Feed Unit) & Chaining from Feed).
  - Teens aren’t relying on Feed for interest consumption as much as they used to, and other products aren’t sufficiently compensating for the loss, posing a significant headwind for the cohort

*Document 52: Darius Kilstein Deposition Exhibit 9, at -7079*

<sup>201</sup> Darius Kilstein Dep. Exhibit 9 at -7079

The reported “headwind” for the “cohort” points to Instagram’s business need to redesign features to better engage younger children and keep up with their competitors (particularly TikTok) as Kilstein says later in the same email.<sup>202</sup>

213. Particularly when, as here, Meta has operationalized its “time spent” metric in order to “make projections for monetization.”<sup>203</sup> This fact is further supported by Meta’s Form 10k submissions to the SEC which state that its “advertising revenue can also be adversely affected by a number of other factors including: decreases in user engagement, including *time spent on our products*.”<sup>204</sup> This was even acknowledged during Mark Zuckerberg’s deposition in this case when he agreed that the amount of money his companies make is directly related to the amount of time users spend scrolling past and interacting with advertisements on the platform.<sup>205</sup>

214. The purest example of the way in which Meta designed its platforms in order to exploit their users’ attention is the changes to the “News Feed” features. In its original form, Facebook’s “feed” simply allowed people to update their profiles with new events and presented them in chronological order. Later, it was re-engineered to present information based on what Meta’s algorithm predicted a user wanted to see. In 2017, several years after the first algorithm was introduced, a Meta researcher asked whether “algorithms to blame for Facebook addiction?”<sup>206</sup> They concluded that while “research hasn’t addressed this [question],” an algorithm

---

<sup>202</sup> Darius Kilstein Dep. Exhibit 9 at -7080

<sup>203</sup> Darius Kilstein Dep. Tr. at 409:17-20

<sup>204</sup> Meta 2023 Form 10K Filing

<sup>205</sup> Mark Zuckerberg Dep. Tr. at 194:9-12 (“Q. Mr. Zuckerberg, if users spend more time on their Instagram or Facebook Feed, then generally speaking, they’ll see more ads, yes... Q. And you don’t deny that the more ads get seen, the more ad revenue Meta earns, yes? A. .... In general, the more ads that people see, the more opportunities we have to show people relevant ads. Q. Which means more advertising revenue you have a chance to earn? A. Yeah. In general, from a business perspective, I think that’s roughly correct.”).

<sup>206</sup> META3047MDL-005-00000001, -0003

that “favor[s] content or functionality that encourages people to spend more time on Facebook, then it's possible that this will by its nature tap into addictive mechanisms” and lead to “addictive/compulsive usage more severe and more widespread.”<sup>207</sup>

215. As Kan-Xing states in his deposition “The problem that I think news feed was trying to solve was if you had maybe a hundred or 200 friends, it was actually pretty time-intensive and not that efficient a use of time to --like, if you're just trying to figure out what's new or what's happening with your friends, to go through all 100 or 200 of them.”<sup>208</sup> Later in his deposition, reading from an internal Meta document, he quotes “In essence, Facebook users didn’t think they wanted constant up-to-the-minute updates on what other people are doing, yet when they experienced this sort of omnipresent knowledge, they found it intriguing and addictive.”<sup>209</sup> In 2015, as quoted in the same deposition, Zuckerberg himself states “I’ve spent a lot of time recently thinking about the decline in content production, and I wanted to upgrade our sense of urgency around this. I think this is the biggest issue we must now address as a company.”<sup>210</sup> Teen engagement in particular was optimized by making the default news feed only include posts from people in their age range with the rationale that parental posts would be of less interest.<sup>211</sup> The “infinite feed” invented at Facebook was exported and adapted to Instagram. In her deposition, Dr. Lee states:

---

<sup>207</sup> META3047MDL-005-00000001, -0003

<sup>208</sup> Kang-Xing Jin Deposition Tr. at 169:16-22

<sup>209</sup> Kang-Xing Jin Deposition Tr. at 171:17-22; *see also* Kang-Xing Jin Dep. Exhibit 10 at -6106

<sup>210</sup> Kang-Xing Jin Deposition Tr. at 234:23-235:3; *see also* Kang-Xing Jin Dep. Exhibit 17 at -6298

<sup>211</sup> Kang-Xing Jin Deposition Tr. at 236:9-14; *see also* Kang-Xing Jin Dep. Exhibit 17 at -6302

12	Let's assume that you have been on
13	Instagram all day and you've been scrolling all day
14	and you finished seeing all the possible content
15	that you could possibly see from your connected
16	accounts. The people that you follow. Once you hit
17	the very bottom of that, there's another surface
18	called end of feed recommendations. That's actually
19	a slightly different set of recommendation
20	algorithms that dictate end of feed, or EOF,
21	compared to in-feed recommendations.

*Document 53: Alison Lee Deposition Transcript at 30:12-21*

The quote starts with the scenario wherein someone has been “scrolling” all day (itself an implicit acknowledgement of overuse) and exhausts their “connected” feeds at which point a new, end of feed algorithm is triggered. Dr. Burke in her deposition talks about the “Friend Paradox” whereby someone’s friends have a higher percentage of likes than they do and says “[i]f News Feed is optimized based on how many likes a post gets, then yes, it could make that paradox appear worse.”<sup>212</sup> That is exactly how News Feed is optimized per Kan-Xing.

216. Raskin invented the infinite scroll. This feature allows a user to scroll indefinitely through their feed, receiving endless posts and intermittent variable rewards. At deposition, Raskin testified that infinite scroll was like “digital cocaine” for the user. ██████████, who self-identifies as one of the technology leads for Instant Articles on Android at Facebook, reports “we only care about things like time spent, open links etc. That’s what we optimized for. That’s what we defined as success and failure.”<sup>213</sup>

---

<sup>212</sup> Moira Burke Dep. Tr. at 143:1-4

<sup>213</sup> Haugen\_00001033, -1033

217. ██████████ (Sr. Director of Data Science and Engineering) goes further in 2018 when he poses what he deems a “blasphemous” question: “Is ranking good?”<sup>214</sup> He is apparently referring to the algorithm that determines the sequence of posts on a user’s feed. He goes on to write, “The experimental and experiential data in support of ranking is extensive and nearly universal. When we switch a random set of cues to a pure chronologically News Feed, their usage and engagement immediately drops.”<sup>215</sup>

218. This both demonstrates and acknowledges that algorithms are designed to promote and sustain engagement. He goes on to say, “**If we abruptly stopped ranking News Feed Tomorrow, the results would be disastrous for the company by most metrics we care about.**”<sup>216</sup> (emphasis present) But the design features that maximized engagement were not limited to the news feed alone. In his deposition, George Volichenko (Software Engineer) refers to “engagement triggers” and lists as an example a red dot over the app icon on the home screen with a number of notifications.<sup>217</sup>

219. Even as they actively worked to maximize teen engagement, Facebook’s internal documents evince that teens want help managing their time on the app:

---

<sup>214</sup> Haugen\_00002372, -2372

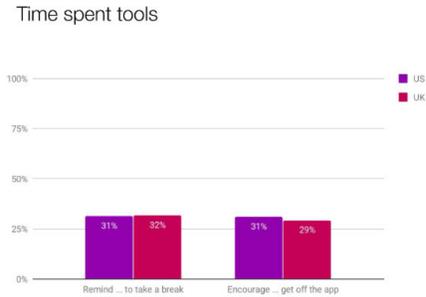
<sup>215</sup> Haugen\_00002372, -2374

<sup>216</sup> Haugen\_00002372, -2386

<sup>217</sup> George Volichenko Dep. Tr. at 131:13-24

## Teens want help controlling the time they spend on the app

- Teens talk about the amount of time they spend on Instagram as one of the “worst” aspects of their relationship to the app.
- They have an addicts’ narrative about their use -- it can make them feel good, feel bad. They wish they could spend less time caring about it, but they can’t help themselves.
- Teens recognize the amount of time they spend online isn’t good for them but at the same time know they lack the willpower to control the time spent themselves



Q: Now you're going to see some things that Instagram could do to help teens. Please select your top 3 for what Instagram should do.  
 US n = 1296; UK n = 1308

Document 54: META3047MDL-003-00109173, -9221

220. In fact, in 2018, Meta acknowledged that “we can make changes to Facebook so it has less potential to be habit-forming and provide support for people to “break” Facebook “habits” they don’t want.”<sup>218</sup> Further, the document goes so far as to propose changes in four areas:

<sup>218</sup> META3047MDL-014-00359270, -9307

<p><b>Awareness</b></p> <p>Increase awareness by <b>getting feedback</b> on behavior, receiving <b>alerts</b>, and <b>tracking behavior</b> over time</p>	<p><b>Attentiveness</b></p> <p>Increase attentiveness by adding <b>new salient stimuli</b> or "speed bumps" to <b>add friction to automatic behaviors, such as making it harder to find unwanted distractions</b></p>
<p><b>Intentionality</b></p> <p>Increase intentionality by <b>prompting more choices, reconsidering defaults</b>, and providing <b>support for setting and pursuing goals</b></p>	<p><b>Control</b></p> <p>Increase control over behavior with <b>tools to set limits, prompts for alternative behaviors, and reduced barriers for exercising control</b> (e.g., make it easy to flexibly change settings)</p>

39

ENTIAL META3047MDL-014-00359308

*Document 55: META3047MDL-014-00359270, -9308*

221. A 2018 email from Margaret Stewart VP of Product Design at Facebook states:

I expect the team has already explored this, but have we thought about ways to allow people to limit their time per day on our products? I know this sounds a bit extreme, but if we were to be able to set a budget of time and then be notified when we go over it (similar to ways people manage their food intake), that would go a long way in terms of giving people a sense of control. When I first saw mention of the "time out" feature we launched, I thought that's what it was going to be 😊

*Document 56: META3047MDL-014-00071620, -1621*

222. Kang-Xing in his deposition addressing the "mechanics" of news feed and their role in "problematic use," was asked explicitly about steps Facebook took to address it:

Q: And you're basically saying -- I mean, at a high level, just to explain this in layman's terms for the jury, I mean, you're saying there's a lot of work to be done to get where we need to be on problematic use. I mean, isn't that what this is Page 443 trying to communicate at the bottom of page 1?

THE WITNESS: "I think at a high level, yes, although I think the "where we need to be" definition is one that there probably wasn't broad alignment on either within the company or outside. So I think different people may, like, have different opinions on that.

Q" But, really, by any measure -- wherever that ultimate endpoint was, by any measure, Meta had a lot of work to do to make progress on this issue; right?

THE WITNESS: It was my opinion that there was a meaningful amount of work that still needed to be done in this area, yes.

Q: Was that an informed opinion?

THE WITNESS: It was informed based on the context that I had, yes.

*Document 57: Kang-Xing Dep. Tr. at 442:21-443:22*

223. In 2022, Meta launched MYST, a prospective, longitudinal study to follow 4000 teens (13-17 years of age) using the highly respected National Opinion Research Center (NORC) at the University of Chicago for survey data and Meta logs of activity on their platform.

224. Using validated measures for a variety of outcomes and actual usage patterns on their platforms, Meta intends to shed additional light on many of the outcomes discussed in this report. Having reviewed the analytic plan and approach as well as the *a priori* stated hypotheses, I can say that I am impressed with the design. Though it is not an experimental design (as they attest) it is a very robust, well-designed longitudinal study—the kind that only industry could conduct based on their access to actual SM data. Among other things, they reported that they will test the following negative hypotheses:

Teens who engage in problematic use on Meta platforms will report poorer social-emotional functioning and more mental health symptoms over time.

Teens who engage in more late-night sessions on Meta platforms will report poorer social-emotional functioning and more mental health symptoms over time.

Teens who watch more autoplay videos on Meta platforms will report poorer social-emotional functioning and more mental health symptoms over time.

Teens who use appearance-altering Meta platform products (e.g., photo filters) more often will report poorer social-emotional functioning and more mental health symptoms over time.

*Document 58: META3047MDL-072-00327080, -7089*

225. As the editor of the leading academic pediatric research journal, I would love to see the results of this study come across my desk. However, while the initial plan for the study contemplated its results being released publicly, I have found no evidence that Meta has publicly acknowledged that the study was conducted or released any results (to the contrary, I have seen documents indicating a desire to not make this study public).<sup>219</sup> Indeed, as the study moved forward, they seemed to have become explicitly fearful of it being publicly released. Leaders directed researchers to only refer to the study by its full name, because they would “really prefer not to have a cute acronym (MYST). That makes things sound more interesting when/if it leaks. Can we please keep our names as boring and uninteresting as possible.”<sup>220</sup>

226. Although I was not presented with any of the longitudinal data, since Meta did not include it in the documents they produced to the Plaintiffs, one baseline cross-sectional finding was as follows:

---

<sup>219</sup> META3047MDL-072-00317597

<sup>220</sup> META3047MDL-047-00488494, -8494.

Teens whose parents say they perform more digital supervision/caregiving activities (the sum of their 'yes' responses to the list in Figure 7) spend more time on the platform and have a marginally higher composite score of potential PU-related behaviors (p=0.07). Furthermore, teens who say their parents follow them on social media spend more time on the platform and have more sessions. Teens who say their parents monitor them on social media have a larger composite score of potential PU-related behaviors, spend more time, and have more sessions. Regression models available in [\[\[ HYPERLINK \l "\\_7midhep5cwit" \h \]\]](#).

*Document 59: META3047MDL-050-00215015, -5027*

227. These findings are cross-sectional, and in the absence of additional data, causal relationships cannot be determined. However, the most plausible explanation is that parents are supervising/monitoring their children's activities on social media more when they are concerned about their child's "problematic usage" and those higher levels of supervision are not correlated to lower levels of time spent or "problematic use" of the platform. Thus, these findings are at odds with the notion that more parental monitoring is a potential solution, mediator, or even effective prevention strategy for the untoward effects of social media usage. While Meta's Instagram Parents' Guide instructs parents and kids to "come to an agreement about what is an appropriate amount of time on the platform each day or each week" and "not let screentime rules slide,"<sup>221</sup> internally their researchers were reporting that "[t]here is no association between either parental reports or teen reports of parental digital caregiving/supervision and teens' survey measures" of either habitual use or their ability to control their use of social media.<sup>222</sup>

228. Perhaps more illustrative is the tension that is recognized in Meta's documents discussing MYST. Meta's internal documents reflect that "there is a fundamental tension between doing rigorous research to identify whether there are potential opportunities for

---

<sup>221</sup> Meta, *A parent and guardian's guide to Instagram* 47-48, available for download at <https://about.instagram.com/community/parents>.

<sup>222</sup> META3047MDL-050-00215015, -5027

change/intervention and advancing external/narrative goals.”<sup>223</sup> This tension with external goals is reflected in Meta’s many decisions to not share publicly what their research or outside experts conveyed internally. External goals included the goal to “[a]dvance the credibility of research that finds small or null correlations between social media well-being and/or well-being relative to other factors,” and to “deflate the conversations about research claiming causal connections between social media and mental health and well-being.”<sup>224</sup> Perhaps most critically, Meta’s stated external goal was to “[m]ake the issue bigger than social media. Research should be designed in a way that provides a holistic understanding of the relationship between youth well-being and technology and other factors (e.g., cultural and environmental factors that lead people to use technology and their use of technology, rather than focusing on Meta-specific product offerings.”<sup>225</sup> Thus, the impetus for their research appears to be advocacy or public relations rather than scientific discovery or public safety.

229. YouTube documents also provided evidence of problematic usage by virtue of design. YouTube’s 2019 Strategy offsite includes the following observations about their usage and app features:

---

<sup>223</sup> META3047MDL-072-00317597 at -7598

<sup>224</sup> META3047MDL-072-00317597 at -7598-9

<sup>225</sup> META3047MDL-072-00317597 at -00317599

- Home and Watch Next are by far the biggest feeds on the platform, generating more than 75% (46% Watch Next, 30% Home) of feed content impressions on mobile. The next largest “feed” is Search, with 11% of impressions.
- Direct Response ads are our fastest growing, least budget constrained business. DR ads should be a focus and the DR Ads paper highlights how we will grow demand and evolve our formats and quality. We recently launched video ads throughout the home feed and will launch more demand into these slots over the next 2 quarters.
- Watch Next and comments compete directly for user attention in their current, adjacent configuration on mobile. This produces a zero-sum trade-off for user attention. We’ve historically viewed watch time & Community as similarly important priorities, which has constrained Watch Next to only 16 videos.
- Inline Playback in the Home feed experiments were metrics negative when they focused on shifting watch time to the feed (in 2015). Our current approach, which is metrics positive, provides a preview experience designed to bring users to the watch page, with an emerging secondary use case of watching videos with captions “on the go”.

*Document 60: GOOG-3047MDL-00937887, -7898*

All of these point to design elements that promote increased time on the platform.

230. Soon after, in 2015, YouTube launched Autoplay on desktop (followed by a launch in 2016 to app), which automatically and continuously played the next video – without any need for user action.<sup>226</sup> Autoplay set to on be default because it generates more watch time.<sup>227</sup> Autoplay watch time doubles at night.<sup>228</sup>

231. “This was the single most impactful launch in YouTube history, with +8% desktop watchtime, +4% overall watchtime increase.”<sup>229</sup>

232. These objectives led to a steady and significant increase in watchtime from 2012 onward, and helped achieve the CEO’s 2015 ambitious goal of reaching 1 billion hours of watch time by the end of 2016.

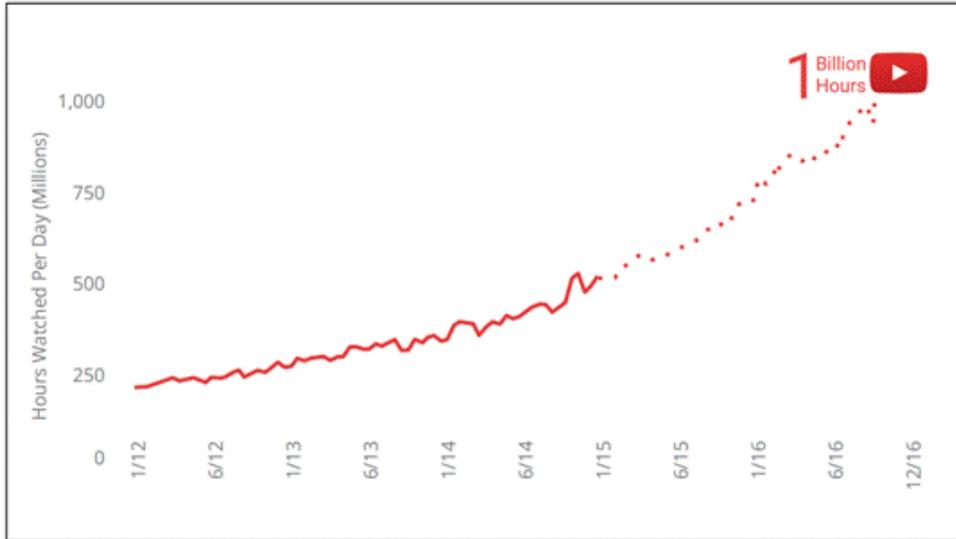
---

<sup>226</sup> GOOG-3047MDL-00767071 at Slide 26; GOOG-3047MDL-04652560 at -5565.

<sup>227</sup> GOOG-3047MDL-00223583 at Slide 9.

<sup>228</sup> GOOG-3047MDL-00223584 at Slide 10.

<sup>229</sup> GOOG-3047MDL-00767071 at slide 27; *See also* Kim Ex. 13, GOOG-3047MDL-04613300 (“The [Autoplay desktop] launch’s impact has been huge. Autoplay now generates 16% of YouTube’s desktop watch time...”).



Document 61: GOOG-3047MDL-00579554 at slide 2.

233. A February 4th, 2016, YouTube presentation included the following “Vision” slide for the app that was then still early in development:

**Vision**  
We aspire to create an app that is.....

**Best in class** Viewers should prefer to watch videos in our app even if they're available elsewhere, and they should consider our recommendations superior to those they find elsewhere

**Robust** Our app should set and constantly raise the standard for reliability and snappiness throughout the entire experience

**Innovative** We should introduce delightful new features big and small before other apps do. We value rapid-fire experimentation and taking risks by trying new things we're not sure about

**Cohesive** All user journeys should be intuitive and effortless, and the level of quality is always consistent throughout the app

**Addictive** Our app experience should compel users to come back more and more often

**Polished** Unparalleled UI is all about details, and no detail is too small to get right. We are proud perfectionists, and we hold other teams to the same standard

*To be refined through XFN team offsites - Product Excellence; Product Playbook*

Document 62: GOOG-3047MDL-00767071 at Slide 11

As seen above (highlighting added), making the app “addictive” was a core design feature. This was consistent with the aspiration articulated in slide 51 of the same presentation of “Building the world’s most powerful and delightful video consumption experience.” YouTube internally acknowledged the potential negative effects of digital videos in a 2018 presentation entitled “Literature Review: Effects of watching digital videos and viewer well-being” by [REDACTED] (User Experience Researcher). Below is a screen capture from later in that presentation:

- **Problematic Internet Use (PIU):** multidimensional syndrome that consists of cognitive, emotional, and behavioral symptoms that result in difficulties with managing one’s offline life.
  - Overlaps with addiction
  - Often irrational and not under conscious control
  - Descends into dysfunction & causes one to avoid working on an intended task
- **“Just One More Video” Effect**
  - Very simple to watch an ongoing sequence of videos (autoplay)
  - Often followed by feelings of guilt

*Document 63: GOOG-3047MDL-00874191 at Slide 8*

234. The slide calls out problematic internet use and states that it overlaps with “addiction” consistent with my belief that the entire continuum to the right of casual use (Figure 20) can be viewed as problematic. Furthermore, it highlights how “autoplay,” a key feature of YouTube, drives the “just one more video effect.” Slide 10 summarizes data from an internal survey of “265 respondents” that calls out the “stickiness” of the app and states that its interactivity and notifications “causes users to feel that they must be aware of what is happening on the platform” which “keeps users on the platform longer.”<sup>230</sup>

235. The sampling frame and methodology of this survey are not evident from the documents provided to me, but it is not material since the slide presents the findings as if they are conclusive, or at least sufficiently robust to take as factual. The same survey yielded the findings

---

<sup>230</sup> GOOG-3047MDL-00874191 at Slide 10

noted below. While cat videos can create a more positive mood, “it is difficult to stop watching the videos.”

**Videos are *initially* used for quick mood management, but result in excessive viewing**



- Survey research suggests that **video watching is a common technique for mood management.**
- Respondents reported watching **cat videos to be in a more positive mood more quickly.**
- After one video is over, it is difficult to stop watching the videos.
- Ultimately, viewers **experience feelings of guilt** for spending so much time doing non-meaningful tasks.

(Myrick, 2015)

*Document 64: GOOG-3047MDL-00874191 at Slide 11*

236. There is something else that is notable in this slide: “videos are used for quick mood management.” In other words, people bring an affect to their search: angst, sadness, depression; or conversely happiness, enthusiasm, excitement. The point is that the content is tailored to the pre-morbid mood and driven by the platform’s algorithms. This is a fundamental way in which YouTube is different from other, “pre-internet” or “analogue” viewing experiences (TV, Cable, VHS/DVD) where options were infinitely more limited and not “auto played” based on one’s mood and prior preferences. Slide 14 of the presentation acknowledges these salient differences while acknowledging that “notifications are a critical part of YouTube and contribute to addiction.” (see below).

## YouTube users control what they want to watch

- YouTube is different from TV because **users can decide what they want to watch.**
- Users will **spend more time on the platform** because they continue to watch things that interest them.
- Studies show the **content people watch correlates with their personality characteristics** (i.e. sensation seeking).
- **Notifications** are a critical part of YouTube and **contribute to addiction.**
  - Users are tempted to watch videos the moment they are uploaded.



(Haridakis & Hansen, 2009; Metro Creative Connection, 2018)

*Document 65: GOOG-3047MDL-00874191 at Slide 14*

237. TikTok’s engagement algorithm is frequently viewed as the most effective in the industry: “more personalized,” “more accurate,” and “more diversified.” They emphasize its effectiveness in their marketing presentations and tout that >53% of suggested videos are “viewed.”<sup>231</sup>

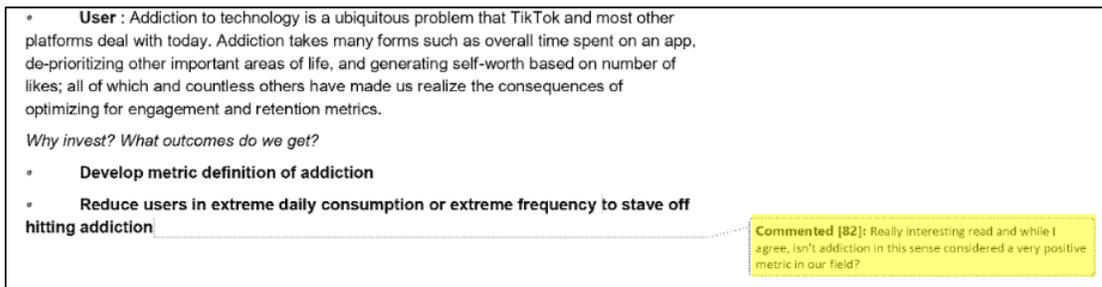
238. An internal memo citing the company’s “vulnerabilities” reports: “In a user survey of 2,300 users in February of 2020, when respondents were asked to give a score of 5 to indicate strength of agreement with the statement ‘I spend too much time on TikTok, the average response was a 4.0.’<sup>232</sup> It goes on to say, “some elements of persuasive design may be unique to TikTok; for instance, the fact that when you click the back button on your phone to leave the For You Feed/all, you get a prompt saying, ‘Tap again to exit,’ which can be seen as increasing friction for

<sup>231</sup> TIKTOK3047MDL-004-00314472, -4483

<sup>232</sup> TIKTOK3047MDL-002-00100441, -0452

users seeking to leave the app.”<sup>233</sup> And still later concedes, “TikTok is particularly popular with younger users, who are seen as more vulnerable to online harms and the negative impacts of compulsive use.”<sup>234</sup>

239. In a “2021 TikTok for Good Business Plan and Vision” document, the following is stated:



The screenshot shows a document snippet with a yellow comment box. The document text includes:

- **User** : Addiction to technology is a ubiquitous problem that TikTok and most other platforms deal with today. Addiction takes many forms such as overall time spent on an app, de-prioritizing other important areas of life, and generating self-worth based on number of likes; all of which and countless others have made us realize the consequences of optimizing for engagement and retention metrics.

Why invest? What outcomes do we get?

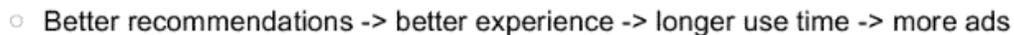
- **Develop metric definition of addiction**
- **Reduce users in extreme daily consumption or extreme frequency to stave off hitting addiction**

The comment box [82] contains the text: "Really interesting read and while I agree, isn't addiction in this sense considered a very positive metric in our field?"

*Document 66: TIKTOK3047MDL-005-00325851, -5862 (emphasis added)*

Notably, the plan acknowledges that “addiction to technology is a ubiquitous problem.” At the same time, comment [82] notes that “addiction” could be “considered a very positive metric in our field.” In other words, for all of its public posturing seeking to discredit or minimize the existence of compulsive or addictive use of screens, internal documents acknowledge its existence and even allude to its “value” to the industry.

240. TikTok’s algorithms are widely considered the “best in the industry” because of their effectiveness at driving engagement. They lay out their business case quite simply:



The screenshot shows a diagram with a single bullet point: "o Better recommendations -> better experience -> longer use time -> more ads".

*Document 67: TIKTOK3047MDL-004-00290146, -0146*

<sup>233</sup> TIKTOK3047MDL-002-00100441, -0452

<sup>234</sup> TIKTOK3047MDL-002-00100441, -0452

In pursuit of that end, they made leaving the app more difficult than others:

◦ This is an issue common for many platforms. However, some elements of persuasive design may be unique to TikTok; for instance, the fact that when you click the back button on your phone to leave the For You feed/app, you get a prompt saying, 'Tap again to exit,' which can be seen as increasing friction for users seeking to leave the app.

*Document 68: TIKTOK3047MDL-060-01158658, -8668*

And for those times when people did succeed in leaving, they refined their “push” approach to, among other things, get people back onto their app:

### **When we send user push?**

Actually, push are mostly used by Algo team and Operation team. Different types of push have different goals:

- ▮ **Interest Push:** based on "no interest push to daily active users" strategy, the goal is to activate users so they will return to the app
- ▮ **Ops Push:** inform users the trending content on TikTok and encourage video view and creation

*Document 69: TIKTOK3047MDL-004-00291835, -1835*

And their “push methodology”, like every change they make to their platform, was apparently subjected to A/B testing where rapid cycle experiments were performed comparing one version to another with users randomized to experience one or the other (i.e. no confounding) and with an eye to ensuring that core metrics were not adversely affected.

- **A few A/B tests here, evidently showing that push notifications and their timings/recipients do matter**

*Document 70: TIKTOK3047MDL-004-00290146, -0149*

241. One can infer from this statement that they tested not just the notifications, but the timings and the characteristics of the recipients. In fact, TikTok is firmly grounded in the idea that every strategic change to their platform should be tested against that metric:

**Q: Can we launch the strategy without conducting AB tests?**

• A: No, only if it's very urgent, for example, safety-related urgent needs. According to our experience, expected outcomes can vary from the actual results drastically. Launching without conducting AB tests can bring damaging influence.

**Q: A strategy can't be launched if the AB tests show no improvement?**

A: No. A strategy can be launched under two circumstances, even if the AB test shows no improvement: when the strategy can only be observed and evaluated through AA; and when a strategy needs to be launched based on our belief and best judgement. There are only a few strategies which can only be observed through AA. Explicit reasons and a clear long-team review mechanism are required for deciding whether a strategy needs to be launched based on the belief and best judgement, the criteria refer to [ [HYPERLINK "https://bytedance.feishu.cn/space/doc/doccnpPffh3n4XlI0HkBHUPyUDa" \h](https://bytedance.feishu.cn/space/doc/doccnpPffh3n4XlI0HkBHUPyUDa) ].

*Document 71: TIKTOK3047MDL-004-00139811, -9822 (emphasis in original)*

242. While focused on the bottom line and rigorous evaluation of changes, the pace of modifications appears to be quite brisk. A frequent phrase, peppered throughout TikTok shared documents is:

**Feedback**

***An absence of feedback after the deadline will be considered as, "Reviewed, no further comments. Defer the decision to the project owner."***

*Document 72: TIKTOK3047MDL-015-00341931, -1934*

243. This “presumptive approval” approach facilitates keeping modifications adherent to deadlines by assuming that no response is assent or agreement. This is further reflected in chats and emails by Julie de Bailliencourt, who was TikTok’s global head of product policy from

June 2021 to May 2024.<sup>235</sup> Bailliencourt's files reflect the tension between growth and safety at TikTok, as reflected in the email below<sup>236</sup>:

**From:** julie.de@bytedance.com [julie.de@bytedance.com]  
**Sent:** 2/17/2022 11:50:20 AM  
**Subject:** Product: we want to launch product G in 4 weeks, can you review and devise safety strategies?

Product: we want to launch product G in 4 weeks, can you review and devise safety strategies?  
T&S: upon review, this is quite a risky product as it stands today. To make it safer, you need to build control Z, control K, and limit people's ability to do X.  
Product: hm. We don't have the time to make these changes.  
T&S: our company will face significant external backlash on the basis of safety if you launch as is, and will put some users at risk.  
Product: but we have an aggressive growth target. You need to come up with a safety strategy.  
T&S: if you took the time to build control Z, control K, and limit people's ability to do X, we could come up with a solid safety strategy.  
Product: we are launching in 4 weeks. Do your job.

3 months later.

Product: hey T&S, we are facing some safety problems with this product. Can you do better?  
T&S: we had advised to make some changes to make this new product safer.  
Product: we need more moderation. We need stricter policies.

*Document 73: TIKTOK3047MDL-081-02351173*

244. And similarly, the culture at Meta was driven by the “move fast and break things.”<sup>237</sup> While the company ethos was designed to enable engagement innovations to move quickly, safety features were held to a different standard. Jayakumar, the youth safety policy lead for Instagram, goes on to say in her deposition that “we had to be very mindful of any impact that it [safety recommendation] might have on growth, and really demonstrate that we were mitigating any potential impact to growth as much as possible.”<sup>238</sup>

---

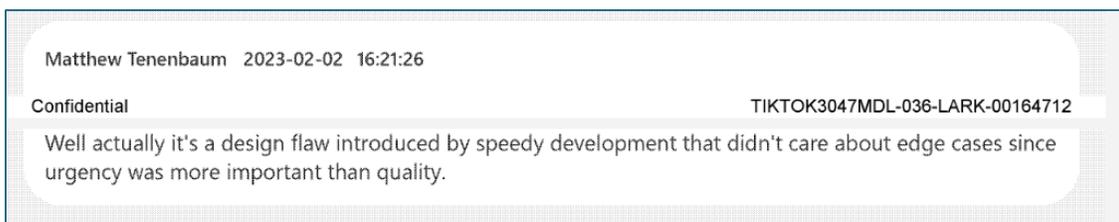
<sup>235</sup> Bailliencourt Dep. at 39:18-23.

<sup>236</sup> This tension between growth and safety is also discussed in Bailliencourt's deposition at pages 164-207 and in exhibit 22.

<sup>237</sup> Vaishnavi Jayakumar Dep. Tr. at 37:7-10

<sup>238</sup> Vaishnavi Jayakumar Deposition Tr. at 30:6-11

245. Simply stated, the approach driving SM companies is fundamentally orthogonal to one that prioritizes safety. In fact, a culture of safety is predicated on verbal affirmation not on tacit or implied agreement. Pilots wait to hear their co-pilots read back their settings; surgeons have checklists they read out and wait to have them confirmed before cutting anything. Similarly, “breaking things” is not a safety motto. It is not surprising then that when reacting to the development of TikTok now, Mathew Tenenbaum (Senior Product Manager) says:



*Document 74: TIKTOK3047MDL-036-LARK-00164712, 4712-13*

In contrast to the breakneck speed at which tech innovations proceed at Meta, “integrity” and “safety” research is on a different track. Dr Lee in her deposition states:

18 A. The pace at which we share at research  
19 takes weeks. It takes weeks to get permission to  
20 share that research. It takes weeks to get an  
21 audience with the product teams, and especially if  
22 you wanted somebody who was in the position to  
23 actually make a change, it would take weeks to get  
24 onto their calendar for them to even be -- get  
25 access to this information.

*Document 75: Alison Lee Dep. Tr. at 94:18-25*

246. In fact, after the Haugen leak to the *Wall Street Journal*, all research was paused for a period of three months and then was subjected to “comms leadership” review.<sup>239</sup> The net effect of fast-paced rollouts and slow-paced integrity research is a vehicle with a gas pedal and no brakes

---

<sup>239</sup> Alison Lee Dep. Tr. at 117:22-25; *see also* Alison Lee Dep. Exhibit 8 at -8879

and it predictably leads to safety problems. An Insta presentation from Dr Lee had the following slide:

**We released a high volume (100+ in H1 2021) of product launches, 17 leading to integrity regressions.**

The pace at which these launches occur make it difficult for us to identify the additive, interaction or long-term integrity effects of each launch.

*Document 76: Alison Lee Deposition Exhibit 4 at Slide 5*

247. In H2, they launched over 200 products with 30+ regressions.<sup>240</sup> If each integrity regression is treated as a safety defect, this would yield a defect rate of approximately 17% which would be shockingly high for any consumer industry that typically sets defect standards between 1 in 100K to 1 in 1 million. For software defects that reach consumers, I could not find clear benchmarks although the Tability blog lists a change failure rate of 10-20% as average and best in class as less than 5%.<sup>241</sup> But Dr Lee in a October 2021 chat reports her “dismay” that Mosseri’s responded to a question about additional resources for integrity by saying that they were “doing enough” and then asking “how much is enough to invest into integrity?”<sup>242</sup>

248. For its part, Snap innovated the concept of “the streak” which it specifically designed to gamify its platform and drive usage and engagement. Streaks are built and maintained by two people “snapping” back and forth on consecutive days. The quality or content of the snap is irrelevant—it’s all about ping-ponging back and forth consistently. Streaks were extraordinarily effective. Less than a year after they were launched, 22% of users had at least one streak, and the

---

<sup>240</sup> Alison Lee Dep. Exhibit 5 at Slide 6

<sup>241</sup> *The 10 Best Metrics For Software Quality*, TABILITY, [https://www.tability.io/templates/m/X4kB\\_LA75HWq](https://www.tability.io/templates/m/X4kB_LA75HWq) (last accessed Apr. 16, 2025)

<sup>242</sup> Alison Lee Dep. Exhibit 12 at -1435

average highest streak value was 76.<sup>243</sup> Snapstreaks were also particularly popular with younger users. 47% Snapstreak users were under 17.<sup>244</sup> (The true number was likely even higher, given how easy Snap made it for children to lie about their age. *Infra* XII.P.iv.)

249. Snapstreaks' importance to users was also tied to Snapscores. While streak counts are private to the participants, every user of Snap has a numerical Snapscore displayed on their profile. The way that this score is generated is opaque, but broadly speaking it reflects how much they engage with the app and how much engagement they receive from others, including Streaks.<sup>245</sup> For users, high Snapscores are a way to "show off their popularity."<sup>246</sup>

250. Snap's research found that "users mainly used Streaks as a social status measure (to increase Snap Scores, have as many Streaks as possible) rather than a communication tool."<sup>247</sup> This is consistent with Snap's conclusion that once users knew how to increase their score, it "may make user care about (if not become addictive to) their [] score."<sup>248</sup> However, Snap saw this as a positive feature of Snapscores because it would "encourage [users] to produce and consume more snaps."<sup>249</sup>

251. But at the same time that Snap was chasing the engagement increases offered by streaks, worries about the addictive effect of Streaks on children were growing. For example, focus groups results forwarded by Rachel Racusen (Sr. Director of Corporate Communications and Public Affairs) to Jennifer Stout (VP of Global Public Policy) found the following:

---

<sup>243</sup> SNAP6759344, -9344

<sup>244</sup> SNAP6759344, -9344

<sup>245</sup> <https://help.snapchat.com/hc/en-us/articles/7012326657044-What-is-a-Snapscore>

<sup>246</sup> SNAP4279961 at -9963-65

<sup>247</sup> SNAP7148843 at 8847

<sup>248</sup> SNAP6118652

<sup>249</sup> SNAP6118652

- While not raised with the same level of concern as ephemerality and Snap Map, Snap Streaks are also mentioned as a feature that compounds parents' displeasure with their teens using the app.
- Snap Streaks are viewed as furthering the already strong grasp that technology broadly, and Snapchat specifically, has on their teen's focus and attention.
- Parents often mention their teens' seemingly uncontrollable need to "keep their streaks alive" associating the practice with ceaseless online communication and screen addiction.

*Document 77: SNAP1251784, -1784*

Another employee put it even more bluntly: “we seem to have tapped into some mass psychosis where 17 million people *must* keep the streaks going.”<sup>250</sup>

252. There was also desperate outreach from individual users:

On Aug 18, 2016 7:20 PM, [REDACTED] wrote:  
Hey all-  
  
Just wanted to bring this to your attention:  
[https://twitter.com/\[REDACTED\]/status/766455497268473856](https://twitter.com/[REDACTED]/status/766455497268473856)  
<https://snapchat.zendesk.com/agent/tickets/9663706>  
  
User has threatened to kill herself multiple times because her streak of 74 days broke. We followed up regarding the Snapstreak via Twitter and email, and we've sent her the suicide prevention macro via DM's.  
  
I'll let y'all know if the situation escalates at all.  
  
Thanks,  
BB

*Document 78: SNAP0857671, -7671*

253. A 2017 study of Snap “Power Users,” commissioned by Snap to better understand features that drove usage, found that “For some, streaks have become a “compulsive behavior” that they are “in too deep” with.”<sup>251</sup>

<sup>250</sup> SNAP6759344, -9344 (emphasis in original)

<sup>251</sup> SNAP0029949, -9959

**Streaks:** Use of streaks is extremely common and complicated for Power Users, who typically send both morning streaks and "goodnight" streaks. For some, streaks serve to help maintain connections among friends and to re-create real-life relationships in the app. However, for others it has become compulsive behavior, and many users feel they are "in too deep" to get out of a streak. While streaks are being held across ages, younger Power Users are slightly more inclined to use them. There is also strong social pressure to maintain a streak, and breaking a streak can negatively affect personal relationships. As part of maintaining streaks, users will allow their friends to access their accounts to keep the streak going in case they are unable to. Their friends will have access to their username and password, and will maintain a streak on behalf of the account owner. *"I have streaks over a year old but I messed one up because I was out of town I got confused by the timezones and I was really upset because it was hard work"*

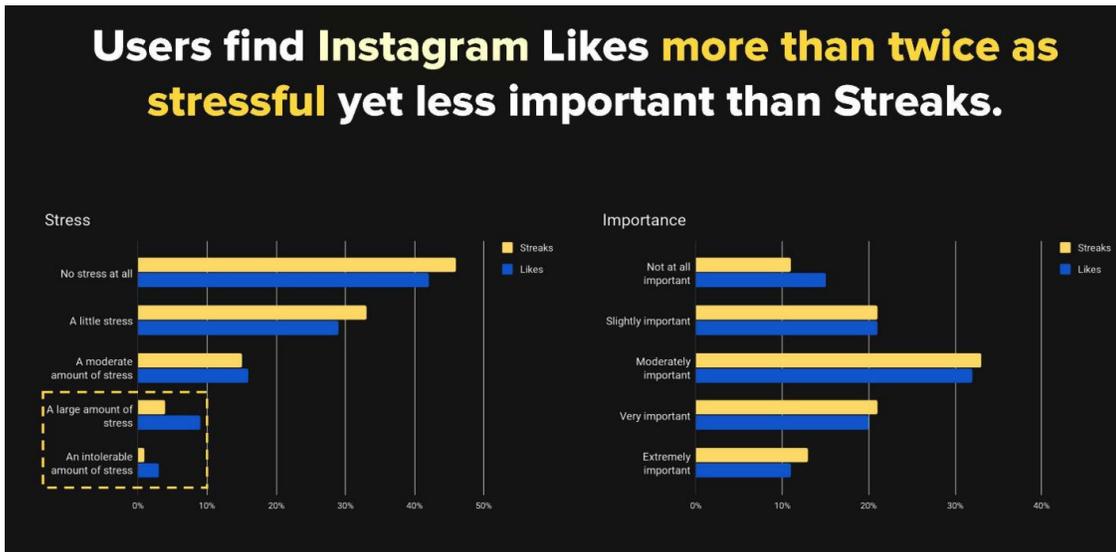
*Document 79: SNAP0666370, -6374*

Similarly, an internal presentation on streaks tried to spin them as a positive tool for building friendships but nonetheless was forced to observe that Streaks can be "really stressful" and "make[] it impossible to unplug for even a day."<sup>252</sup>

254. To better assess the harms of streaks, in 2018 Snap commissioned a survey of 790 13–24-year-old users (its core demographic).

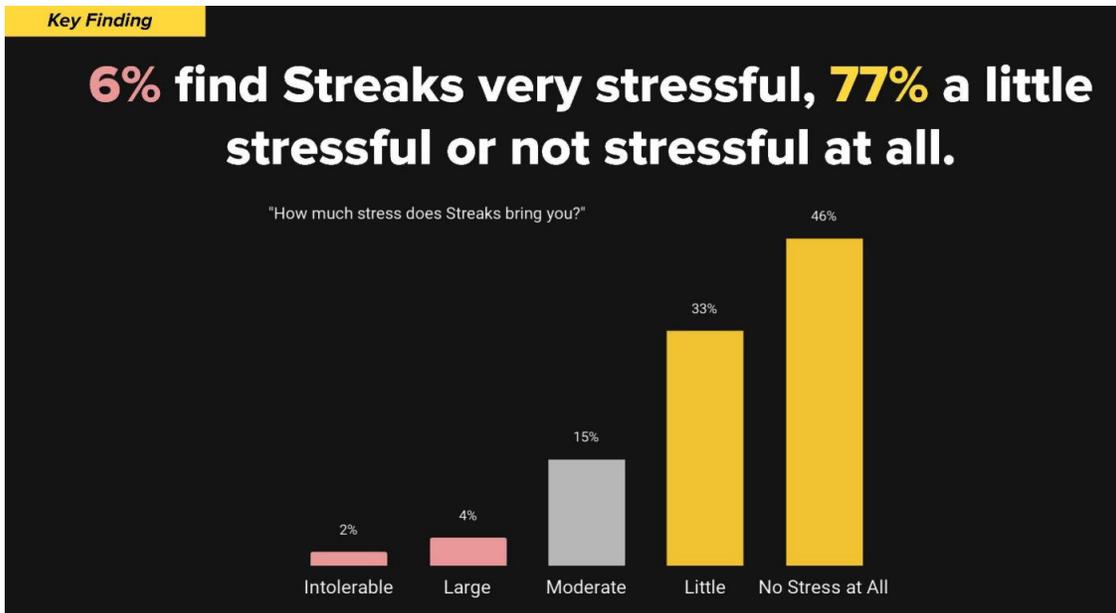
---

<sup>252</sup> SNAP2183204, -3272



Document 80:SNAP2183204, -3231

In Snap's telling, this study showed that that only a minimal number of users found streaks stressful.



Document 81: SNAP2183204, -3234

Snap presented this same conclusion to the United States Senate, writing that the study showed “the majority of our community did not indicate Streaks were a significant source of stress—but six percent did.”<sup>253</sup>

255. However, both the methodology and Snap’s spin on the results were deeply flawed. Given Snap’s millions of users, 6% still represents hundreds of thousands if not millions of users. And to get to 6%, Snap disregarded the “moderate” stress responses entirely. Snap also, without justification, combined those who found streaks a little stressful with those who found they caused no stress at all. A more accurate assessment of stress levels might be that 21% of users experience at least moderate stress because of snaps. A more holistic statement of the survey’s results would be that 54% of streak users found streaks at least a little stressful. The other problem with sampling “users” to assess experiential stress is what in epidemiological terms is called the “survivor effect.” Many people who found streaks intolerably stressful will have left the platform or opt not participate in a study about how stressful they are. In other words, this approach leads to a biased sample. Much like asking people still in a baseball stadium when the game has gone into 13 innings if games take too long will underestimate the true proportion of fans who think they do, asking regular users of an app how stressful it is underestimates the unease it induces. The study also depended on self-reporting from young people. But as Jennifer Stout, Snap’s Global Vice President for Public Policy, pointed out “Kids like a lot of dumb things and parents are always trying to regulate their activities for their own safety!”<sup>254</sup>

---

<sup>253</sup> SNAP0008117, -8119

<sup>254</sup> SNAP1251784, -1784

256. Snap’s spin on the 2018 study is further undermined by the fact that Snap employees, including Spiegel, continued to express concern that streaks were addictive and harmful for users.

**From:** Evan Spiegel [REDACTED]  
**Sent:** 8/7/2018 8:20:52 AM  
**To:** Bobby Murphy [REDACTED]  
**Subject:** Re: streaks

I definitely agree with your counterpoints and I think when used "as intended" streaks are a lot of fun and a healthy reinforcement of deep relationships. What happened in some cases, however, was that people felt pressured to keep a streak over a long period of time and that created anxiety - it also fueled a sense of competition among friends which isn't in keeping with our philosophy. I think minimizing the prominence of streaks has gone a long way towards solving these issues and we will continue to think about ways of rewarding the depth of a friendship...

*Document 82: SNAP0892766, -2766*

257. A 2023 document containing suggested answers for an employee Q&A with Speigel is clear that Streaks “can cause confusion and/or anxiety. We know this from the millions of support tickets we get every week from people asking us to restore a streak they accidentally lost.”<sup>255</sup>

258. Users themselves make clear that they found streaks addictive. As the Q&A answer explained, users’ frantic commitment to streaks could be tracked in part by tickets seeking to restore lost streaks.<sup>256</sup> But the answer actually understated how desperate users were to have their streaks restored. By 2021, Snap was receiving an average of four hundred thousand streak restore requests *a day*, making up 95% of the total volume of customer service contacts.<sup>257</sup> In some cases, individual users reach out directly to Spiegel to express the harms of streaks.

---

<sup>255</sup> SNAP1197331

<sup>256</sup> SNAP1197331

<sup>257</sup> SNAP0006256, -6256

From: Redacted--PII  
Date: Sun, Jun 5, 2022 at 3:20 PM  
Subject: Snapstreaks  
To: [REDACTED]

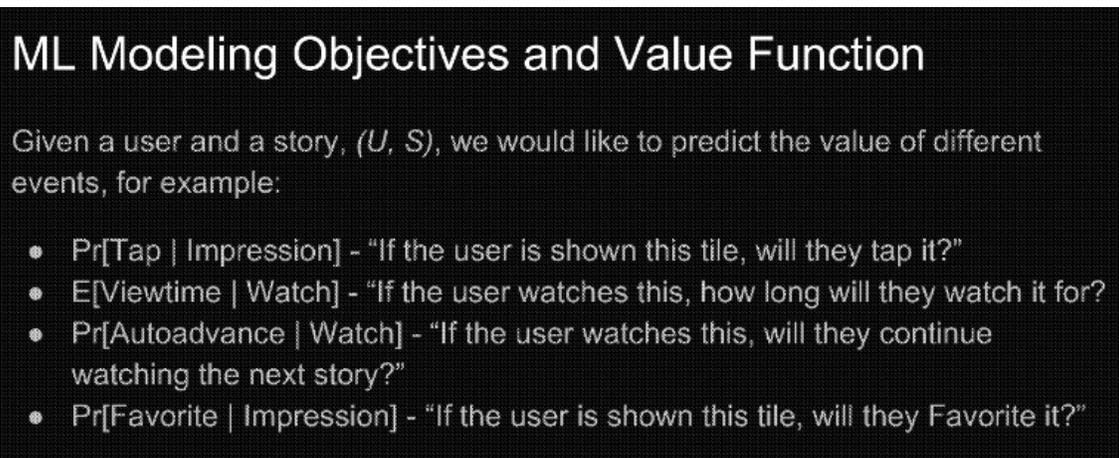
Hi Evan,

I am sorry for this direct email, but I wanted to ask you to consider some changes to Snapstreaks or just make you aware of the influence that this feature has on a generation of kids including my children. Snapchat is a great social media platform and a communication channel for millions of groups and individuals. I know how fun it is to use the different filters and other interactive features and how my children love reading new stories and snap with their friends. I am thankful for the fact that snaps are not stored but disappear after 10s (although this provides protection and risk at the same time), and I understand that this channel is growing stronger by the day. I have been leading IT teams for over 20 years, and I am still developing new applications, games and services to this day and I know how every feature is being tested and considered before going live.

The issue that we have as parents with the Snapstreaks feature, is the effect it has on the mental health of our children. Similar to what happened with 'likes' on Facebook, the anxiety of missing a streak means that millions of young kids feel that they 'must' login every day, even if for one minute, so they won't disappoint their friends. I have personally witnessed the distress and fearfulness and I am sure that this was not the intention of this feature although this is the reality.

*Document 83: SNAP1152337, -2337*

259. Snap's engagement strategy was not limited to streaks. It also conducted internal research to predict and maximize time on the app. The slides below show how they designed algorithms to predict and prioritize the likelihood of specific user behaviors:



**ML Modeling Objectives and Value Function**

Given a user and a story,  $(U, S)$ , we would like to predict the value of different events, for example:

- $\text{Pr}[\text{Tap} \mid \text{Impression}]$  - "If the user is shown this tile, will they tap it?"
- $\text{E}[\text{Viewtime} \mid \text{Watch}]$  - "If the user watches this, how long will they watch it for?"
- $\text{Pr}[\text{Autoadvance} \mid \text{Watch}]$  - "If the user watches this, will they continue watching the next story?"
- $\text{Pr}[\text{Favorite} \mid \text{Impression}]$  - "If the user is shown this tile, will they Favorite it?"

*Document 84: SNAP0224369, -4381*

# ML Modeling Objectives and Value Function

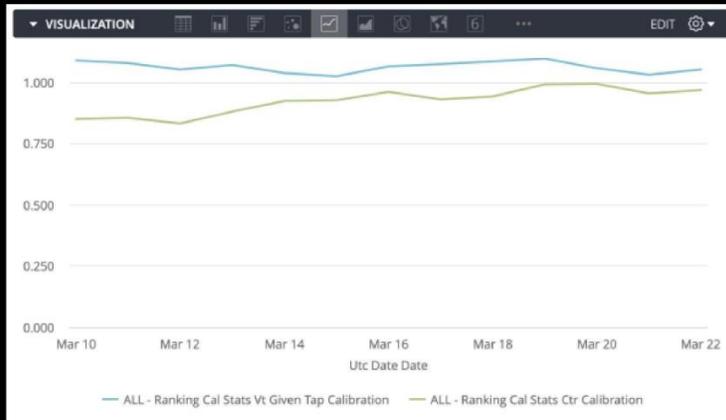
1  
2 These predictions are then linearly combined to produce a score value:

$$\begin{aligned} \text{Score} = & a * \text{Pr}[\text{tap} | \text{impression}] + \\ & b * E[\text{viewtime} | \text{impression}] + \\ & c * \text{Pr}[\text{autoadvance} | \text{impression}] + \\ & d * \text{Pr}[\text{favorite} | \text{impression}] + \dots \end{aligned}$$

Document 85: SNAP0224369, -4382

## ML Prediction - Calibrations

- Model predictions (CTR, predicted view-time given impression) are compared to observed user engagements
- Calibration metrics defined as  $\frac{\text{avg}(\text{predictions})}{\text{avg}(\text{observations})}$ 
  - The closer to 1.0 the better
  - Evaluated for specific models (and can be limited to specific countries/story types)
- Shown are results for the current production for-you model
- Exploration can be done at <https://go.sc-corp.net/dll>



Document 86: SNAP0224369, -4385

Tests showed that algorithmic ranking significantly increased both the numbers of stories viewed and the time spent viewing stories relative to just ordering them at random.<sup>258</sup>

<sup>258</sup> SNAP0224369, -4373

260. Snap also explicitly embraced the idea that social rewards were key to keeping users coming back for more, conducting research that “confirmed the hypothesis that Posters are motivated to post because audience feedback is their ROI reward.”<sup>259</sup>

### Sharing Bartering Posters are not “sharing” but rather bartering their content for feedback

This research confirmed the hypothesis that Posters are motivated to post because audience feedback is their ROI reward

- Users mentioned deleting content that did not elicit the feedback they were seeking
- When given a hypothetical scenario of getting no feedback whatsoever, some users mentioned feeling anxious believing they posted something wrong

*“Only time it doesn't feel rewarding if you post something and no one says anything”*

- Technically, this means users aren't “sharing” because the sharer should have a net loss when sharing (ex. *Sharing a stick of gum is -1 gum*)
  - Rather, posters are bartering their content in exchange for audience feedback/reaction (ex. I give you content that makes you feel something, you give me feedback that makes me feel something)

Takeaway

- Because feedback is the motivator/reward to post, it is imperative that we offer more feedback currencies that the audience can provide the poster (whether explicit or implicit)

*Document 87: SNAP4301491, -1500*

261. Indeed, Snap’s research “confirmed the *causal* relationship between receiving feedback (views and replies) and the poster propensity to post again.”<sup>260</sup> Following these conclusions, Snap’s product team proposed tweaking “Story reactions” and replies to generate more engagement and posting.

We're pushing in this direction because we believe that introducing a lightweight feedback mechanism (and generally increasing the volume of positive feedback posters receive) is the single most important thing we can do to grow Friend Story posting. [Recent analysis](#) shows that declining Friend Story availability (fewer posts) is the primary driver of the declines in consumption.

*Document 88: SNAP0467577, -7577*

---

<sup>259</sup> SNAP4301491, -1500

<sup>260</sup> SNAP0467577, -7579 (emphasis added).

262. And just as feedback made users feel good, Snap’s user research team found that not getting feedback was “discouraging.” Snap researchers found that 44% of Snapchat Story posters didn’t receive any feedback on a given day leaving them more anxious and worse off than not having posted at all.<sup>261</sup>

263. Senior Snap employees were clear-eyed about the consequences of manipulating users’ basic neurobiology and socioemotional responses. Responding to suggestions for ways to increase feedback rewards on Snapchat, CEO Evan Spiegel chimed in to say:

What is not discussed here is the research-based evidence that these sorts of mechanics are harmful for mental health which is one of the reasons we have not enabled this for friend stories on Snapchat. Feeling like you need to post "popular" and "likeable" content can actually contribute to reduced sharing in the future as the bar for sharing becomes higher.

*Document 89: SNAP0467577, -7578*

264. Similarly, when reviewing proposed changes to the way that streaks operate, Stephen Collins, a Director of Public Policy, expressed a similar sentiment, observing that “[r]ewards are known to drive compulsive/addictive behavior among some vulnerable groups.”<sup>262</sup> Of course, Collins’s solution to this was not to reduce the role that rewards-maximizing play in Snap’s design but just avoid using the word.<sup>263</sup>

265. Snap also used other forms of notifications to drive engagement. In 2023, Snap decided to implement an “always on badge” to get more users to try Spotlight.<sup>264</sup> This badge was a red dot that was always present on the spotlight icon, visually drawing users to it.<sup>265</sup> It would only go away once a visitor clicked on Spotlight and would reactivate every 60 minutes, regardless

---

<sup>261</sup> SNAP0467577, -7577

<sup>262</sup> SNAP4783191, -3196.

<sup>263</sup> SNAP4783191, -3196 (“I wouldn’t use the term ‘reward’.”).

<sup>264</sup> SNAP4227244

<sup>265</sup> SNAP4227244

of whether there was any new content or other change to actually notify a user of.<sup>266</sup> This one design change raised the daily active use of Spotlight from 53 million to 71 million, the single biggest increase in daily active use ever seen in an A/B test.<sup>267</sup>

266. Snap was not only made aware of the risks of these features, they were told to remove them by outside social media experts (SME's) they consulted.

Remove the infinity scroll – only allow a pre-set number of posts or amount of time  
Turn off push notifications so you don't open the app for every like or post  
Ability to refuse push notifications on the whole app

*Document 90: SNAP0404262, -4310*

However, not only did they fail to do so, but to my knowledge they made no effort to warn parents, teens, or the public of the risks posed.

267. Ultimately, every site adapted effective addictive and harmful design elements from their competitors. Below is a summary of key features and which platforms deploy them.

---

<sup>266</sup> SNAP4227244

<sup>267</sup> SNAP4227244

**Table 5: Summary of Harmful Design Features by Platform**

<b>Design Feature</b>	<b>Platform</b>
Infinite scroll	TikTok, Insta, FB, YT, Snap
Streaks	Snap, TikTok
Notifications	TikTok, Insta, FB, YT, Snap
FoMo	TikTok, Insta, FB, Snap, YT
Newsfeed prioritization	TikTok, Insta, FB, YT, Snap
Reels/Short form videos	TikTok, Insta, FB, YT, Snap
Variable intermittent rewards	TikTok, Insta, FB, YT, Snap
Likes	TikTok, Insta, FB, YT, Snap
Engagement Algorithm	TikTok, Insta, FB, YT, Snap
Video autoplay	TikTok, Insta, FB, YT, Snap
Appearance Filters/Negative Social Comparison	TikTok, Insta, FB, YT, Snap

268. It is my opinion to a reasonable degree of medical and scientific certainty that certain design features of social media increase usage, including problematic usage and addictive behavior. These design features include engagement algorithms, beauty or appearance filters, metrics such as the like button, comments, infinite scroll, and auto-play. The medical and academic community recognize the harm that flows from the use of these features. Similarly, there are ample Defendant documents providing additional support that these features increase usage, including problematic and addictive usage.

**XI. Social Media and Specific Harms**

269. I will now turn to a discussion of specific harms. Following the above framework, I will first discuss the literature regarding the relationship between social media and the harm

identified. I have not cited every single study reviewed, but rather focused on a synthesis of the totality of the evidence. I will then discuss exemplar internal documents that discuss that harm.

**A. Body Dysmorphic Disorder**

270. In this section, I will review the existing scientific literature supporting a causal relationship between social media use and body image and eating disorders mediated through the pathways of depression, body image, anxiety, and problematic/addictive use. As discussed before, problematic or addictive use remains relevant for these pathways as it drives additional time online (or on sites) which in turn drives the other downstream outcomes.

271. Body Dysmorphic Disorder (BDD) involves obsessive thoughts, repetitive behaviors, and mental acts in response to perceived appearance flaws and may focus on a particular feature of one's body (nose, hair, chin for example). The DSM-5 criteria for the diagnosis are below:

- A. Preoccupation with one or more perceived defects or flaws in physical appearance that are not observable or appear slight to others.
- B. At some point during the course of the disorder, the individual has performed repetitive behaviors (e.g., mirror checking, excessive grooming, skin picking, reassurance seeking) or mental acts (e.g., comparing his or her appearance with that of others) in response to the appearance concerns.
- C. The preoccupation causes clinically significant distress or impairment in social, occupational or other areas of functioning.
- D. The appearance preoccupation is not better explained by concerns with body fat or weight in an individual whose symptoms meet diagnostic criteria for an eating disorder.

*Document 91: DSM-5 Body Dysmorphic Disorder Criteria*

272. Eating Disorders (ED) involve disturbances in thoughts and behaviors related to eating, weight, and shape. Body Dysmorphic Disorder (BDD) and eating disorders, such as anorexia nervosa or bulimia nervosa, share similarities in that both involve a preoccupation with appearance and a distorted self-image. Individuals with BDD focus intensely on perceived flaws

in their physical appearance, often unrelated to weight, which may lead to compulsive behaviors like mirror checking or seeking cosmetic procedures. In contrast, eating disorders specifically center on weight, body shape, and food behaviors, with associated actions like extreme dieting, binge eating, or purging. While both conditions stem from deep psychological distress and can co-occur, their core focus differs: BDD is rooted in an obsession with minor or nonexistent physical imperfections, whereas eating disorders primarily involve concerns about weight and eating patterns. Both require specialized treatments, often involving therapy, to address underlying issues of self-esteem and anxiety.

273. Both ED and BDD cause considerable distress and dysfunction. In many cases, body dysmorphic disorder precedes the onset of eating disorder.<sup>268</sup> This finding suggests that body dysmorphic concerns may serve as a risk factor for the development of some eating disorders. Individuals with BDD and ED experience functional impairment in their daily lives but those with BDD often suffer more than those with ED. They have a higher rate of suicidality, including suicide ideation and suicide attempts, and more severe levels of depression.<sup>269</sup>

---

<sup>268</sup> Grant JE, Phillips KA. Is anorexia nervosa a subtype of body dysmorphic disorder? Probably not, but read on. *Harv Rev Psychiatry*. Mar-Apr 2004;12(2):123-6. doi:10.1080/10673220490447236

<sup>269</sup> See Krebs G, Fernández de la Cruz L, Rijdsdijk FV, et al. The association between body dysmorphic symptoms and suicidality among adolescents and young adults: a genetically informative study. *Psychol Med*. May 2022;52(7):1268-1276. doi:10.1017/s0033291720002998; Fennig S, Hadas A. Suicidal behavior and depression in adolescents with eating disorders. *Nord J Psychiatry*. 2010;64(1):32-9. doi:10.3109/08039480903265751; Rief W, Buhlmann U, Wilhelm S, Borkenhagen ADA, Brähler E. The prevalence of body dysmorphic disorder: a population-based survey. *Psychological Medicine*. 2006;36(6):877-885. doi:10.1017/S0033291706007264

274. Based on DSM-5 criteria, the lifetime prevalence of eating disorders is approximately 8% in girls and women and 2% in boys and men.<sup>270</sup> One metaanalysis estimated the prevalence of BDD at 11% but noted that there was considerable heterogeneity in the samples meaning that the estimates varied widely depending on the source population.<sup>271</sup> A single population-based study (using random digit dialing) in the U.S. puts the prevalence at about 4% and found it to be equally common in men and women.<sup>272</sup>

275. Evidence suggests that social media usage can increase the risk of or exacerbate existing negative body image body dysmorphia. With respect to body image in particular, the effects are driven by the reactions one gets (“likes” or comments) to posted images of oneself. The effects can be positive or negative. Multiple studies have examined the net effects of social media and body image.

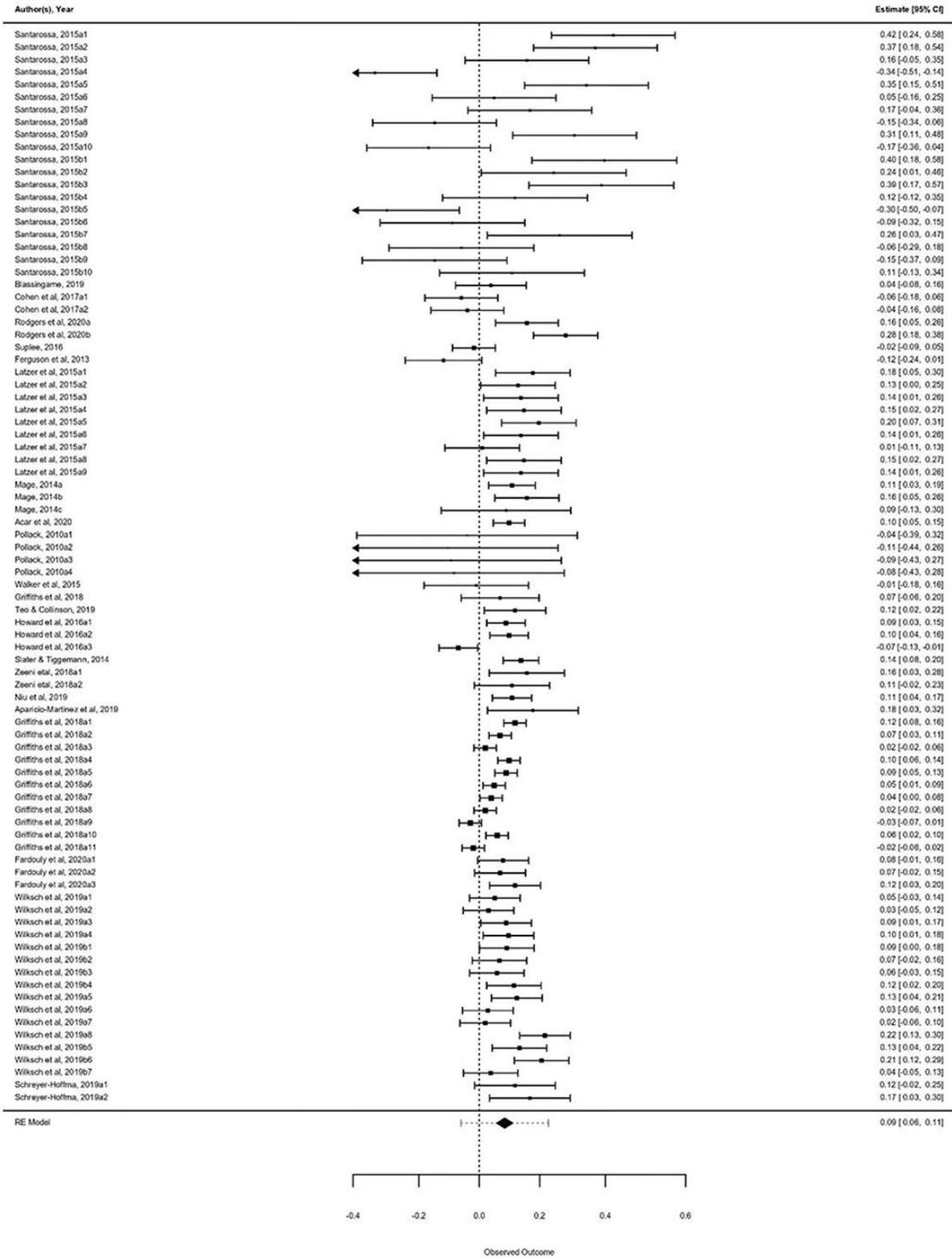
---

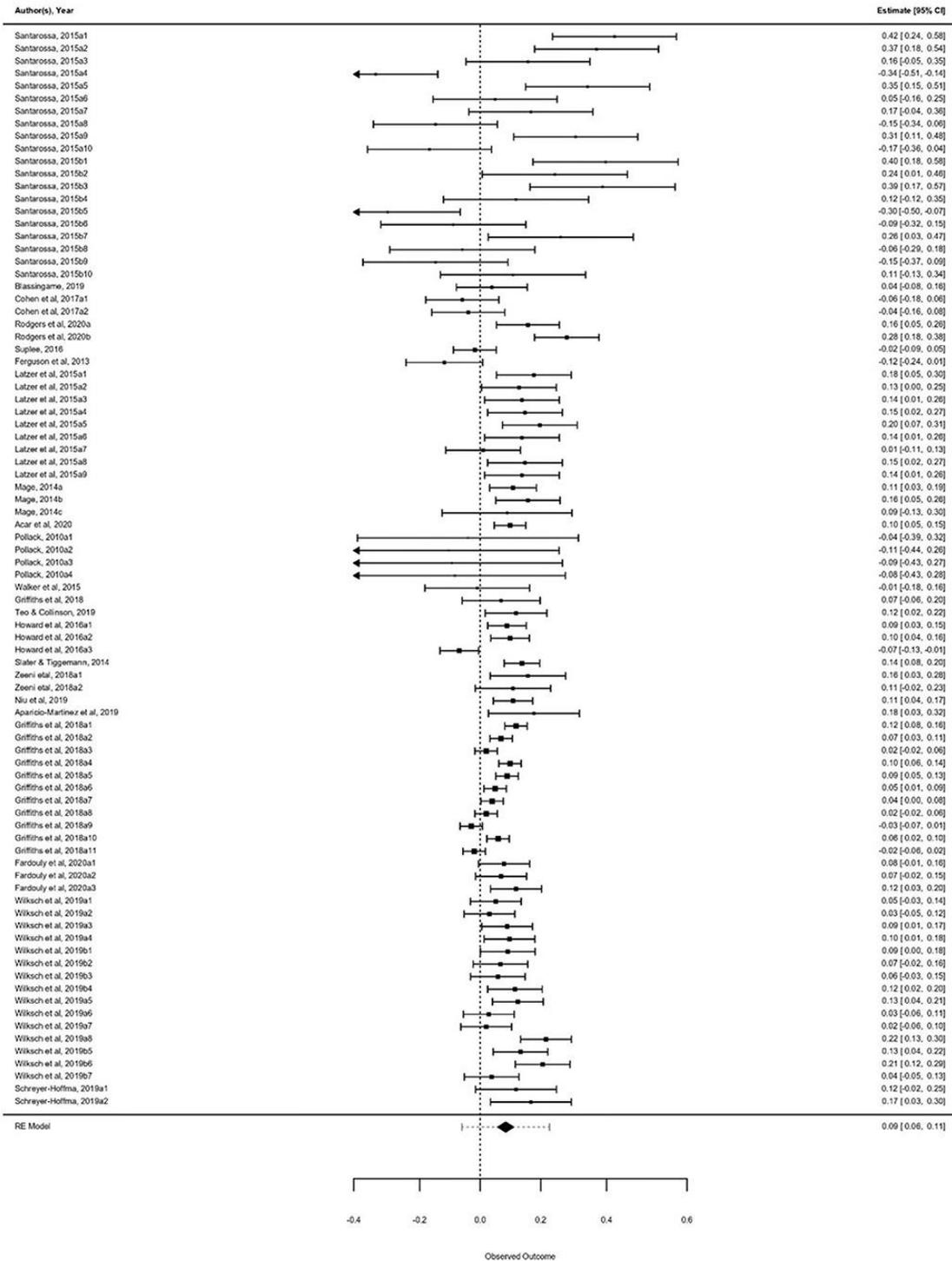
<sup>270</sup> Galmiche M, Déchelotte P, Lambert G, Tavoracci MP. Prevalence of eating disorders over the 2000–2018 period: a systematic literature review. *The American Journal of Clinical Nutrition*. 2019/05/01/ 2019;109(5):1402-1413. doi:<https://doi.org/10.1093/ajcn/nqy342>

<sup>271</sup> McGrath LR, Oey L, McDonald S, Berle D, Wootton BM. Prevalence of body dysmorphic disorder: A systematic review and meta-analysis. *Body Image*. 2023/09/01/ 2023;46:202-211. doi:<https://doi.org/10.1016/j.bodyim.2023.06.008>

<sup>272</sup> Koran LM, Abujaoude E, Large MD, Serpe RT. The Prevalence of Body Dysmorphic Disorder in the United States Adult Population. *CNS Spectrums*. 2008;13(4):316-322. doi:10.1017/S1092852900016436

Figure 29: Metanalysis of SM and Disordered Eating Behaviors





276. The Figure above is from a metaanalysis limited to the studies that allowed for pooling and combining data relating social media usage and body image problems. Although the figure is presented in a way that makes it too small to see the individual studies, all bars to the right of the central bar show a positive correlation between social media usage and BDD. The overall (summary) effect size was 0.11 (small).<sup>273</sup>

277. A more comprehensive systematic review of 40 studies examining the relationship between social media use and BDD found that the two are correlated.<sup>274</sup> The studies were too heterogenous to perform a metaanalysis and derive a summary estimate. A separate metaanalysis that evaluated 48 studies that experimentally manipulated social comparisons revealed a negative significant effect size of .24 (moderate) on body image wellbeing and self-esteem in spite of the heterogeneity of the included studies.<sup>275</sup>

278. An experimental manipulation of Instagram selfies conducted in 144 adolescent girls in which they were exposed to original or enhanced Instagram photos (see the figure below) found that body image and satisfaction were lower upon viewing the manipulated ones (effect size .17).<sup>276</sup>

---

<sup>273</sup> Zhang J, Wang Y, Li Q, Wu C. The Relationship Between SNS Usage and Disordered Eating Behaviors: A Meta-Analysis. *Front Psychol.* 2021;12:641919. doi:10.3389/fpsyg.2021.641919

<sup>274</sup> Ryding FC, Kuss DJ. The use of social networking sites, body image dissatisfaction, and body dysmorphic disorder: A systematic review of psychological research. *Psychology of Popular Media.* 2020;9(4):412-435. doi:10.1037/ppm0000264

<sup>275</sup> McComb C, Vanman E, Tobin S. A Meta-Analysis of the Effects of Social Media Exposure to Upward Comparison Targets on Self-Evaluations and Emotions. *Media Psychology.* 2023;26(5)

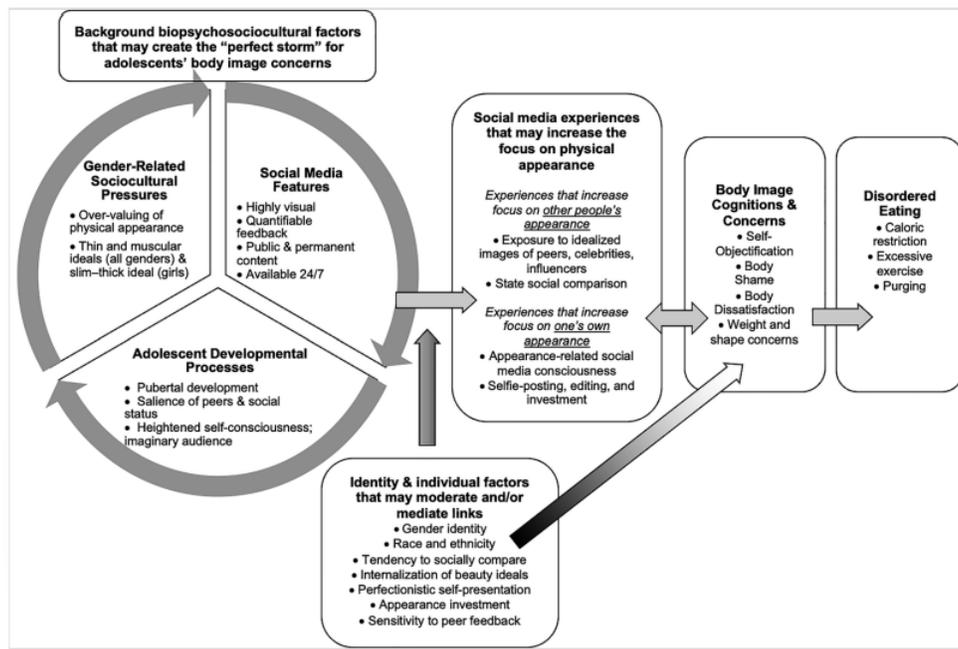
<sup>276</sup> Kleemans M, Daalmans S, Carbaat I, Anschütz D. Picture Perfect: The Direct Effect of Manipulated Instagram Photos on Body Image in Adolescent Girls. *Media Psychology.* 2018/01/02 2018;21(1):93-110. doi:10.1080/15213269.2016.1257392

**Figure 30: Real or Enhanced Instagram Images Presented to Teen Subjects**



279. The Handbook on Children and Screens included the following figure, which captures the relationship between social media features and the development of body image conditions and eating disorders:

Figure 31: Handbook on Children and Screens at 151



280. This figure, adapted from the version first published in Choukas-Bradley et. al. (2022) in *Clinical Child and Family Psychology Review*, indicates the perfect-storm of social media use, vulnerable adolescent development, and subsequent eating disorder and body image conditions.<sup>277</sup> Academics have recognized that specific features such as highly visual posts, quantifiable feedback or metrics (such as the like button), the public and permanent nature of social media, and social media's 24/7 availability all drive body image concerns. Notably, appearance and beauty filters are missing from this list, though they are recognized in the literature as contributing to body image issues and eating disorders.<sup>278</sup>

<sup>277</sup> Handbook of Children and Screens page 151.

<sup>278</sup> Fioravanti G, Bocci Benucci S, Ceragioli G, Casale S. How the exposure to beauty ideals on social networking sites influences body image: a systematic review of experimental studies. *Adolesc Res Rev.* 2022;7(3):419–58. <https://doi.org/10.1007/s40894-022-00179-4>.

281. Internal documents reflect that Defendants studied the effect of these design features on the development of body image issues and eating disorders. For example, Margaret Gould Stewart, “Head of Responsible Innovation” at Meta, testified regarding what experts they consulted told them about the relationship between appearance filters and potential harms as follows: “My recollection is that the significant majority of them confirmed our hypothesis that these had the potential to be very harmful, in particular to young people, and some of them specifically called out young women.” She went on to say the following:

My recollection, and I would need to reread the documents to get specific, but that the sense of negative comparison, so how I look versus how other people look, or the naturalistic presentation of myself versus the augmented, manipulated, affected view of myself over time, would feed into anxiety, body dysmorphia, depression related to one's appearance.<sup>279</sup>

282. Her recollection is accurate. All but one of the experts that Meta consulted stated that beauty filters that simulated plastic surgery would pose significant risks of negative outcomes especially for girls.<sup>280</sup> Internal Meta documents assert that “Filtering leads to unrealistic beauty standards, mental health issues, and a rise in plastic surgery.”<sup>281</sup> As discussed above, one of the mechanisms by which filters lead to adverse outcomes is through the phenomenon of social comparison. In January of 2020, Jennifer Guadagno, PhD, a lead researcher on Meta’s well-being team, states that “social comparison is the most important driver of well-being.”<sup>282</sup>

283. Dr. Guadagno commissioned her own independent analysis of filters, led by Dr. Diane Moscovitz, a PhD Psychologist at Duke and an expert on body image in teens. Her opinion, after reviewing the existing literature, is summarized below:

---

<sup>279</sup> Margaret Gould Stewart Dep. Tr. at 46:24-47:8

<sup>280</sup> Vaishnavi Jayakumar Dep. Exhibit 17 at -7136

<sup>281</sup> META3047MDL-020-00609936, -9937

<sup>282</sup> Jennifer Guadagno Deposition Exhibit 4 at -4847

#### SUMMARY

Selfie-manipulation is a new yet common, cross-cultural phenomenon. An analyses of the costs and benefits of editing selfies and viewing manipulated photos indicates the risks far outweigh the benefits. Research to date suggests these behaviors exacerbate risk and maintenance of several mental health concerns including body dissatisfaction, eating disorders and body dysmorphic disorder, among vulnerable populations (Kleemans et al., 2018, McLean et al., 2015, Mills et al., 2018, Rajanala et al., 2018). Data also indicates that editing selfies may have a paradoxical effect with regards to social connection. Rather than increasing acceptance, editing photos may actually increase social rejection (Vendemia & DeAndrea, 2018). Public outcry also indicates Facebook may alienate users by incorporating selfie-editing tools. Many argue photo-editing apps are harmful and culturally insensitive, (e.g., byrslf.co, 2016) and social media companies have received backlash for endorsing these westernized standards of beauty (Huffingtonpost.com, 2016 & 2017). Rather than bringing people together, selfie manipulation tools risk propagating unrealistic standards of beauty that are cross-culturally harmful and divide more than they unite.

*Document 92: Jennifer Guadagno Deposition Exhibit 11 at -6302*

284. In her deposition, Dr. Guadagno concurred that there are significant risks, stating, “I believe the main findings or some of the core summary was that social media can increase the risk of some of these concerns including poor body image.”<sup>283</sup> Despite this, Meta reversed its initial decision and allowed beauty or appearance filters.<sup>284</sup>

285. Interestingly, Dr. Guadagno’s team proposed an entire research plan (longitudinal or experimental) to test the effects of cosmetic surgery effects on body image. As stated previously, only industry could do the type of research study her team proposes, namely experimentally manipulating access to filters for a large group of social media users and following them prospectively over time.<sup>285</sup> It is not clear from the documents I reviewed why such studies were not done. However, once the decision to remove the filter ban was made, Dr. Guadagno “made the unpopular (among Responsible Innovation Team) decision to not have my team continue deep research in the area. Even though they were asking for it, I identified that the research requested would be significant work without a lot of gain and likely wouldn’t add anything that would change

---

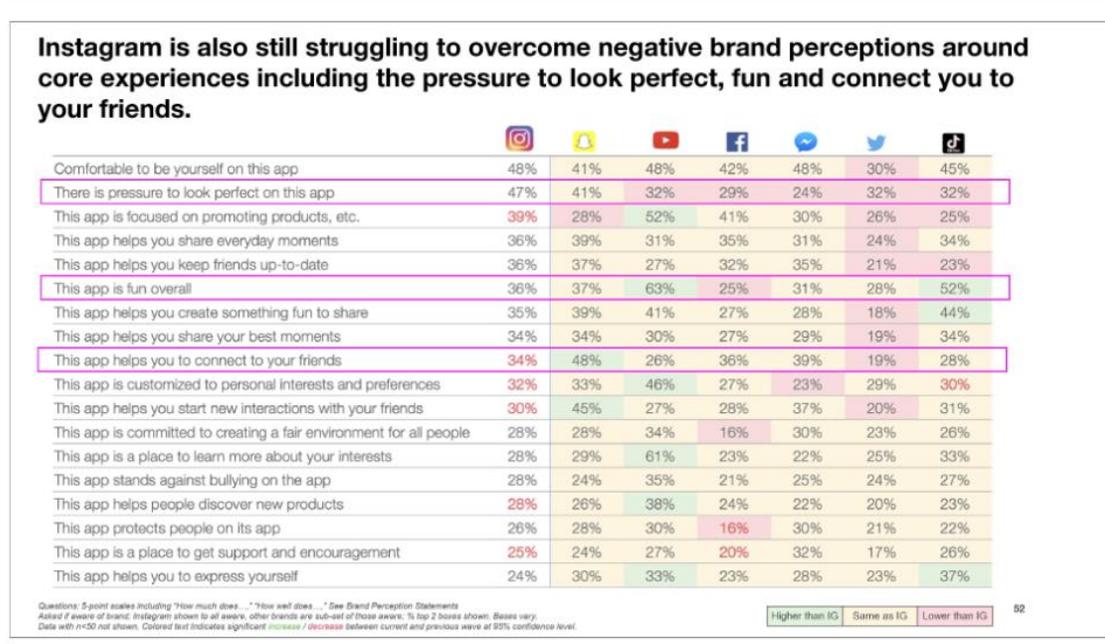
<sup>283</sup> Jennifer Guadagno Dep. Tr. at 136:10-13

<sup>284</sup> Vaishnavi Jayakumar Dep. Tr. at 103:8-11

<sup>285</sup> See e.g., Jennifer Guadagno Dep. Exhibit 16

leads' minds."<sup>286</sup> Implicit in her assertion is her belief that even robust experimental evidence showing harm would not be persuasive suggesting yet again that bottom line metrics trump well-being at Meta.

286. Meta did conduct a survey of teens' attitudes about Instagram and it revealed the following:

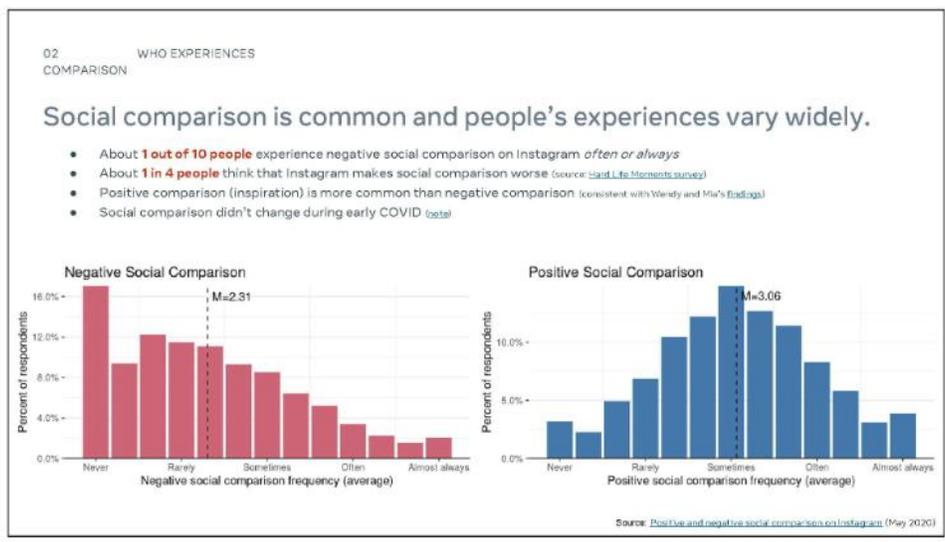


Document 93: META3047MDL-020-00350316, -0367

A sizeable percentage (for Instagram almost half) of teens report that “There is pressure to look perfect” on various social media apps and that pressure in turn motivates some to resort to cosmetic surgery effects.

287. The “underlying social comparison” that motivates users is common. In a global survey of 100K users of Instagram, Meta reported the following:

<sup>286</sup> META3047MDL-020-00265122, -5123



Document 94: META3047MDL-003-00001890, -1895

Ten percent of users experience negative social comparison on Instagram “often” or “always” and 25 percent think that Instagram makes social contagion worse.

288. Further, Diego Castaneda, in an email exchange with an Instagram analyst, reported the following:

Dr. Diego Emiliano Castaneda (8/31/2021 07:45:41 PDT):  
 >Hi Gargi- That sounds right. I'm looking at the previous leads review from two week ago (which is where I think you got it, right?) to make sure it's aligned.

Dr. Diego Emiliano Castaneda (8/31/2021 07:46:37 PDT):  
 >"Teens feel that aspects of Instagram exacerbate each other and lead to vicious cycles of scrolling and comparing, causing them to feel bad about themselves" this slide has similar language.  
[https://docs.google.com/presentation/d/1iQk88yc0rhT5wIqnjwYyVL8UBPo4a3wFHI5KDTsvDU/edit#slide=id.g8009939c5\\_3\\_451](https://docs.google.com/presentation/d/1iQk88yc0rhT5wIqnjwYyVL8UBPo4a3wFHI5KDTsvDU/edit#slide=id.g8009939c5_3_451)

Document 95: Diego Castaneda Deposition Exhibit 27 at -9652

289. Again, this is an explicit acknowledgement of the negative feedback loop that social comparison and algorithms can have on teens’ self-esteem and body image. Meta’s own analysis provides support that Instagram is worse than its competitors in fostering negative social comparison:

- **Social comparison is worse on Instagram.** It is perceived as real life, but based on celebrity standards. Explore and profile stalking enables never-ending rabbit holes. **Celebrity content** is more frequent but **friends' content** is more impactful in terms of social comparison.
- **Other apps are shielded by fun filters.** TikTok is grounded in dance, fun. Snapchat is sheltered by the element of fun that keeps focus on the face and not the body and sharing with close friends. At the other extreme, VSCO is a detached dream that is 100% edited.

*Document 96: Haugen\_00015958, -5964*

290. Concerns over the potential harms of cosmetic surgery effects, especially to teenage girls, led Meta to temporarily ban them in 2019 motivated in part by recommendations from outside academic experts, including a psychologist at Duke University.<sup>287</sup> But soon thereafter, a movement to lift the bans emerged motivated by the risk bans posed to adoption of the platform, a movement which was opposed by Sheryl Sandberg herself in her email below:

**From:** Sheryl Sandberg <sheryl@fb.com>  
**Sent:** Monday, March 30, 2020 9:40:18 PM  
**To:** Margaret Stewart <margarets@fb.com>  
**Subject:** Re: Meeting on cosmetic surgery effects

I am strongly against lifting this ban. That is why I am joining the meeting

*Document 97: Margaret Gould Stewart Deposition Exhibit 13*

291. Stewart's concerns about cosmetic surgery effects emerged in part from her experience as a parent. She stated "Yes, I feel that raising teenagers during that team [sic] period gave me a perspective on why we needed to be looking at this issue because of my personal proximity to it."<sup>288</sup> And later, "As a parent of two teenage girls, I can tell you the pressure on them

<sup>287</sup> Margaret Gould Stewart Dep. Tr. at 67:9-13; *see also* Margaret Gould Stewart Dep. Exhibit 9

<sup>288</sup> Margaret Gould Stewart Dep. Tr. at 34:3-7

and their peers coming through social media is intense with respect to body image.”<sup>289</sup> Her concerns were shared by other tech executives. As reported in *The Atlantic*:

Steve Jobs limited his children’s use of technology. TikTok CEO Shou Zi Chew doesn’t let his children on TikTok. Bill Gates restricted his kids’ screen time and did not give them a phone until they were 14. Google CEO Sundar Pichai didn’t give his 11-year-old a phone. Mark Zuckerberg has carefully monitored his kids’ screen time and avoided sharing identifying photos of them on Instagram. Snap CEO Evan Spiegel limited his 7-year-old’s technology use to 90 minutes a week.<sup>290</sup>

292. Ms. Stewart’s and Ms. Sandberg’s concerns were confirmed by a consultation with Google which Ms. Stewart forwarded in a group email:

**Google Engagement:** On Dec 10, we met with representatives from Google’s well-being research team, camera engineers and PMs. They have been investing heavily in research in camera beautification and their conclusions have been similar to ours - highlighting that global experts believe this to be a mental health risk. They’re acting on these findings and their focus is on user agency in deciding when and if to engage with effects and camera editing tools, particularly at the device and OS level (e.g. Pixel’s camera; Android OS controls.)

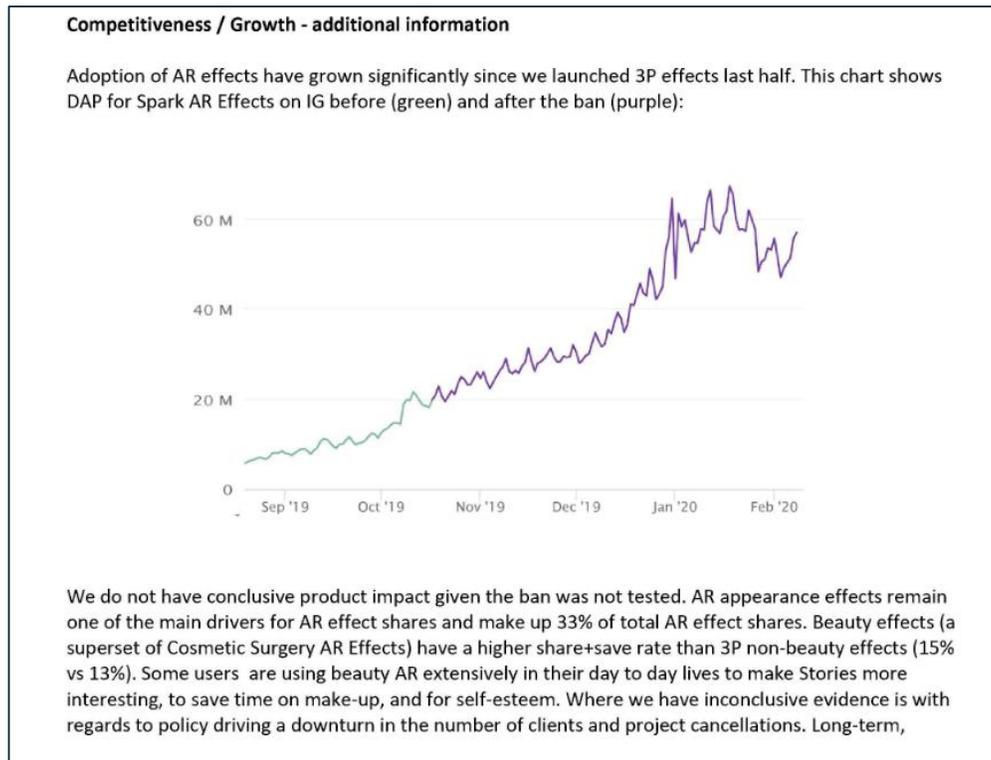
*Document 98: Margaret Gould Stewart Deposition Exhibit 22 at -8779*

293. Leaning against the ban were data presented on the adoption effects of Augmented Reality (AR) and how the ban might impact growth and engagement with the platform. The data are below:

---

<sup>289</sup> Margaret Gould Stewart Dep. Tr. at 58:22-25; *see also* Margaret Gould Stewart Dep. Exhibit 8

<sup>290</sup> Rausch Z, Haidt J, Torres L, *Social-Media Companies’ Worst Argument*, THE ATLANTIC (Sep. 12, 2024), <https://www.theatlantic.com/ideas/archive/2024/09/social-media-lgbtq-teens-harms/679798/>.



*Document 99: Margaret Gould Stewart Deposition Exhibit 6 at -9488*

Although as the report acknowledges, the data does not assess what the effect of the ban might be, they suggest that augmented reality is a major driver of Instagram uptake. The decision to lift this ban was allegedly made by Mr. Zuckerberg himself in May of 2020.<sup>291</sup>

294. For its part, Snap also researched the effect and reach of its filters using a sample of 13–25-year-old users. As part of that study, Snap asked users to explain “why they felt they looked best in Snap Chat Camera.”<sup>292</sup> Users overwhelmingly listed “Beautification lenses” as the reason.<sup>293</sup> The same report also found:

<sup>291</sup> META3047MDL-014-00053599, -3600.

<sup>292</sup> SNAP0640776, -0777

<sup>293</sup> SNAP0640776, -0777

- Many of these open-end responses were troubling, with responses like “Bc the filters actually make me look decent instead of being a horrendous ugly black girl,” “because it hides blemishes and makes my face look thinner and lighter than it is,” and “I love the filters it hides my ugliness”
- Users were very quick to point to their own perceived physical ‘flaws’ with some highlighting the benefit of looking lighter in a Lens. While it is a positive users feel confident when using Lenses, users are simultaneously left feeling bad about how they look without one; this underlines a key issue mentioned in research with employees around ‘Lens Dysmorphia’

*Document 100: SNAP0640776, -0777*

295. Beyond the effects that cosmetic surgery effects and AR might have on BDD, Instagram also explored the role that “likes” might play in exacerbating negative social comparison.<sup>294</sup> In 2019, Instagram launched “Project Daisy” which as originally proposed would test turning off the “like count” in general for all users **PRIV** but in the end this was made a setting people could opt into rather than a default.<sup>295</sup>

296. The linkage between social media use and BDD is highly psychologically plausible. Body dissatisfaction has long been recognized as influenced by a variety of sociocultural factors including media. Historically, this has happened in the “real” world or via magazine advertisements or television programs. Prior to the widespread use of social media, media exposure, and in particular exposure to advertisements was associated with BDD. A 2008 meta-analysis of 77 studies (predating social media) found small to moderate correlations between media exposure and body dissatisfaction.<sup>296</sup> Meta’s own research concludes/accepts this, stating:

The media has long held a role in establishing and perpetuating “standards” of beauty and attractiveness, communicating how we *should* look and act to gain

<sup>294</sup> Vaishnavi Jayakumar Dep. Tr. at 167:5

<sup>295</sup> Vaishnavi Jayakumar Dep. Tr. at 167

<sup>296</sup> Grabe S, Ward LM, Hyde JS. The role of the media in body image concerns among women: a meta-analysis of experimental and correlational studies. *Psychol Bull.* May 2008;134(3):460-76. doi:10.1037/0033-2909.134.3.460

acceptance and find happiness. **These message about ideal attractiveness pose a significant risk for body dissatisfaction, eating disorder behaviors, depression and anxiety among vulnerable populations who adopt these standards as their own.**<sup>297</sup>

In addition, internal emails referencing research done by Meta with their own data reports the following.<sup>298</sup> (Their data included ~6,000 respondents in 7 countries, matched with their log data, something no independent scientist could do.)

297. Social comparison is common on Instagram. 51% of people experience social comparison on Instagram. They either compare their accomplishments to others or observe other people to decide how they should act “sometimes” or “more often”;<sup>299</sup>

298. Women and teens are more prone to social comparisons, especially negative social comparison;<sup>300</sup>

299. 33% of people have been feeling worse about themselves on Instagram for “several months to a year;”<sup>301</sup>

300. Women on average engage in more social comparison then men (53% vs. 43%) and those comparisons on average make them feel worse about themselves whereas they make men feel better about themselves. The precise percentage difference is not discernable because of the low-resolution graphic in the document;<sup>302</sup>

301. The negative effects of comparison for females are considerably greater for teenagers. Again, precise differences are not discernable;<sup>303</sup>

---

<sup>297</sup> META3047MDL-014-00376298 (emphasis added)

<sup>298</sup> Haugen\_00000797

<sup>299</sup> Haugen\_00000797, -0797

<sup>300</sup> Haugen\_00000797, -0797

<sup>301</sup> Haugen\_00000797

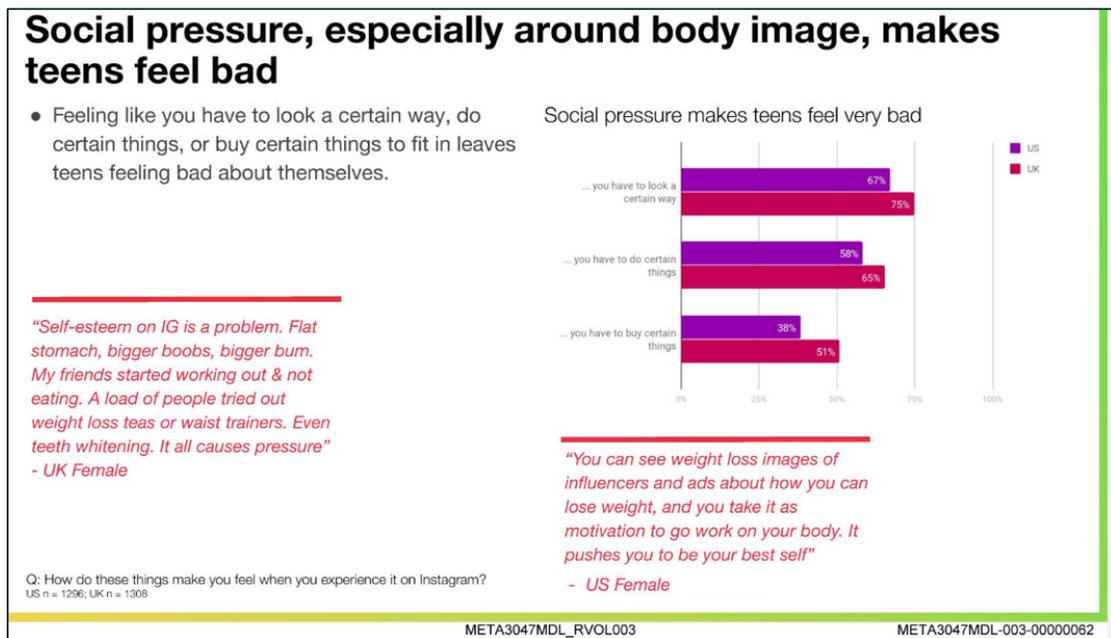
<sup>302</sup> Haugen\_00000797, -0822

<sup>303</sup> Haugen\_00000797, -0823

302. A logistic regression of their data shows that 13-17 year olds (OR 4.4), 18–24-year-olds (OR 2.0), and females (OR 1.8) all have statistically increased odds of negative comparisons; and<sup>304</sup>

303. 1 in 3 teen girls, according to Meta’s own data, report that Instagram use makes their body image issues worse.<sup>305</sup>

304. And a separate Meta survey of 2,500 13–17-year-olds in the US and UK reported the following:



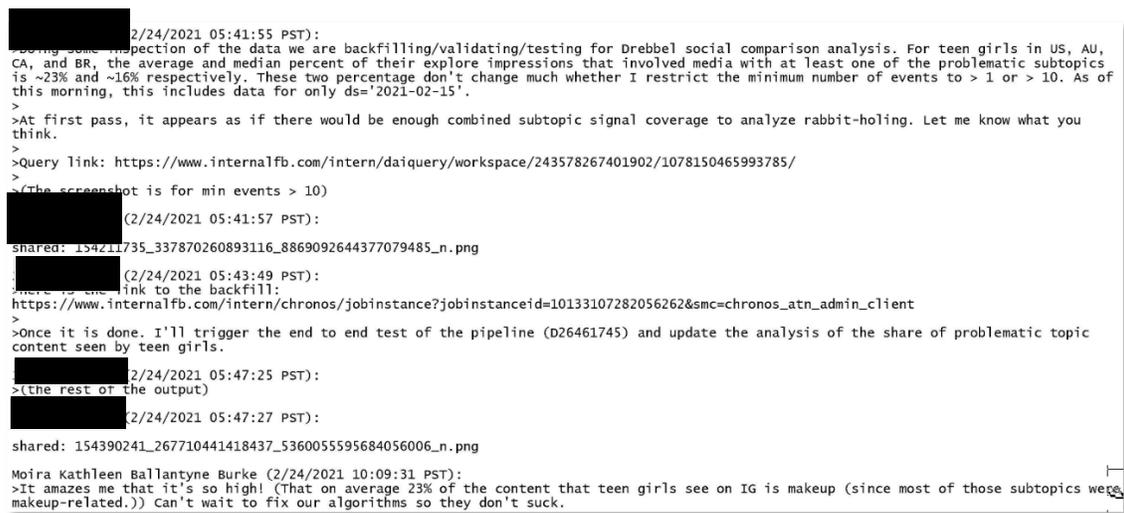
Document 101: META3047MDL-003-00000029, -0063

305. In 2018, [REDACTED] emailed Moira Burke, a Research Scientist, and the “Well-being (+Meaningful Interactions & Agency/Control) Research XFN” team, to propose: “What if IG suggests photo filters selectively, e.g. only to certain types of photos such as landscape but not

<sup>304</sup> Haugen\_00000797, -0828

<sup>305</sup> Haugen\_00016707

to people portraits (major trigger for social comparison)? And, “What if we launch a campaign such as “No Filter Friday,” #beyourtrueself to promote authentic expression on IG”?<sup>306</sup> Both of these proposed solutions, neither of which to my knowledge were implemented, may, based on available evidence, have mitigated the untoward effects on body image. But more to the point, they demonstrate that Meta was at least aware of, and one researcher acknowledged, the problem they were exacerbating. And it is something their platform, specifically their algorithms, is responsible for. Below is an internal text exchange between [REDACTED] (IG Well-Being Sr. Data Scientist) and Dr. Burke:



[REDACTED] (2/24/2021 05:41:55 PST):  
>Doing some inspection of the data we are backfilling/validating/testing for Drebbel social comparison analysis. For teen girls in US, AU, CA, and BR, the average and median percent of their explore impressions that involved media with at least one of the problematic subtopics is ~23% and ~16% respectively. These two percentages don't change much whether I restrict the minimum number of events to > 1 or > 10. As of this morning, this includes data for only ds='2021-02-15'.  
>  
>At first pass, it appears as if there would be enough combined subtopic signal coverage to analyze rabbit-holing. Let me know what you think.  
>  
>Query link: <https://www.internalfb.com/intern/daiquery/workspace/243578267401902/1078150465993785/>  
>  
>(The screenshot is for min events > 10)

[REDACTED] (2/24/2021 05:41:57 PST):  
shared: 154211735\_337870260893116\_8869092644377079485\_n.png

[REDACTED] (2/24/2021 05:43:49 PST):  
>Here's the link to the backfill:  
[https://www.internalfb.com/intern/chronos/jobinstance?jobinstanceid=10133107282056262&smc=chronos\\_atn\\_admin\\_client](https://www.internalfb.com/intern/chronos/jobinstance?jobinstanceid=10133107282056262&smc=chronos_atn_admin_client)  
>  
>Once it is done, I'll trigger the end to end test of the pipeline (D26461745) and update the analysis of the share of problematic topic content seen by teen girls.

[REDACTED] (2/24/2021 05:47:25 PST):  
>(The rest of the output)

[REDACTED] (2/24/2021 05:47:27 PST):  
shared: 154390241\_267710441418437\_5360055595684056006\_n.png

Moira Kathleen Ballantyne Burke (2/24/2021 10:09:31 PST):  
>It amazes me that it's so high! (That on average 23% of the content that teen girls see on IG is makeup (since most of those subtopics were makeup-related.)) Can't wait to fix our algorithms so they don't suck.

*Document 102: META3047MDL-003-00123974, -3974*

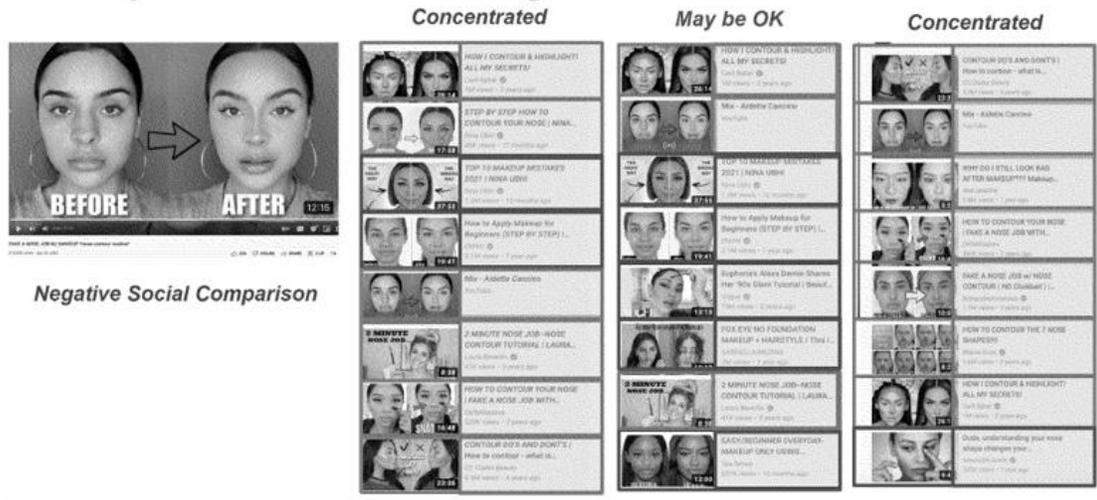
In response to the content that is being feed to teen girls as a result of their searches, Burke responds that she “can’t wait to fix our algorithms so they don’t suck.” As a slide from Alison Lee’s deposition asserts, “Prior research on rabbit holing suggests that particular kinds of accounts and topics may serve as gateway entities for falling into problematic Rabbit holes.”<sup>307</sup>

<sup>306</sup> META3047MDL-014-00035646, -5646

<sup>307</sup> Alison Lee Dep. Exhibit 28 at Slide 2

306. For its part, YouTube also recognized a design feature that amplified content and creates rabbit holes that could be harmful.

## A feed in a teen session can have a high volume of videos that repeat the same message



Document 103: GOOG-3047MDL-01372619

The notes to this slide state “Once you watch a few of these, your feed might become concentrated w/ a high volume of content that repeats the same message,” which is to say that the algorithm itself creates, perpetuates, and exacerbates the problem.

307. From a scientific standpoint, there is consistency across multiple studies as the overwhelming majority, but not all of them find an association between social media and BDD. Although the preponderance of existing research linking social media to BDD is cross-sectional and relies on self-report of social media usage, there have been some longitudinal and laboratory based experimental studies. Much like smoking, experimentally testing whether social media exposure leads to BDD is impractical to the point of impossibility. Accordingly, experimental manipulations typically bring subjects into a laboratory, randomize them to see (or not see) various

idealized images of themselves or others, and conduct pre- and post-exposure assessments of how they feel about their bodies comparing their feelings before and after.

308. For example, one laboratory study randomized 130 undergraduate female students to see 18 Instagram images. The intervention group saw “fitspiration” ones of women in fitness clothing or engaging in exercise and the control arm saw women at travel destinations. Mood and body dissatisfaction as well as self-esteem were measured both at baseline and after the exposures. Both body dissatisfaction and self-esteem were significantly lower in the “fitspiration” group.<sup>308</sup> A systematic review (again a metaanalysis was not feasible) of 43 experimental studies found moderate to large effects of such approaches.<sup>309</sup> Meta’s own documents refer to (and do not dispute) the existence of “snapchat dysmorphia” wherein affected individuals present idealized selfies of themselves to plastic surgeons and ask to be made to look like them.<sup>310</sup> And the Haugen documents, detailing their study of ~6000 users of IG, found that “Beauty, Fitness, and Fashion are the top three contents that trigger negative comparisons for women.”<sup>311</sup>

309. There also appears to be a dose response relationship: the more time spent on social media sites, the greater the risk of BDD. Studies that assessed the correlation between appearance-focused social media vs. general social media usage found larger effect sizes ( $r=0.31$  vs  $0.11$ ).<sup>312</sup> While problematic, idealized images have existed in media for decades, the advent of social media

---

<sup>308</sup> Tiggemann M, Zaccardo M. “Exercise to be fit, not skinny”: The effect of fitspiration imagery on women's body image. *Body Image*. 2015/09/01/ 2015;15:61-67. doi:<https://doi.org/10.1016/j.bodyim.2015.06.003>

<sup>309</sup> Fioravanti G, Bocci Benucci S, Ceragioli G, Casale S. How the exposure to beauty ideals on social networking sites influences body image: A systematic review of experimental studies. *Adolescent Research Review*. 2022:No Pagination Specified-No Pagination Specified. doi:10.1007/s40894-022-00179-4

<sup>310</sup> META3047MDL-014-00376300

<sup>311</sup> Haugen\_00000797, -0834

<sup>312</sup> Saiphoo AN, Vahedi Z. A meta-analytic review of the relationship between social media use and body image disturbance. *Comput Hum Behav*. 2019;101:259-275.

– with features that promote social comparison, 24/7 access, filters, and engagement algorithms that create rabbit holes – is particularly harmful to mental health. In other words, while the images are problematic, their capacity to cause harm is greatly increased by the characteristics and algorithms of social media.

310. Finally, there appears to be some ability to reverse or reduce the risk by mediating social media usage. A systematic review of media literacy programs’ effects on body dissatisfaction found some modest evidence of benefit. In particular, training adolescents to recognize that online portrayals are unrealistic or even doctored reduced the risk of BDD.<sup>313</sup> In addition, a cornerstone of effective BDD psychotherapeutic treatment is Cognitive Behavioral Therapy which, among other things, counsels subjects to be aware of and actively mitigate the effects SM has on them.

311. An internal Facebook email from 2019 states:

**Topline Research and Outreach:** While the research in this field is new and there is not yet robust causal evidence, experts and academic research from around the world -- APAC included -- generally agree that these effects are cause for concern for mental health and wellbeing and there’s strong support for us to get out in front of the issue and be industry leaders. The Spark creators from APAC were largely accepting of the policy, though recognized the potential difficulty in enforcing it.

*Document 104: Margaret Gould Stewart Deposition Exhibit 22*

In other words, Facebook’s own “take” on the academic consensus was that there was “cause for concern” based on the use of cosmetic surgery effects. In fact, even as they state internally that we don’t “yet” know if high Negative Appearance Comparison (NAC) content (content that promotes NAC in teens) is causally related to teens experiencing NAC, an internal document states:

---

<sup>313</sup> McLean SA, Paxton SJ, Wertheim EH. The role of media literacy in body dissatisfaction and disordered eating: A systematic review. *Body Image*. Dec 2016;19:9-23. doi:10.1016/j.bodyim.2016.08.002

**We don't yet know.** But we believe it's very likely that seeing High-NAC content causally contributes to experiencing NAC, and that reducing exposure to High-NAC content will result in reducing experiences of NAC.

*Document 105: META3047MDL-019-00066693 at Slide 11*

312. If they “believe” that there is a causal relationship between seeing “negative appearance content” and experiencing it, something the preponderance of the scientific literature supports, they surely were obligated to act on it. But I did not find credible evidence that they did. Furthermore, the same report acknowledges that “High NAC content is prevalent, and our systems make it more common.”<sup>314</sup> The net result, per Meta, is that “High NAC is 18% of what teen girls and women see on IG.” And then it states:

**- The top accounts that we recommend to new users produce more High-NAC content than the top accounts we recommend to existing users (19% vs. 13%). We may be setting up new users to greater High-NAC exposure because popularity is the key factor in the algorithm.**

*Document 106: META3047MDL-019-00066693 at Slide 12*

313. Meta implicates its own algorithm in promoting NAC content. The content is ubiquitous, but the unique and effective means of promoting it – feeding it – is their own invention.

---

<sup>314</sup> META3047MDL-019-00066693, Slide 12

Furthermore, Meta’s own assessment determined that 23% of 13–15-year-olds have felt worse about themselves because of people’s posts on Instagram.<sup>315</sup>

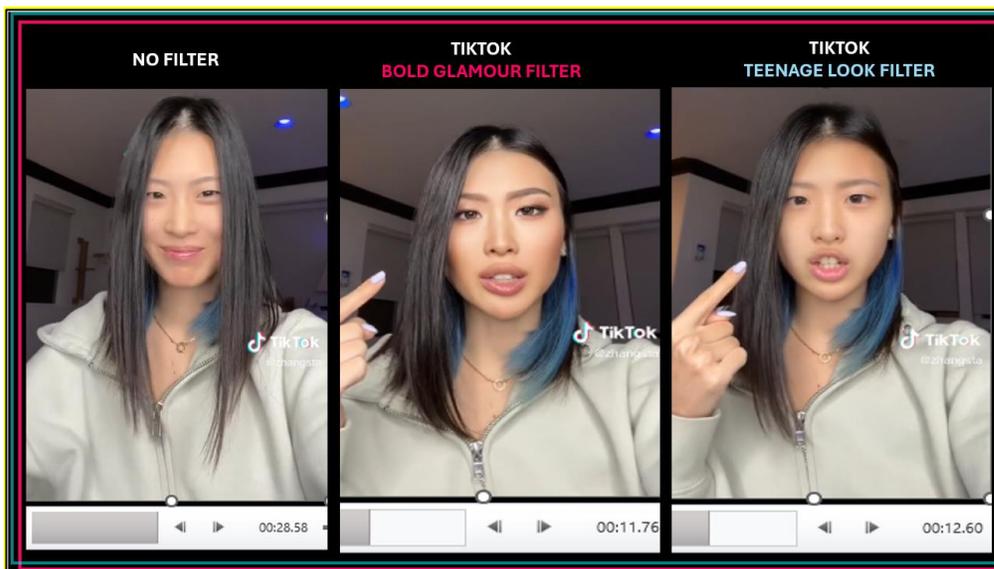
314. In 2023, in spite of the public backlash for Meta’s filters, TikTok launched its own image enhancing product, “The Bold Glamour Filter.” It too faced harsh criticism:

\* TikToker Rosaura Alvarez called the beauty filter “a problem,” and called out TikTok for taking it “a little too far.” In her video, which now has 8.7 million views, Alvarez says, “you can’t even tell it’s a filter anymore!” “As someone who experienced body dysmorphia growing up, this makes me sick to my stomach,” read the caption of her video. “TikTok u can’t be enabling this...it’s sickening for our youth.”

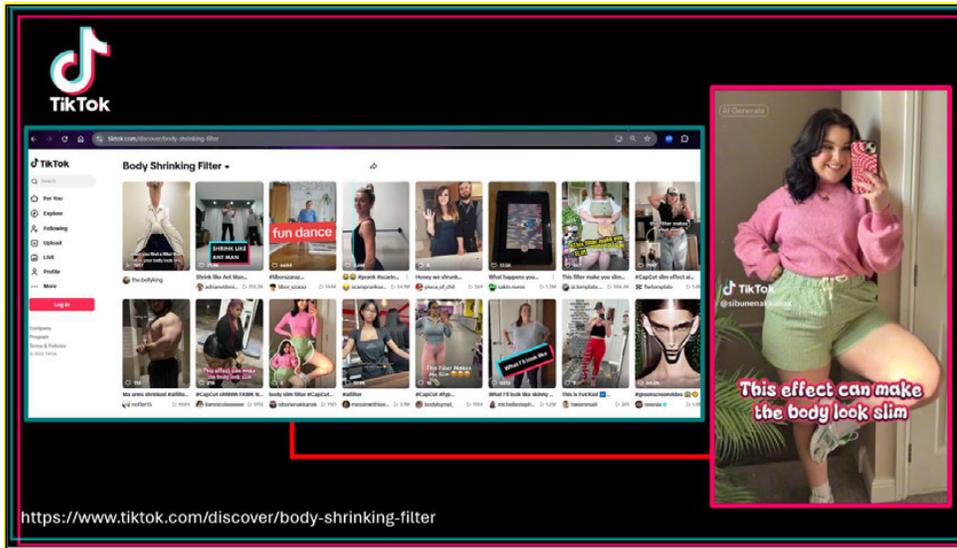
\* The Bold Glamour filter has also been described as [ HYPERLINK "<https://www.thetimes.co.uk/article/the-ugly-truth-of-tiktoks-new-beauty-filter-jlp9nvlmw>" \h ] for setting unrealistic beauty standards amongst young women and girls.

*Document 107:TIKTOK3047MDL-056-00987601*

315. Three particularly problematic filters, illustrated below, included “teenage look” filter, “bold glamour” filter, and “body shrinking” filter.



<sup>315</sup> META3047MDL-031-00118103, -8104



*Document 108: Jordan Furlong Dep. Exhibits 25 and 28*

316. Based upon a review of internal TikTok documents note, it appears that “Beauty Retouching” effects (including those that mimicked “minimally-invasive cosmetic procedure[s]” such as “skin smoothing and retexturing,” “teeth whitening,” and “undereye lightening”) were defaulted to “on.”<sup>316</sup> The document goes on to warn that these decisions “directly contradict expert recommendations” that were “among the most unanimous set of recommendations the Mental Health issue vertical has ever received from external consultations” and their own trust and safety team that reported these filters have “high potential to lead to negative body image and increase body dysmorphia symptoms, especially among teenage girls” and have a “disproportionate impact on minors” more generally.<sup>317</sup>

317. It is my opinion to a reasonable degree of medical and scientific certainty that the overwhelming evidence supports a causal relationship between social media use and body dysmorphia. In my opinion this occurs through a variety of mechanisms, including social

<sup>316</sup> TIKTOK3047MDL-054-LARK-00552309, -2309-10.

<sup>317</sup> TIKTOK3047MDL-054-LARK-00552309, -2311-12.

comparison that is enhanced and facilitated by the design of the SM platforms. A review of the internal Defendant documents provides further support for a causal relationship between social media and the development of body dysmorphia and body image conditions. Of note, it appears that on more than one occasion the Defendants utilized design features that increased the risk of harm, rather than safer alternatives.

#### **B. Eating Disorders**

318. In addition to the link between social media use and BDD discussed above, there are reasons to plausibly believe that social media usage can cause eating disorders. A brief summary of clinical eating disorders taken from Dane and Bhatia's derivation is in figure below.<sup>318</sup>

---

<sup>318</sup> Dane A, Bhatia K. The social media diet: A scoping review to investigate the association between social media, body image and eating disorders amongst young people. *PLOS Global Public Health*. 2023;3(3):e0001091.

**Figure 32: Summary of Clinical Eating Disorders**

CLINICAL EATING DISORDERS	
<b>Anorexia</b>	An intense fear of weight gain and/or a disturbed body image that motivates severe dietary restriction or other weight loss behaviours
<b>Bulimia</b>	Recurrent episodes of binge eating and compensatory behaviours, e.g., purging, to prevent weight gain
<b>Binge eating disorder</b>	Recurrent episodes of compulsive overeating that leads to distress without attempts to compensate for weight gain
<b>Avoidant/restrictive food intake disorder</b>	The avoidance or restrictive intake of food in the absence of body image concerns and fear of weight gain
<b>Pica</b>	Eating non-nutritive or non-food substances for a period of one month or more
<b>Rumination disorder</b>	Involves regurgitation of food after eating in the absence of nausea, involuntary retching, or disgust
SUBCLINICAL OTHER SPECIFIC FEEDING AND EATING DISORDERS	
<b>Orthorexia Nervosa</b>	A pathological fixation with healthy or ‘clean’ eating, avoidance of unhealthy foods and rigid dietary and exercise practices- violations of which cause severe emotional distress
<b>Atypical anorexia</b>	Majority of symptoms of anorexia are present, but the individual is classified as being within the normal BMI range
<b>Atypical bulimia</b>	Mimics clinical bulimia but occurs less frequently and with shorter duration
<b>Atypical binge eating disorder</b>	Mimics clinical binge eating disorder but occurs less frequently and with shorter duration
<b>Purging disorder</b>	Purging or using laxatives as a mean to control weight
<b>Night eating disorder</b>	Repeatedly eating at night, either after an evening meal or waking up from sleep
COMMON PATHOLOGY	
Dieting, bingeing, purging, restricting, avoidance of certain food groups, compulsive or compensatory exercise behaviours and the use of laxatives or weight loss pills	
Adapted from [1, 2, 5, 6]	
<a href="https://doi.org/10.1371/journal.pgph.0001091.t001">https://doi.org/10.1371/journal.pgph.0001091.t001</a>	

319. In terms of psychobiological plausibility, much of the basis for a linkage between social media usage and eating disorders is discussed in the BDD section (X.A) immediately prior and relates to social comparisons either with one’s “friends” or with other idealized images that are frequently posted to, or shared via, social media. It also includes risk created by tailored algorithms that can create rabbit holes or filter bubbles and provide extensive exposure of harmful content to a user.

320. A 2022 study by Fairplay, circulated and cited within Meta emails shared with me, investigated the so called “Pro-Eating Disorder” Bubble, namely an ecosystem of Instagram accounts that actively promote unhealthy eating habits and techniques to both optimize them and avoid detection. The researchers in this study created test accounts that showed an interest in pro-

eating disorder content by using vocabulary such as “Thinspo” and “TW (trigger warning)” in their biographies. Over 5 weeks of inactivity, one such account gained an additional 686 followers who had been algorithmically directed to the account and chose to follow it presumably in search of content.<sup>319</sup> Within that same report, a first-hand account of “Kelsey” a 17-year-old Southern California high school student is reported. Among other things, she alleges that “pro-eating disorder” content was “always pushed towards [her] from the moment I opened my account.” She goes on “I have never searched for those things and yet they pop up on my screen, whereas images or reminders of positive things, such as body positivity influencers, et cetera, I have to actively search for them in order for them to appear on my phone.”<sup>320</sup>

321. When asked if these kinds of “never searched for” things are an example of product mechanics that cause harm, Kang-Xing Jin responds, “Yes. If this user’s experience was that they were getting essentially pushed this content without having prior indication that they wanted it, then that would be, in my opinion, on the high-responsibility side of things for recommendations.”<sup>321</sup> Notably, he does not take issue that such unwelcomed content *could* be pushed towards a user. Indeed, in follow-up questioning, when asked if Meta should have done more to optimize for teen safety, Jin replies “Yes”<sup>322</sup> Indeed, Meta’s own data show that “18% of teens have seen someone promoting eating disorders or unhealthy weight loss on Instagram in the last week.”<sup>323</sup>

322. In terms of effect size, a recent metaanalysis of 87 effect sizes from 22 studies produced a summary estimate of the association between self-reported social media use and

---

<sup>319</sup> Kang-Xing Jin Dep. Exhibit 43

<sup>320</sup> Kang-Xing Jin Dep. Tr. at 493

<sup>321</sup> Kang-Xing Jin Dep. Tr. at 495

<sup>322</sup> Kang-Xing Jin Dep. Tr. at 710

<sup>323</sup> META3047MDL-003-00156702, -6718

various validated measures of eating disorders.<sup>324</sup> All studies included were published between 2010 and 2020 and the included sample had 5,031 males and 8,270 females. Notably, there was considerable heterogeneity in the studies driven in large part by BMI of respondent, sample source (e.g. college students, children, adolescents, clinical populations), and survey methods (e.g. online vs. paper and pencil). The summary estimate showed a “weak” but significant correlation between social media usage and disordered eating (0.09 [95% CI .06, .11]). The overall association, while small, must be taken in the context of the limitations of the data collection methodologies as well as the very high likelihood that there is significant variability in susceptibilities amongst included participants.

323. The primary predictor variable was some self-reported measure of social media usage and the primary outcome was some measure of disordered eating. For many people, social media usage may not include much (or any) portrayals of idealized bodies. In fact, for some people, social media usage might be affirming of their current body habitus and thus have a “protective” effect against disordered eating. Again, the limited granularity about what content is pushed and therefore consumed limits the ability to discern both a more precise estimate of effect but also who is more likely to be affected and makes it difficult for independent scientists to study this problem. This is (yet) another example of the kind of lacuna in research that could be filled with industry data.

324. Several longitudinal studies confirm that there is a temporal association between exposure to social media images of ideal bodies and subsequent eating disorder symptoms.<sup>325</sup> From

---

<sup>324</sup> Zhang J, Wang Y, Li Q, Wu C. The Relationship Between SNS Usage and Disordered Eating Behaviors: A Meta-Analysis. *Front Psychol.* 2021;12:641919. doi:10.3389/fpsyg.2021.641919

<sup>325</sup> See Brown Z, Tiggemann M. Attractive celebrity and peer images on Instagram: Effect on women's mood and body image. *Body Image.* 2016/12/01/ 2016;19:37-43.

a dose response (biologic gradient) standpoint, if increased exposure to content might drive (or be associated with) eating disorders, one might posit that more time spent on the internet might increase risk. To that end, a metaanalysis of studies that tested the association between problematic internet use and eating disorders including 39 studies from 21 countries found an effect size of (0.21 [95% CI 0.14, 0.28]).<sup>326</sup> This effect size is still small but twice the size of the overall effect size of general usage and eating disorders.

325. Experimental data are sparse because, much as with BDD, manipulations of exposure with sufficient frequency and long enough follow up to document differences in the development of eating disorders is sufficiently impractical as to be essentially impossible. But shorter-term laboratory based experimental approaches such as the one performed by Tiggman and summarized above, shows that there are at least short-term effects on one's body image based on what one is presented on social media.<sup>327</sup>

326. Internally documents reflect a recognition that unconnected content either in "image" or "reel" format on users' "explore" tab could be problematic for people at risk for eating

---

doi:<https://doi.org/10.1016/j.bodyim.2016.08.007>; Kim M, Park W. Who is at risk on Facebook? The effects of Facebook News Feed photographs on female college students' appearance satisfaction. *The Social Science Journal*. 2016/12/01/ 2016;53(4):427-434, doi:<https://doi.org/10.1016/j.soscij.2016.08.007>; Puccio F, Kalathas F, Fuller-Tyszkiewicz M, Krug I. A revised examination of the dual pathway model for bulimic symptoms: The importance of social comparisons made on Facebook and sociotropy. *Computers in Human Behavior*. 2016/12/01/ 2016;65:142-150. doi:<https://doi.org/10.1016/j.chb.2016.08.018>.

<sup>326</sup> Ioannidis K, Taylor C, Holt L, et al. Problematic usage of the internet and eating disorder and related psychopathology: A multifaceted, systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*. 2021/06/01/ 2021;125:569-581. doi:<https://doi.org/10.1016/j.neubiorev.2021.03.005>

<sup>327</sup> See Kim M, Park W. Who is at risk on Facebook? The effects of Facebook News Feed photographs on female college students' appearance satisfaction. *The Social Science Journal*. 2016/12/01/ 2016;53(4):427-434. doi:<https://doi.org/10.1016/j.soscij.2016.08.007>; Tiggemann M, Barbato I. "You look great!": The effect of viewing appearance-related Instagram comments on women's body image. *Body Image*. 2018/12/01/ 2018;27:61-66. doi:<https://doi.org/10.1016/j.bodyim.2018.08.009>

disorders. Below is an exchange between Jayakumar and ██████████ from Meta's policy department:

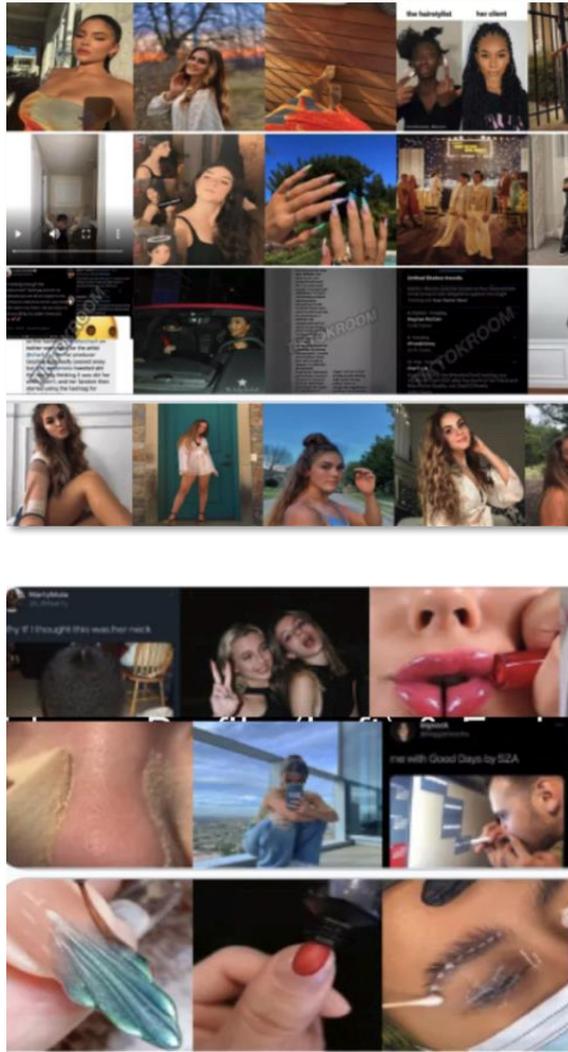
```
██████████ (4/09/2021 10:03:45 PDT):
>would we consider not surfacing certain content in our rec surfaces that we know to be triggering?
P. ██████████ (4/09/2021 10:04:05 PDT):
> we've identified borderline harmful in other categories
Vaishnavi Jayakumar (4/09/2021 10:04:15 PDT):
>Totally - we're working on a borderline ED policy right now to address this and are just kicking that
off
██████████ (4/09/2021 10:04:22 PDT):
Vaishnavi Jayakumar (4/09/2021 10:05:19 PDT):
>we had a first round of brainstorming this week that was really interesting - bucketing potential types
of behaviours into borderline / not-borderline
██████████ (4/09/2021 10:06:03 PDT):
>The images that Satish sent yesterday of an explore feed is really problematic, and the ownership of
that is also on us since it's an unconnected, recommended surface.
Vaishnavi Jayakumar (4/09/2021 10:07:21 PDT):
>Yeah - the challenge is that when you zoom in, any one of those photos probably wouldn't be triggering.
But in an explore grid with all these photos stacked against one another, it's pretty overwhelming
```

*Document 109: META3047MDL-040-00541685*

327. Below are the referenced images taken from the explore feed of a teenage Instagram user sent by ██████████ who stated: "The question is that there is content that will continue to be created on the platform that makes teens feel back [sic] about themselves."<sup>328</sup>

---

<sup>328</sup> Vaishnavi Jayakumar Dep. Tr. at 181:3-6



*Document 110: META3047MDL-145-0000005*

328. Jayakumar in her deposition asserts that while viewing these images would not necessarily cause harm to everyone, “One could argue that if you already struggle with negative social comparison, being flooded with a wave of images like that could be -could make you feel really bad about yourself.”<sup>329</sup> Fair enough. But the design of social media is such that certain

---

<sup>329</sup> Vaishnavi Jayakumar Dep. Tr. at 184:2-6

people with pre-existing vulnerabilities are precisely the ones that will be more likely to receive a “flood” of posts that tap into this pre-existing vulnerability. This amplification of content and design that pushes the content to vulnerable children is part of the mechanism by which social media simultaneously increases engagement and the risk of harm. In other words, the problem is that teens with body image issues or those prone to eating disorders are fed such posts and it happens in the “explore” tab which is amassed by Meta’s algorithms. In fact, Jayakumar goes on to say that the “aggregation” is part of the problem.

329. YouTube’s algorithms also curate ED content for its viewers. Google’s own internal focus group study had the following quote:

**EXPERT POV**



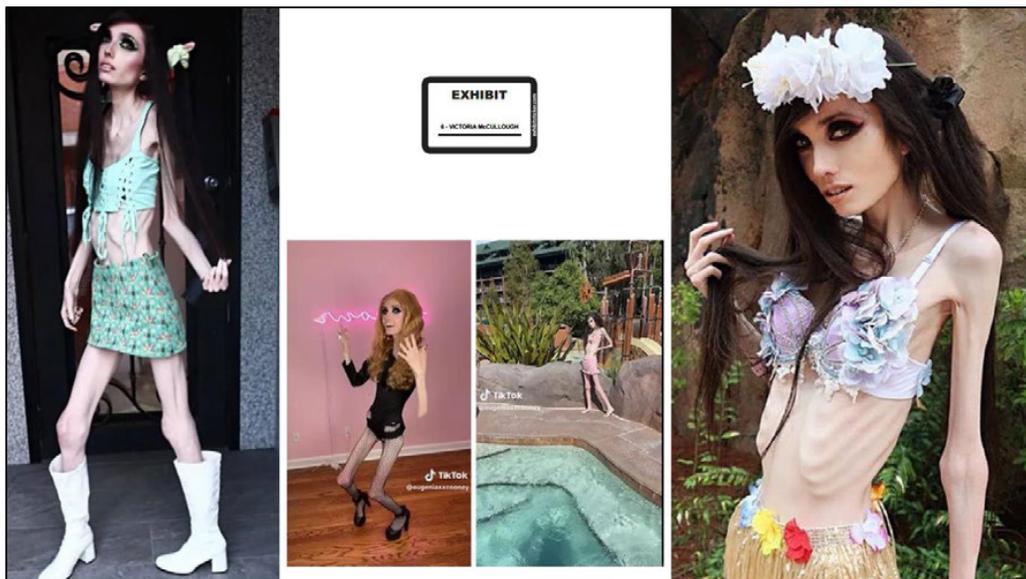
*“I worked with a 13-year-old who was watching [YT videos] to maintain her anorexia... once you find one video, it’s so easy to see more and more.*

*“Most products come with guidelines for healthy use: ‘Here’s how to use this safely without harming yourself or others.’ Tech products should be no different.*

*- Nicole, Family Therapist*

*Document 111: GOOG-3047MDL-00236723 at Slide 12 (emphasis in original)*

330. Even when alerted about accounts that posed risks to teens, TikTok continued to promote them. For example, Eugenia Cooney’s account was repeatedly reported as being problematic for its glamorizing of a profoundly emaciated and unhealthy figure:



*Document 112: Victoria McCullough Dep. Exhibit 7*

331. In fact, her account was “a part of TikTok’s Creator Fund.”<sup>330</sup> Her skeletal appearance—as unhealthy as it is—is hard for me to look at it even as a physician. It is shameful that it was promoted.

332. I have seen no evidence that TikTok informed the medical community, parents, or underage users of the increased risks of eating disorders that arise from use of their product. This decision is particularly concerning given evidence that outside experts requested that they do so. Dr. Bryn Austin, a pediatrician at Harvard Medical School, “suggested that TikTok should incorporate ‘warning labels (access to help lines, etc.)’ in ED-like content.”<sup>331</sup> Further, Dr. Austin suggested that “promoting/havi[n]g accessibility to ED content on teen’s FYP” is “more likely to

---

<sup>330</sup> Victoria McCullough Dep. Exhibit 7

<sup>331</sup> Victoria McCullough Dep. Exhibit 7.

cause harm.”<sup>332</sup> At deposition, McCullough stated she was not aware of TikTok ever informing the public of this risks of eating disorders despite this advice.<sup>333</sup>

333. It is my opinion that to a reasonable degree of medical and scientific certainty, social media usage increases the risk of eating disorders and body image conditions in adolescents. This is in part due to several design features that exacerbate social comparison, provide idealized beauty standards, and pull children into filter bubbles or rabbit holes. A review of the internal documents provides further support that Defendants recognized that social media increased the risk of or exacerbated eating disorders, negative social comparison, body image conditions, and low self-esteem.

### **C. Sleep Problems**

334. Sleep is essential for maintaining overall health and well-being. It plays a crucial role in various bodily functions, including the repair and rejuvenation of cells, muscles, and tissues. During sleep, the brain consolidates memories and processes information which can help improve cognitive function, memory retention, and problem-solving skills. Lack of sleep, on the other hand, has been linked to a higher risk of developing chronic health conditions such as heart disease, diabetes, and obesity. It also weakens the immune system, making the body more vulnerable to infections. Moreover, sleep is vital for mental and emotional health. A well-rested brain is better equipped to manage stress, regulate emotions, and maintain a positive mood.

335. Sleep deprivation, however, can lead to irritability and difficulty concentrating, which negatively impacts daily performance, including at school. In the long term, insufficient sleep can contribute to mental health disorders like depression and anxiety. Therefore, maintaining

---

<sup>332</sup> Victoria McCullough Dep. Exhibit 7.

<sup>333</sup> Victoria McCullough Dep. Tr. at 81:10-83:11

a healthy sleep routine is crucial for both physical and mental well-being. As presented above (Section VII.C), a metaanalysis found that the mere presence of a device in a child’s bedroom is associated with a 79% increased risk of sleep problems irrespective of addiction.<sup>334</sup>

336. For simplicity, we will focus only on the direct path between SM use and sleep (which again is enhanced by compulsive/addictive screen use). But both depression and anxiety (discussed in the next section of this report) have reciprocal relationships with sleep, and in fact sleep disturbance is a core symptom of both.<sup>335</sup>

337. There are several plausible psychobiological/environmental mechanisms by which this relationship is causal:

338. Psychological arousal: Social media, by virtue of the engaging (and enraging) content its algorithms are designed to surface, can inhibit the relaxation process that is essential to inducing sleep.

339. Displacement: Peri-bedtime social media use effectively displaces time that could be spent sleeping. Simply put, a child in bed on their phone is postponing or delaying sleep.

340. Melatonin inhibition: Decades of research have established that emitted light suppresses secretion of the sleep inducing hormone, melatonin.<sup>336</sup>

341. Alerts or Notifications: Audible or visual alerts from social media sites indicating updates can disrupt or delay sleep.

---

<sup>334</sup> Carter B, Rees P, Hale L, Bhattacharjee D, Paradkar MS. Association Between Portable Screen-Based Media Device Access or Use and Sleep Outcomes: A Systematic Review and Meta-analysis. *JAMA Pediatr.* Dec 1 2016;170(12):1202-1208. doi:10.1001/jamapediatrics.2016.2341.

<sup>335</sup> Gregory AM, Sadeh A. Sleep, emotional and behavioral difficulties in children and adolescents. *Sleep Med Rev.* Apr 2012;16(2):129-36. doi:10.1016/j.smrv.2011.03.007

<sup>336</sup> Lewy AJ, Wehr TA, Goodwin FK, Newsome DA, Markey SP. Light suppresses melatonin secretion in humans. *Science.* Dec 12 1980;210(4475):1267-9. doi:10.1126/science.7434030

342. Fear of Missing Out (FoMo): As discussed elsewhere in this report, FoMo mitigation drives social media use.

343. In terms of the association between social media and sleep problems, there have been several systematic reviews performed.<sup>337</sup> A comprehensive systematic review summarized 42 studies of social media use and sleep quality including both cross sectional and longitudinal studies.<sup>338</sup> It reported that five cohort studies found excessive social media use at baseline to be a risk factor for poor sleep quality a follow up. In addition, among the 24 cross sectional studies identified, 23 found positive associations between frequent social media use and poor sleep quality.

344. In terms of temporality and effect size, a few longitudinal studies have examined the relationship between social media and insomnia. In one study, 1098 adolescents (13-19 years of age) were followed serially over 4 months with monthly assessments of nomophobia, social media addiction, and insomnia.<sup>339</sup> Insomnia was associated with both nomophobia ( $B=.20$   $p<.001$ ) and addictive social media usage ( $B=.49$   $p<.001$ ). A systematic review and metaanalysis of 23 *longitudinal* studies of screen use and sleep in adolescents found an effect size of  $-.12$  for social media usage  $-.19$  for dysfunctional media usage at baseline and subsequent sleep health at follow

---

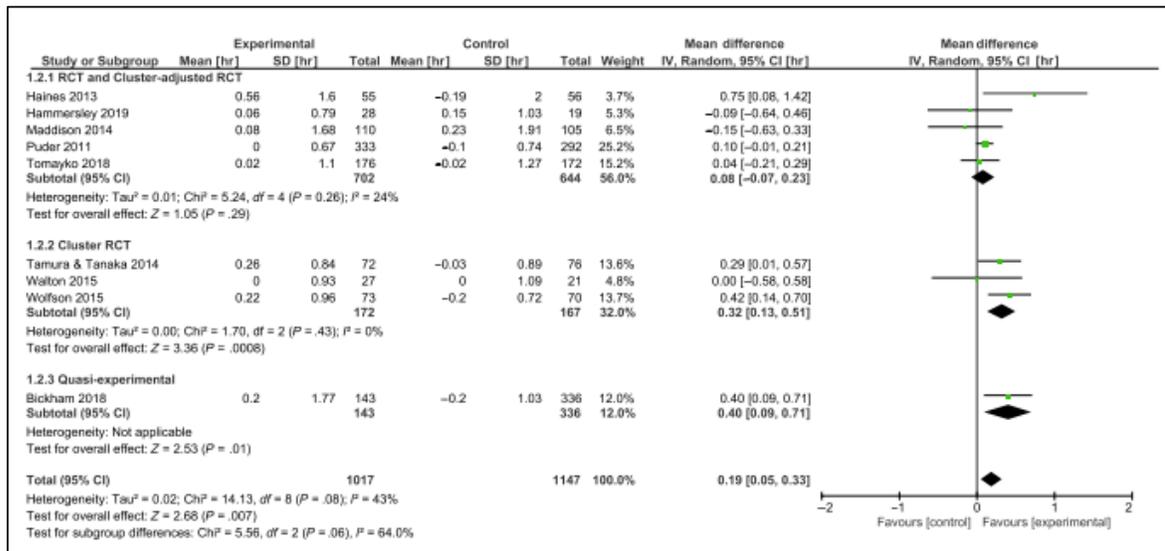
<sup>337</sup> A distinction can be made between studies that looked at SM use and sleep and those that looked at screen time and sleep. While some studies regarding screen time and sleep have inconsistent results regarding association between use and sleep health outcomes, the use of SM that interrupt nighttime sleep are associated with a variety of adverse sleep outcomes in the literature. This distinction is discussed in the Handbook of Children and Screens.

<sup>338</sup> Alonzo R, Hussain J, Stranges S, Anderson KK. Interplay between social media use, sleep quality, and mental health in youth: A systematic review. *Sleep Med Rev.* Apr 2021;56:101414. doi:10.1016/j.smrv.2020.101414

<sup>339</sup> Lin C-Y, Potenza MN, Ulander M, et al. Longitudinal Relationships between Nomophobia, Addictive Use of Social Media, and Insomnia in Adolescents. *Healthcare.* 2021;9(9):1201.

up.<sup>340</sup> A systematic review and metaanalysis of 16 *experimental* studies of screen time reduction in children is summarized below.<sup>341</sup>

**Figure 33: 28 Metanalysis Of Experimental Studies Of Screen Use And Sleep**



345. The overall effect of screen time reduction interventions resulted in a significant increase of 11 minutes of sleep. Since that metaanalysis was published, Bartel and colleagues performed a within person experiment in 14–18-year-old adolescents wherein their baseline sleep was monitored for a week after which they were given individual phone stoppage time 1 hour before bed for one school week. During the phone restriction week, adolescents stopped using their phones 80 min earlier, turned their lights out 17 minutes earlier and slept 21 minutes longer.<sup>342</sup>

<sup>340</sup> Pagano M, Bacaro V, Crocetti E. “Using digital media or sleeping ... that is the question”. A meta-analysis on digital media use and unhealthy sleep in adolescence. *Computers in Human Behavior*. 2023/09/01/ 2023;146:107813. doi:<https://doi.org/10.1016/j.chb.2023.107813>

<sup>341</sup> Martin KB, Bednarz JM, Aromataris EC. Interventions to control children's screen use and their effect on sleep: A systematic review and meta-analysis. *Journal of Sleep Research*. 2021;30(3):e13130. doi:<https://doi.org/10.1111/jsr.13130>

<sup>342</sup> Bartel K, Scheeren R, Gradisar M. Altering Adolescents’ Pre-Bedtime Phone Use to Achieve Better Sleep Health. *Health Communication*. 2019/03/21 2019;34(4):456-462. doi:10.1080/10410236.2017.1422099

346. Finally, the Davis and Goldfield study (detailed in the FoMo section of this report, Section VII.F), also assessed sleep after 3 weeks of social media reduction. The intervention group slept almost 30 minutes more per night based on self-report. To understand the implications of this effect size, consider that among teens, evidence from the school start time literature show that average differences as small as 10-20 minutes of sleep per night are adversely associated with academic, cognitive, and motor vehicle crash outcomes.<sup>343</sup>

347. FoMo occurs throughout the day of course but there are, at least in theory, other structural societal constraints that lean against checking social media continuously (e.g. school for children although school cell policies have until very recently been nonexistent or not enforced effectively).<sup>344</sup> However, at the end of the day, when children are in bed, those other constraints are largely gone and FoMo can preclude the mental relaxation that is requisite to induce sleep.

348. In addition, there is some evidence of coherence. Functional magnetic imaging studies have shown that various areas of the brain are differentially activated by episodic or chronic usage of social media.<sup>345</sup> In particular, both the nucleus accumbens discussed earlier, and the amygdala which is part of the “arousal” circuitry of the brain, could plausibly impact sleep and both are preferentially stimulated. Additional features of social media that interrupt sleep include

---

<sup>343</sup> Wheaton AG, Chapman DP, Croft JB. School Start Times, Sleep, Behavioral, Health, and Academic Outcomes: A Review of the Literature. *J Sch Health*. May 2016;86(5):363-81. doi:10.1111/josh.12388; Bowers JM, Moyer A. Effects of school start time on students' sleep duration, daytime sleepiness, and attendance: a meta-analysis. *Sleep Health*. Dec 2017;3(6):423-431. doi:10.1016/j.sleh.2017.08.004; Danner F, Phillips B. Adolescent sleep, school start times, and teen motor vehicle crashes. *J Clin Sleep Med*. Dec 15 2008;4(6):533-5.

<sup>344</sup> Tandon PS, Zhou C, Hogan CM, Christakis DA. Cell Phone Use Policies in US Middle and High Schools. *JAMA Netw Open*. May 1 2020;3(5):e205183. doi:10.1001/jamanetworkopen.2020.5183

Christakis DA, Mathew GM, Reichebberger DA, Rodriguez IR, Ren B, Hale L. Adolescent smartphone use during school hours. *JAMA Peds*. In press

<sup>345</sup> Wadsley M, Ihssen N. A Systematic Review of Structural and Functional MRI Studies Investigating Social Networking Site Use. *Brain Sciences*. 2023;13(5):787.

receiving nighttime notifications, addictive engagement through the algorithmic delivery of intermittent variable rewards, and increased anxiety and depressive symptoms that are also a byproduct of the design of social media.

349. Furthermore, while the evidence reviewed herein will be deliberately limited to social media use and sleep, there is a broader literature linking *overall* screen use to disordered sleep and as discussed in section elsewhere in this report, social media use constitutes a significant percentage of that use. In fact, 60% of emerging adults report using screens and social media prior to bedtime.<sup>346</sup> Reviewing the totality of existing data, a recent expert consensus panel (which I was a member of) convened by the National Sleep Foundation reached consensus meaning that at least 80% of the 16 experts agreed that: (1) in general, screen use impairs sleep health among children and adolescents, (2) the content of screen use before sleep impairs sleep health of children and adolescents, and (3) behavioral strategies and interventions may attenuate the negative effects of screen use on sleep health.<sup>347</sup> I would assert that it's the presentation of content that impairs sleep, as well as addictive features such as infinite scroll that directly contribute to impairing sleep.

350. Let us now apply the distinction between confounding and mediating (discussed above) where the outcome is sleep. Consider the summation of the findings of a study by Viner et al<sup>348</sup> that is referenced in a TikTok memo:

---

<sup>346</sup> Hale L, Kirschen GW, LeBourgeois MK, et al. Youth Screen Media Habits and Sleep: Sleep-Friendly Screen Behavior Recommendations for Clinicians, Educators, and Parents. *Child Adolesc Psychiatr Clin N Am*. Apr 2018;27(2):229-245. doi:10.1016/j.chc.2017.11.014

<sup>347</sup> Hartstein LE, Mathew GM, Reichenberger DA, et al. The impact of screen use on sleep health across the lifespan: A National Sleep Foundation consensus statement. *Sleep Health*. Aug 2024;10(4):373-384. doi:10.1016/j.sleh.2024.05.001

<sup>348</sup> Viner RM, Gireesh A, Stiglic N, et al. Roles of cyberbullying, sleep, and physical activity in mediating the effects of social media use on mental health and wellbeing among young people in England: a secondary analysis of longitudinal data. *Lancet Child Adolesc Health*. Oct 2019;3(10):685-696. doi:10.1016/s2352-4642(19)30186-5

We found that strong, longitudinal associations between very frequent social media use and mental health and well-being in girls were largely *mediated* by cyberbullying and the displacement of sleep and physical activity..... Our data suggests that interventions to reduce social media use to improve mental health might be misplaced. Preventive efforts should consider interventions to prevent or increase resilience to cyberbullying and ensure adequate sleep and physical activity in young people.”<sup>349</sup>

The memo heralds the “mediation” finding as proving that social media does *not* play a role in adverse mental health effects claiming instead that sleep, physical activity, and cyberbullying are to blame. But the paper reports that the association between social media use and mental health was *mediated* by cyberbullying, physical activity, and sleep, not *confounded* by them.<sup>350</sup> Social media sites can lead to cyberbullying, sleep disruptions, and reduced physical activity which in turn can negatively impact mental health. It is *through* them (along with other mechanisms) that social network sites affect mental health.

351. As for “preventive efforts” that might increase sleep and physical activity and promote resilience to cyberbullying, reduced screen time during the day (for physical activity) and at night (for sleep) would surely be part of any effective intervention. In fact, the same TikTok memo referenced above goes on to say, “When assessing and addressing TikTok’s potential wellbeing impacts **we should consider not just how long children are using TikTok for but also when in the day they are doing so.**”<sup>351</sup> And later still, the memo acknowledges: “Currently we send notifications to users during the school day in some cases, up until midnight, which could interfere with sleep.”<sup>352</sup> Finally, social media mediation can actively reduce cyberbullying. For

---

<sup>349</sup> TIKTOK3047MDL-002-00100441, -0445 (emphasis added)

<sup>350</sup> Viner RM, Gireesh A, Stiglic N, et al. Roles of cyberbullying, sleep, and physical activity in mediating the effects of social media use on mental health and wellbeing among young people in England: a secondary analysis of longitudinal data. *Lancet Child Adolesc Health*. Oct 2019;3(10):685-696. doi:10.1016/s2352-4642(19)30186-5

<sup>351</sup> TIKTOK3047MDL-002-00100441, -0448 (emphasis in original)

<sup>352</sup> TIKTOK3047MDL-002-00100441, -0452

example, Instagram began using AI in 2017 to detect and suppress cyberbullying as TikTok itself calls out.<sup>353</sup>

352. Meta was aware of the concern that Instagram usage was displacing sleep. Shruti Bhutada, wellbeing lead at Meta, made and circulated a graphic of the underlying ways she saw social media disrupting sleep:

how it's disrupting sleep	contributing factors	activities
<ul style="list-style-type: none"><li>• <b>Delaying</b> or reducing sleep hours because of using FB</li></ul>	<ul style="list-style-type: none"><li>• Content is addictive!</li></ul>	<ul style="list-style-type: none"><li>• Endlessly scrolling</li></ul>
<ul style="list-style-type: none"><li>• <b>Waking up</b> in the middle of their sleep and checking FB, which prolongs their return to sleep</li></ul>	<ul style="list-style-type: none"><li>• Time zone differences</li><li>• Notifications pull you in</li></ul>	<ul style="list-style-type: none"><li>• Watching videos</li></ul>
<ul style="list-style-type: none"><li>• <b>Upsetting</b>, content or interactions can make it harder to sleep</li></ul>	<ul style="list-style-type: none"><li>• Time to catch up on what missed</li><li>• Self-control is lower at night</li></ul>	<ul style="list-style-type: none"><li>• Checking up on notifications</li></ul>

*Document 113: META3047MDL-019-00106590, -6592*

353. Another former Meta employee, Kang-Xing Jin, testified in his deposition: “I think that’s a reasonable high-level summary of what I would term the displacement of beneficial behaviors category of harms. And I think among those, for example, sleep is a very common one where, even if using Instagram or any other app on your phone late at night isn’t inherently harmful, if it’s actually causing you to get less sleep, my understanding is there’s a fair amount of

<sup>353</sup> TIKTOK3047MDL-002-00100441, -0457

research that suggests enough sleep is important; right?”<sup>354</sup> Jin goes on to discuss the “platform’s” vs the “individual’s” agency with regard to features of Meta. He states:

So using late-night notifications as an example, the platform exerts a fair amount of control through its defaults around, you know, whether, when, and to whom to send those notifications. And so that would fit into mechanics. And you could see a place where, you know, decisions there may amplify some of these issues. And that’s maybe somewhat distinct from tools and resources, where you could imagine the platform could also give people explicit controls to manage those notifications themselves and/or just upsell the system-level controls that already exist right?<sup>355</sup>

354. Instagram went so far as to document “late night use” which they defined as being on the site between 12 and 4 AM local time. They found that 43% of teen users had at least one late night session per week and 4.6% have one or more sessions every night.<sup>356</sup> As recently as 2022, over 2/3 of these late-night sessions are initiated by notification or a badge.<sup>357</sup> Mr. Volichenko’s deposition quotes Ms. Hanko who states “For late night use, the negative impacts such as on work/school performance and mood, of little or poor sleep, especially on young people are well-documented.”<sup>358</sup>

355. According to Volichenko, Meta considered instituting a “quiet mode” at night so that teens would not get notifications that might impede their sleep.<sup>359</sup> In fact, according to documents I have seen, that idea was first proposed before 2018. Meta knew that users--especially teens--were looking for solutions to curb late night use of their platforms. An early 2018 presentation titled “Time Spent: The Research Journey” reports that “teens said their phones were a constant distraction [during both the night and the day], both because of notifications and the

---

<sup>354</sup> Kang-Xing Jin Dep. Tr. at 397:10-20

<sup>355</sup> Kang-Xing Jin Dep. Tr. at 438:1-14

<sup>356</sup> META3047MDL-035-00005017 at Slide 12

<sup>357</sup> META3047MDL-035-00005017 at Slide 18

<sup>358</sup> George Volichenko Dep. Exhibit 5 at -2192

<sup>359</sup> George Volichenko Dep. Tr. at 141:14-143:2

convenience of immediate use.”<sup>360</sup> Those teens also unanimously agreed that they would “use a feature that limited notifications to certain times a day.”<sup>361</sup> In response to this feedback, Meta’s researchers recommended “resurrect[ing] a product” internally referred to as “Quiet Mode” that was “shelved because product focused on the negative.”<sup>362</sup> Instagram did not launch Quiet Mode until 2023.

356. Launching “quiet mode” took a fair amount of convincing of senior leadership.

Jayakumar states in her deposition:

I think it was a little challenging. I think while we generally had consensus that quiet mode, you know, sounded like a good idea in theory, we needed to really do a lot of research into what the impact might be on the platform, on engagement, on, you know, growth, I suppose is the shorter way of putting it.<sup>363</sup>

357. Instagram finally launched quiet mode in 2023 (at least five years after it was first proposed). But it was not launched as a “default” setting and as Jayakumar attested 6 out of 10 teens report never consider changing their settings from the “default ones.”<sup>364</sup> Furthermore, as part of “quiet mode” the team considered and rejected making the screen be black and white – rather than color—to be less stimulating after hours.

358. Not surprisingly then, once launched and tested, “quiet mode” was found to result in a 0.3 percent (and not statistically significant) reduction in late night usage over 7 days.<sup>365</sup> When they looked at a 28-day window they found a reduction of 0.38% that was statistically

---

<sup>360</sup> Kenzie Snyder Dep. Exhibit 28 at Slide 5.

<sup>361</sup> Kenzie Snyder Dep. Exhibit 28 at Slide 5.

<sup>362</sup> Kenzie Snyder Dep. Exhibit 28 at Slide 9; *see also* Kenzie Snyder Dep. Ex. 24 at -6820 (Proposing an “Off Mode setting one can enable to lock oneself out of FB at set times (like at school, work, or at bedtime” as early as 2017.).

<sup>363</sup> Vaishnavi Jayakumar Dep. Tr. at 150:13-19

<sup>364</sup> Vaishnavi Jayakumar Dep. Tr. at 74:13-19

<sup>365</sup> George Volichenko Dep. Exhibit 7 at -4278

significant.<sup>366</sup> But both of these effect sizes are *de minimis* or, in Volichenko's words "very small."<sup>367</sup> Quiet Mode was launched as a proven ineffective mechanism to reduce late night engagement but it "checked the box" of giving the appearance that Instagram was taking steps to curb problematic use to mitigate litigation risk.<sup>368</sup> Volichenko himself states that his supervisor, Ms. █████ did "not put teen safety first."<sup>369</sup>

359. TikTok also recognizes (admits) that its app negatively impacts sleep which motivated them to develop some functionality to diminish use late at night.

**Why build it?**

**Problem statement:** The use of TikTok at night delays sleep for some users, disturbing their sleep patterns and preventing them from getting the minimum recommended amount of sleep.

**Problem validation:**

*Moved due to sensitivity:* [ HYPERLINK "<https://bytedance.feishu.cn/docx/doxcnhT1vqjACSORbkiSOhR0bTf>" \h ]

**Business and Reputational Impact:**

- Users are more likely to churn or reduce activity on TikTok if the app negatively impacts their ability to manage their time and tend to their basic needs
  - 23% of inactive users cite too much time spent on TikTok [ HYPERLINK "<https://bytedance.us.feishu.cn/docs/doccnZJO57KClCbRhNGb9uldE0>" \h ]
- TikTok has received negative media coverage as "the worst social media app for sleep" ([ HYPERLINK "<https://au.lifestyle.yahoo.com/social-media-app-tikok-keeping-you-awake-bad-sleep-041323202.html>" \h ], [ HYPERLINK "<https://www.sleepjunkie.com/revange-sleep-procrastination/>" \h ])

Document 114: TIKTOK3047MDL-010-00329290, -9294

<sup>366</sup> George Volichenko Dep. Exhibit 7 at -4271

<sup>367</sup> George Volichenko Dep. Tr. at 151:1-2

<sup>368</sup> George Volichenko Dep. Tr. at 74:23-75:8

<sup>369</sup> George Volichenko Dep. Tr. at 154:8-13

In particular, TikTok was cognizant of the UK’s Age-Appropriate Design Code (AADC) and its California counterpart that both list “interrupted or inadequate sleep patterns” as a potential risk to be assessed and mitigated.<sup>370</sup>

360. Consistent with their corporate culture, designing a sleep reminder system was first subjected to A/B testing to ensure that it did not hit the guardrail of hurting daily active use (DAU).

Success - Secondary	Adoption of Sleep Reminders	Sleep Reminder DAU (% of total DAU with a sleep reminder configured)	0.10% of DAU, measured 2 weeks after launch
	Increase adoption of Screen Time Management features	Screen Time Management DAU (% of Total DAU)	Small increase
	Positive brand safety impact	Positive media/regulator feedback	-
Guardrail	Observe trade-offs and impact	Total stay duration / user	Fluctuate / Very small decrease
		Publish / user	No impact
		App uninstall rate	Small decrease

*Document 115: TIKTOK3047MDL-010-00329290, -9297*

361. YouTube data from 2016, early in the platform’s history, found that 27% of 18–24-year-olds report that it is cutting into their sleep time.<sup>371</sup> Relatedly, 45.9% of young adults report staying on YouTube longer than they should which, per Google’s own report, led to the “insight” that “some of the heaviest users on our platform don’t report any well-being effects and aren’t

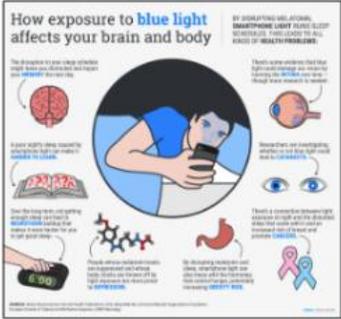
<sup>370</sup> TIKTOK3047MDL-010-00329290, -9295

<sup>371</sup> GOOG-3047MDL-00236723 at Slide 9

aware of how usage affects them.<sup>372</sup> Further, a 2018 YouTube internal presentation on what is “known about video effects,” includes the following:

## Blue-light from screens causes sleep deprivation; ultimately affects the brain's mental processing

- Technology devices **emit light** at multiple (unnatural) wavelengths **that alter our brain chemistry.**
- **Blue-light reduces cortisol and melatonin**, the hormones responsible for our sleep/wake cycles.
  - This keeps the brain alert
  - Tricks the brain into thinking it needs to be awake
- **Lack of sleep can result in poor executive functioning**
  - Humans have several stages of sleep
  - REM sleep (the deepest one) is most important for synaptic rejuvenation and memory consolidation...basically, the brain's housekeeping mechanisms.
  - Inability to experience REM sleep can cause memory loss, neural circuit damage, and slower mental processing.
  - Results in lower academic performance in students/teens



(National Sleep Foundation, *n.d.*; Rosen, 2016)

Document 116: GOOG-3047MDL-00874191 at Slide 23

<sup>372</sup> GOOG-3047MDL-00236723 at Slide 10

## 18-24 year olds report YouTube cutting into their sleep time twice as much as other adults on the platform

### IN THE LAST WEEK WHICH OF THE FOLLOWING APPLIED TO YOU?

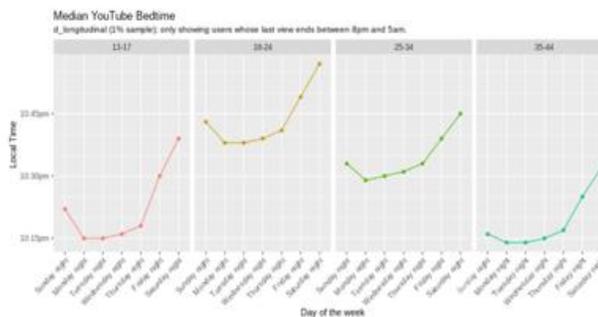
#### 18-24 YEAR OLDS



#### YOUTUBE AUDIENCE 25+



### TIME OF DAY WHEN USERS STOPPED WATCHING YOUTUBE, BY AGE



Document 117: GOOG-3047MDL-01371645 at Slide 73

27.3% of YouTubers believe it cuts into their sleep and the median “bedtime” for 13–17-year-olds on weeknights is after 10:15 PM local time. Remember that median is the midpoint meaning over ½ are on it after that time. 12% of total watchtime occurs between 12 AM to 6 AM.<sup>373</sup> The full range of their data, while known to them, is not in this document. “Autoplay watch time contribution triples during the night.”<sup>374</sup> Finally, in a presentation allegedly adjudicating allegations made about YouTube, the slide to the right which speaks for itself appears.

<sup>373</sup> Cristos Goodrow Dep. Exhibit 7 at p. 15

<sup>374</sup> Cristos Goodrow Dep. Exhibit at 7 at p. 16

### Autoplay disrupts sleep patterns

Status: TRUE

Autoplay potentially disrupts sleep patterns. Autoplay contribution to daily watchtime increases significantly.



*Document 118: Cristos Goodrow Dep. Exhibit 7 at p. 34*

362. In my opinion, to a reasonable degree of medical and scientific certainty, a causal relationship exists between social media use and sleep problems. Furthermore, there is ample evidence internally that use of Meta, Snap, TikTok and Google social media platforms disrupt sleep. The actions Defendants took to mitigate these problems, if any, were weighed against the impact they would have on their core metrics, and ultimately on their bottom line, and were minimally effective by design.

#### **D. Depression and Anxiety**

363. Depression and anxiety are closely intertwined, with many individuals experiencing symptoms of both conditions simultaneously. While they are distinct mental health

disorders, they often co-occur, creating a cycle of emotional and physical challenges that can amplify their individual impacts. Depression is primarily characterized by persistent feelings of sadness, hopelessness, and a loss of interest in previously enjoyable activities. On the other hand, anxiety involves excessive worry, fear, and a heightened state of arousal or tension. Despite these differences, they share common symptoms such as fatigue, difficulty concentrating, and sleep disturbances, making it difficult to differentiate between the two at times.

364. The relationship between depression and anxiety can be explained through shared biological, psychological, and environmental factors. Biologically, both conditions are associated with dysregulation in the brain's neurotransmitters, particularly serotonin, dopamine, and norepinephrine. This overlap explains why individuals predisposed to one condition are at increased risk of developing the other. Psychologically, negative thought patterns such as catastrophizing or ruminating can fuel both anxiety and depression, creating a feedback loop where worry about the future exacerbates feelings of hopelessness and vice versa. Environmental factors, such as chronic stress, trauma, or significant life changes, can also act as triggers for both conditions.

365. This interrelationship is reflected in the scientific literature where studies often evaluate both depression and anxiety simultaneously in the context of the same research protocol often labeling them as “internalizing” symptoms because they are often not obvious to observers. Depression and anxiety are mutually enhancing. Therefore, studies showing an increase in either can be interpreted as demonstrating an increase in both. Similarly, sleep problems, body image and self-esteem issues (addressed in other parts of this Section X) also increase the risk of both anxiety and depression. Shruti Bhutada (IG Well-Being User Experience Researcher) states,

“Social Media is often not the cause of problems related to mental health. However, it can and does both attenuate and exacerbate a user’s experience with mental health issues.”<sup>375</sup>

**i) Depression**

366. Depression is a clinically diagnosable mental health disorder characterized by persistent and pervasive feelings of sadness, hopelessness, and a lack of interest or pleasure in activities, often accompanied by physical symptoms such as changes in appetite, sleep disturbances, and fatigue. A clinical diagnosis of “depression” entails having a constellation of “depressive” symptoms of sufficient intensity and frequency absent an alternative explanation for them. The DSMV-R criteria for depression are below:

---

<sup>375</sup> META3047MDL-019-00106371, -6371

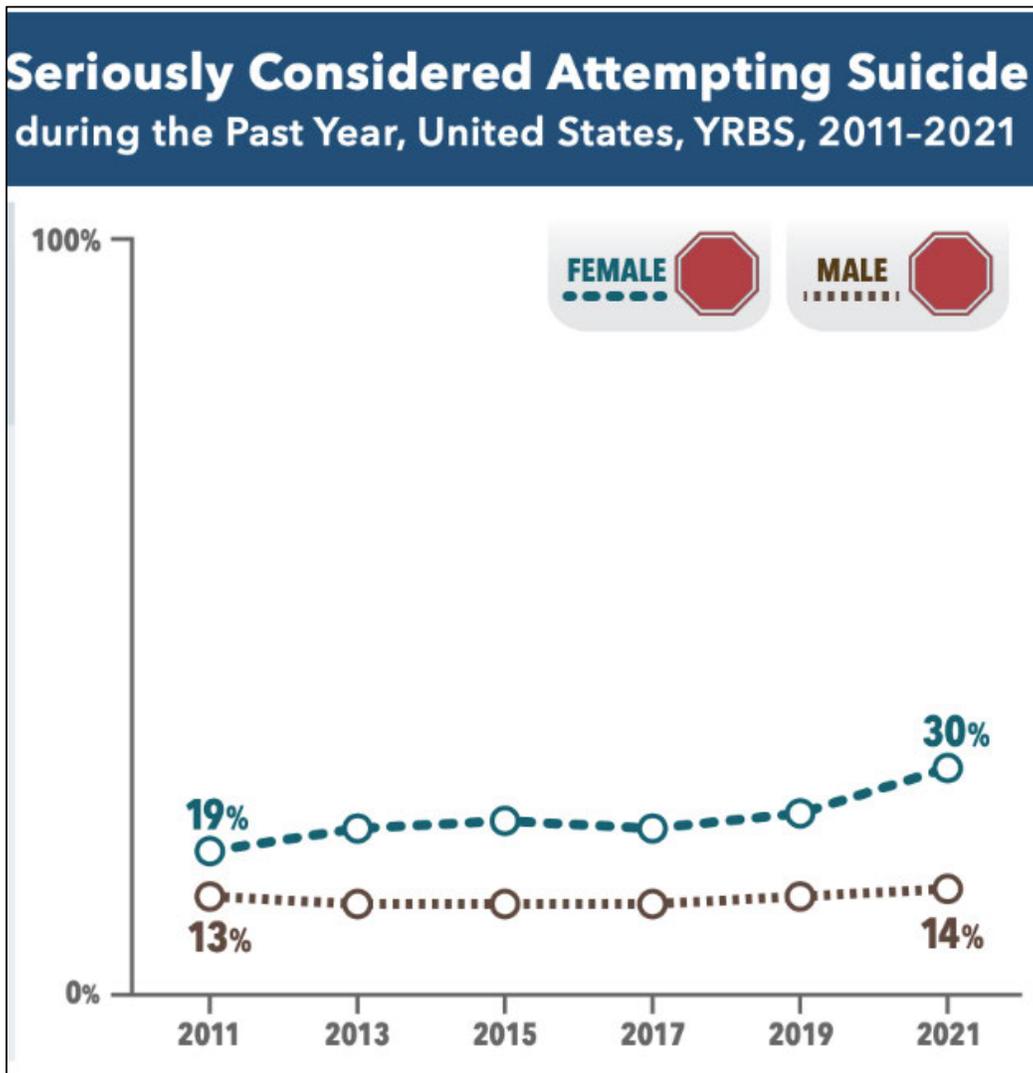
**Figure 34: DSMV-R Criteria for Depression**

- Five or more of the following A Criteria (at least one includes A1 or A2)
- ✓ A1 Depressed mood—indicated by subjective report or observation by others (in children and adolescents, can be irritable mood)
  - ✓ A2 Loss of interest or pleasure in almost all activities—indicated by subjective report or observation by others.
  - ✓ A3 Significant (more than 5 percent in a month) unintentional weight loss/gain or decrease/increase in appetite (in children, failure to make expected weight gains).
  - ✓ A4 Sleep disturbance (insomnia or hypersomnia).
  - ✓ A5 Psychomotor changes (agitation or retardation) severe enough to be observable by others.
  - ✓ A6 Tiredness, fatigue, or low energy, or decreased efficiency with which routine tasks are completed.
  - ✓ A7 A sense of worthlessness or excessive, inappropriate, or delusional guilt (not merely self-reproach or guilt about being sick).
  - ✓ A8 Impaired ability to think, concentrate, or make decisions—indicated by subjective report or observation by others.
  - ✓ A9 Recurrent thoughts of death (not just fear of dying), suicidal ideation, or suicide attempts.
- ✓ The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
- ✓ The symptoms are not due to the direct physiological effects of a substance (e.g., drug abuse, a prescribed medication's side effects) or a medical condition (e.g., hypothyroidism).
- ✓ There has never been a manic episode or hypomanic episode.

367. Depression typically lasts for weeks or longer and significantly impairs an individual's ability to function in daily life. While depressive symptoms can be part of a broader diagnosis of depression, they may also appear independently and be a cause of distress without reaching a clinical diagnostic threshold. Accordingly, some studies, and hence some meta-analyses, use depressive *symptoms* while others use the clinical diagnosis of depression. For example, the CDC as part of its ongoing Youth Risk Behavior Survey (YRBS) has been assessing

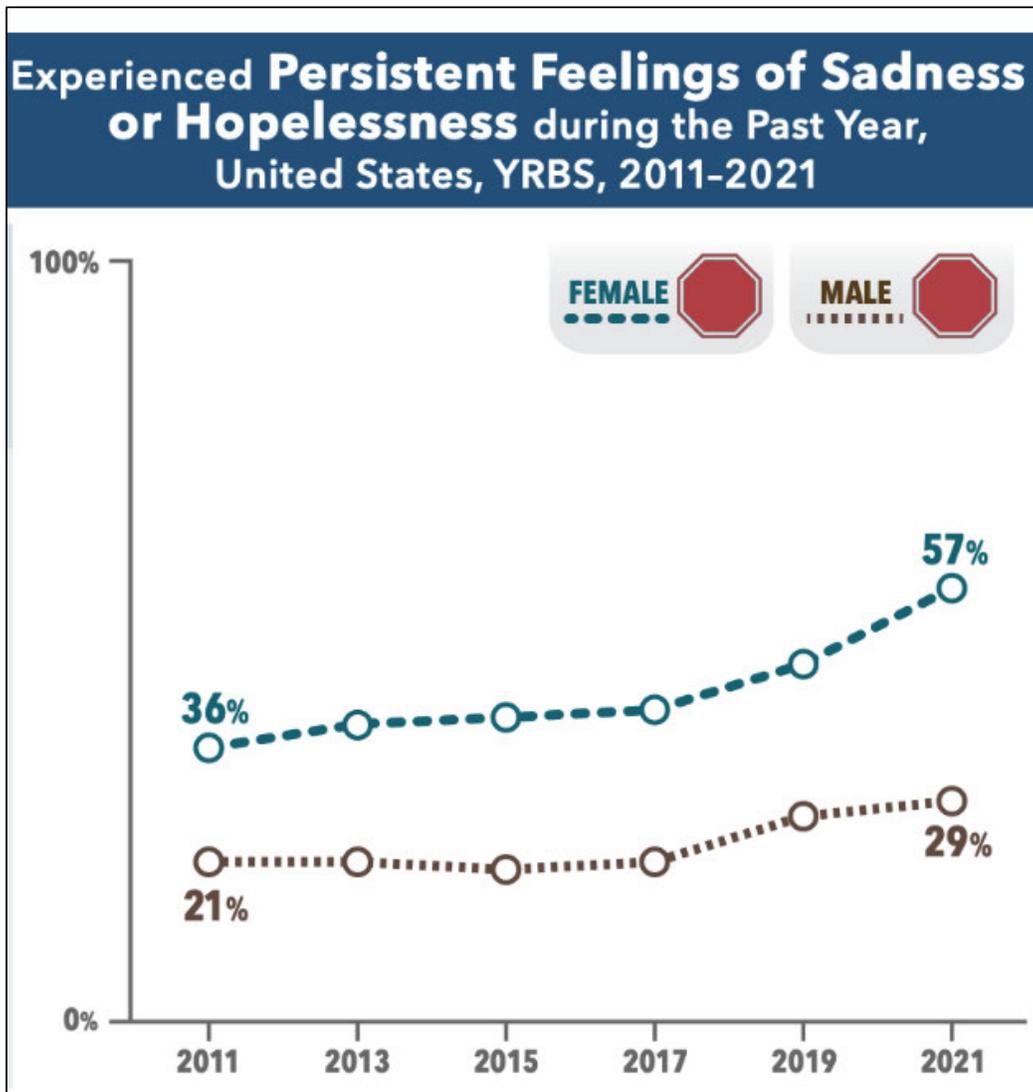
adolescents' feelings of sadness or hopelessness during the past year as well as whether they have seriously considered attempting suicide. It recently published 10-year trend data shown below.<sup>376</sup>

**Figure 35: Prevalence of Suicidal Ideation Over Time**



<sup>376</sup> Black MH, Milbourn B, Chen NTM, et al. The use of wearable technology to measure and support abilities, disabilities and functional skills in autistic youth: a scoping review. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*. 2020;8(1):48-69. doi:doi:10.21307/sjcapp-2020-006

Figure 36: Prevalence of Sadness over Time



368. While not measures of clinical (DSMV-R) “depression,” these data (particularly for females) have caused many to declare a “public health crisis” for teenagers in America and some

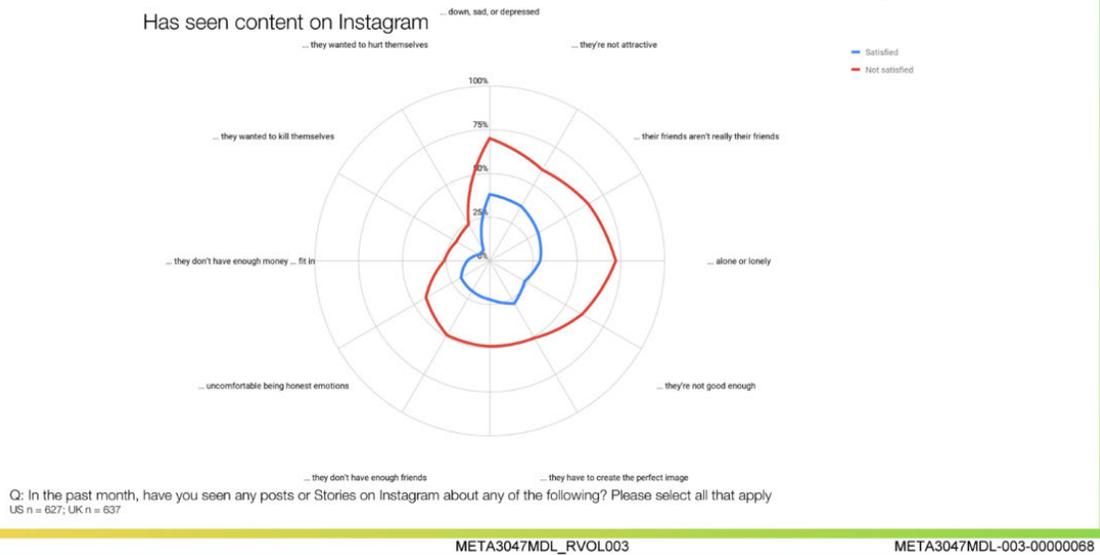
have implicated social media usage as the cause.<sup>377</sup> Although it is true that the rise of these depressive symptoms and the increase in depression and anxiety diagnoses mirror the rise of social media usage, that correlation in and of itself does not prove causation.

369. Sadly, as the slide below shows, Meta's own internal analysis reveals that the algorithms bring mental health related content to the fore more frequently for those who report being unsatisfied with their lives. For teens already prone to—or exhibiting—depressive symptoms, this sets up a very real potential negative feedback loop that both exacerbates symptoms and promotes negative content.

---

<sup>377</sup> Twenge JM. *iGEN: Why Today's Super-Connected Kids Are Growing Up Less Rebellious, More Tolerant, Less Happy-- And Completely Unprepared For Adulthood And (What This Means For The Rest Of Us)*. Atria Books; 2017:viii, 342 pages; Haidt J. *The anxious generation : how the great rewiring of childhood is causing an epidemic of mental illness*. Penguin Press; 2024: 385 pages; Twenge JM. Increases in Depression, Self-Harm, and Suicide Among U.S. Adolescents After 2012 and Links to Technology Use: Possible Mechanisms. *Psychiatr Res Clin Pract*. Summer 2020;2(1):19-25. doi:10.1176/appi.prcp.20190015.

## Teens who are unsatisfied with their lives are more likely to see content related to mental health on Instagram



Document 119: META3047MDL-003-00000029, -0068

370. Instagram leaders called the process their algorithms created “fee[d]ing the spiral,” as documented in this chat exchange between ██████████ and Wendy Gross (Head of Marketing Insights):

```

██████████ (11/02/2020 14:57:32 PST):
>interesting -- do you think it's because the content influences the mood, or vice versa, or both equally intertwined?
Wendy Tegge Gross (11/02/2020 15:00:20 PST):
>from qualitative work, it's intertwined
Wendy Tegge Gross (11/02/2020 15:00:34 PST):
>people feel low and then content makes them feel worse
██████████ (11/02/2020 15:05:10 PST):
>but this should support our "feeing the spiral" theory right? regardless of who starts it, someone feeling bad sees content that makes them feel bad, they engage with it, and then their IG is flooded w it
Wendy Tegge Gross (11/02/2020 15:05:18 PST):
>yup
    
```

Document 120: META3047MDL-003-00121808, -1808

371. A 2021 metanalysis of 62 studies evaluating social media usage and depression symptoms found a significant but “weak” association ( $r=.11$ ). However, the association between

problematic social media use and depressive symptoms was moderate ( $r=.29$ ).<sup>378</sup> The majority of the included studies were observational, cross-sectional, and relied on self-report of social media usage, weaknesses that reflect the current state of existing knowledge as well as the quality of the data available to most scientists outside of industry. A recent meta-analysis found that the risk of depression increased 13% with each hour of daily social media use.<sup>379</sup> In the chapter regarding social media use and depressive symptoms, the authors concluded that there are “consistent links between social media use and depression and evidence of causality.”<sup>380</sup>

372. One of the largest and best longitudinal, observational studies followed 6,595 subjects between 2013 to 2016 as part of the Population Assessment of Tobacco and Health (PATH) study. Participants’ self-reported social media usage at ages 13-16 was used to predict depression and anxiety (internalizing) symptoms at ages 14-17 adjusting for baseline (ages 12-15) internalizing problems. Three to six hours of social media at ages 13-16 was associated with a 60% increased relative risk of internalizing problems at ages 14-17 and greater than 6 hours per day was associated with a 78% increased risk.<sup>381</sup> Recall that the median use of TikTok per day is approximately 2 hours and the 75<sup>th</sup> percentile is 3.

373. A subsequent systematic review of longitudinal studies of “screen time” and mental health in young people reported that 1.5 out of 4 studies of social media usage found associations

---

<sup>378</sup> Cunningham S, Hudson CC, Harkness K. Social Media and Depression Symptoms: a Meta-Analysis. *Res Child Adolesc Psychopathol*. Feb 2021;49(2):241-253. doi:10.1007/s10802-020-00715-7

<sup>379</sup> Liu M, Kamper-DeMarco KE, Zhang J, Xiao J, Dong D, Xue P. Time spent on social media and risk of depression in adolescents: a dose–response metaanalysis. *Int J Environ Res Public Health*. 2022:19

<sup>380</sup> Handbook of Children and Screens, p.139.

<sup>381</sup> Riehm KE, Feder KA, Tormohlen KN, et al. Associations Between Time Spent Using Social Media and Internalizing and Externalizing Problems Among US Youth. *JAMA Psychiatry*. 2019;76(12):1266-1273. doi:10.1001/jamapsychiatry.2019.2325

for depression and 2 out of 2 found associations for internalizing symptoms (including the PATH study cited above).<sup>382</sup> Overall, the authors conclude that there are small ( $r=.10$ ) but significant associations between screen time and depressive symptoms in adolescents.<sup>383</sup>

374. Not included in that systematic review, because it was published subsequent to it, is perhaps the best observational study to date. Braghieri et al deployed a “quasi-experimental” design wherein they tested the effect of the “rollout” of Facebook to 775 college campuses between 2004 (the year it launched) to 2006.<sup>384</sup> Mental health outcomes were derived from the NHCA survey that is administered on a semi-annual basis to US college students since 1998. It includes many questions related to psychological health and well-being and the researchers used all of them to create a composite score. The individual outcomes and the composite one are displayed below in Figure 35. The blue “dots” represent the point estimates and the bars represent the 95% confidence intervals. Being on the right of the red line means the outcome is worse than before. Dots with bars that do not cross the red midline are statistically significant.

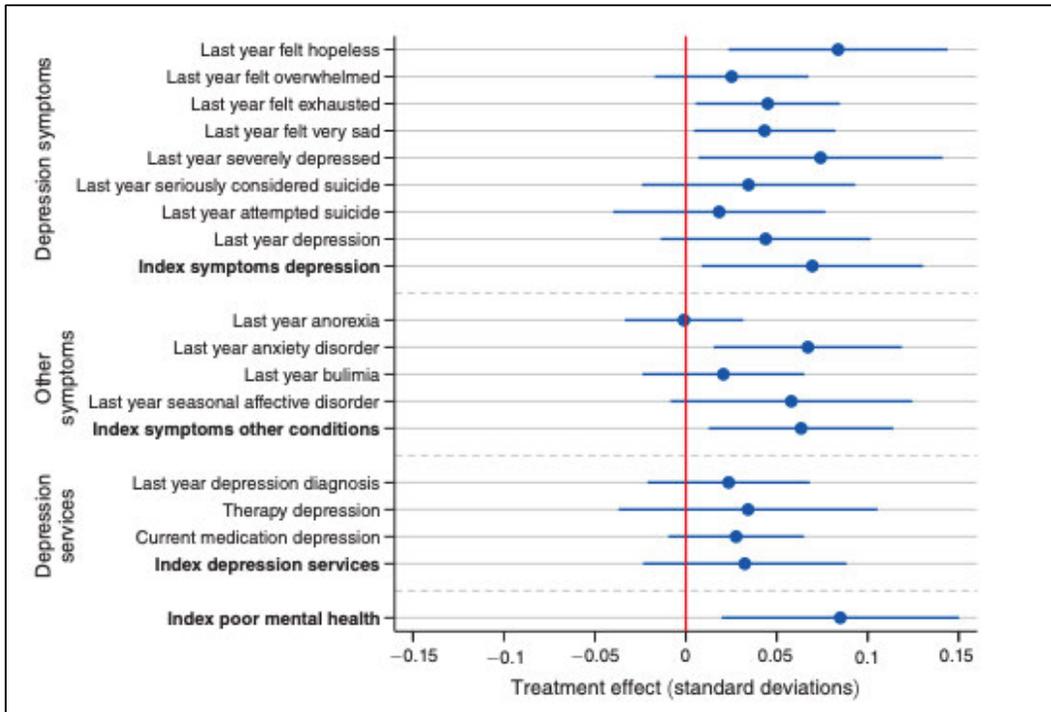
---

<sup>382</sup> Tang S, Werner-Seidler A, Torok M, Mackinnon AJ, Christensen H. The relationship between screen time and mental health in young people: A systematic review of longitudinal studies. *Clinical Psychology Review*. 2021/06/01/ 2021;86:102021. doi:<https://doi.org/10.1016/j.cpr.2021.102021>

<sup>383</sup> Tang S, Werner-Seidler A, Torok M, Mackinnon AJ, Christensen H. The relationship between screen time and mental health in young people: A systematic review of longitudinal studies. *Clinical Psychology Review*. 2021/06/01/ 2021;86:102021. doi:<https://doi.org/10.1016/j.cpr.2021.102021>

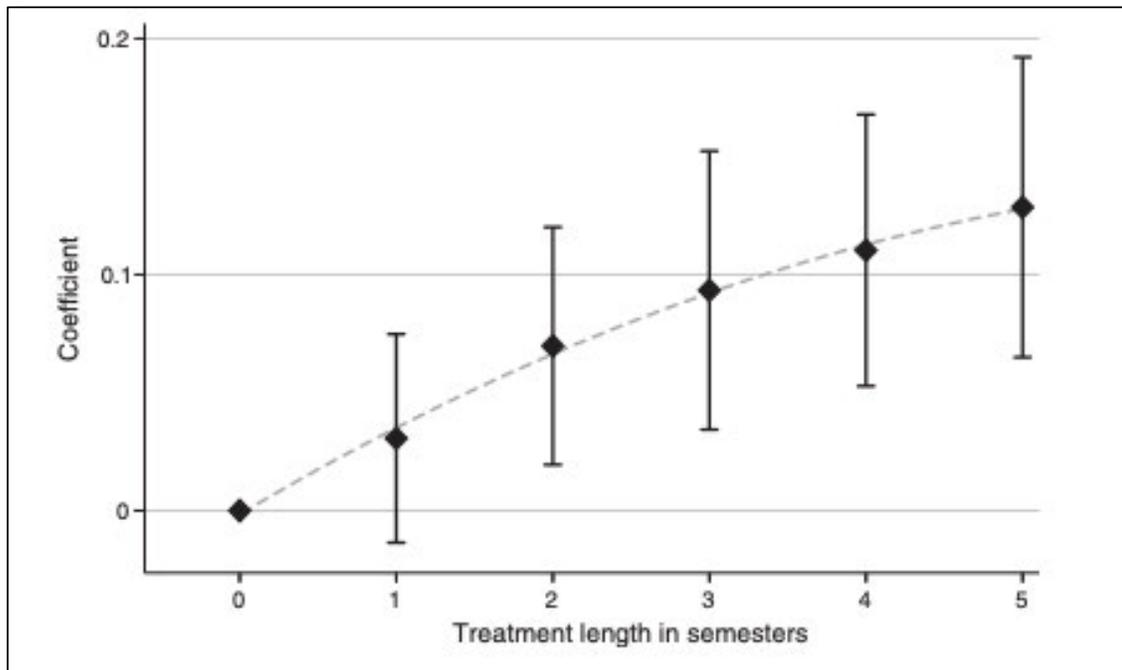
<sup>384</sup> Braghieri L, Levy Re, Makarin A. Social Media and Mental Health. *American Economic Review*. 2022;112(11):3660–93. doi:10.1257/aer.20211218

**Figure 37: Effects of Introduction of Facebook to College Campuses**



375. As can be seen, all but one of the dots falls to the right of the red line and the summary estimate is that the introduction of Facebook to a college campus resulted in statistically significant a .085 standard deviation unit decrease in overall mental health at that campus. To help benchmark that estimate, the authors compare it to the effect of job loss and find it to be about 22% of that. They further estimate that the introduction of Facebook to a college campus increases the percent of students who would meet the clinical diagnosis of depression or anxiety by 2%. Given a baseline estimate of 25% for depression and 16% for anxiety, this represents an 8% and 13% relative increase of each respectively. Next, the authors conducted a “dose-response” analysis in which they estimate the mental health effects based on the number of semesters given students were exposed. Those results are shown in the figure below.

**Figure 38: Effect on Mental Health on College Campuses by Exposure to Facebook**



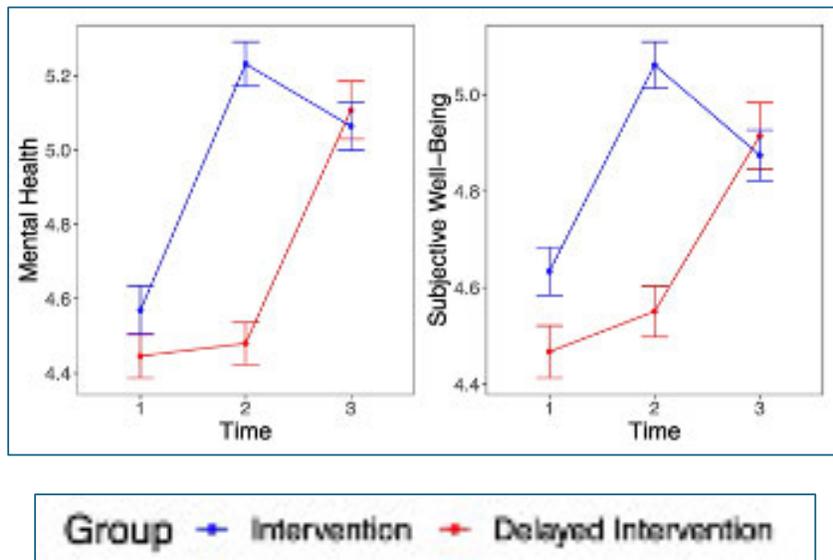
Again, these findings are consistent with the hypothesis that additional exposure to Facebook results in increased risk of poorer mental health at the campus level.

376. Two features of this analysis make the estimates conservative. First, the observed effect is at the *college* level without knowledge of whether or not individuals actually used Facebook or not. For example, if only 50% of students at college Y signed up for it, then the measured effect is diluted by the 50% that did not. Second, the selected years were just after the launch of Facebook when friends were largely limited to other college students, considerably less content was available (infinite scroll had not yet been invented), and sophisticated algorithms had not yet been deployed. In effect then, the study measured the impact of what by today's standards is an anemic version of Facebook and social media in general in terms of its command of one's

attention, its ability to deliver maximally engaging content, and the amount and type of content delivered.

377. A just published experimental trial used a “blocking” technology to effectively make smart phones “dumb” for 2 weeks by preventing them from accessing WiFi and the internet but allowing calls and text messages.<sup>385</sup> Over a four week trial, they assessed attention, mental health, and subjective well-being in 467 adults who were randomly assigned to use the blocker for the first two weeks or the last two weeks. The results on mental health and well-being are presented below.<sup>386</sup>

**Figure 39: Mental Health and Subjective Well-Being**



<sup>385</sup> Castelo N, Kushlev K, Ward AF, Esterman M, Reiner PB. Blocking mobile internet on smartphones improves sustained attention, mental health, and subjective well-being. *PNAS Nexus*. 2025;4(2)doi:10.1093/pnasnexus/pgaf017.

<sup>386</sup> Castelo N, Kushlev K, Ward AF, Esterman M, Reiner PB. Blocking mobile internet on smartphones improves sustained attention, mental health, and subjective well-being. *PNAS Nexus*. 2025;4(2)doi:10.1093/pnasnexus/pgaf017 at 4.

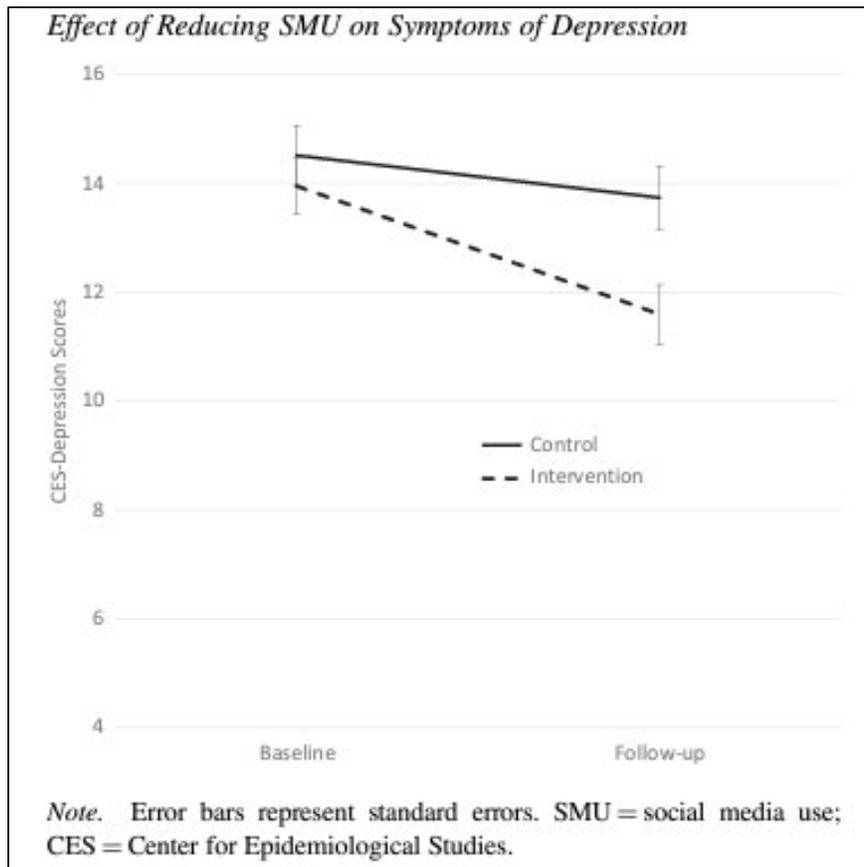
378. Both mental health and subjective well-being improved from T1 to T2 (baseline to 2 weeks later) and then regressed somewhat after the blocking was turned off (Effect Sizes .57 & .46 respectively  $p < .001$ ). Similarly, the delayed intervention group saw considerable improvements in both once they got the blocking program installed at T2. Although the blocking software would effectively prohibit all internet use, based on the preponderance of time spent on social media, it predominately reduced time on them.

379. Finally, the Davis social media reduction trial (discussed in Section VII.F) also assessed depression after three weeks of a 50% reduction in social media use.<sup>387</sup> Those results, presented below, show a difference of 2.36 points on the Center for Epidemiologic Studies of Depression Scale 10 (CES-D10) after the 3-week intervention.

---

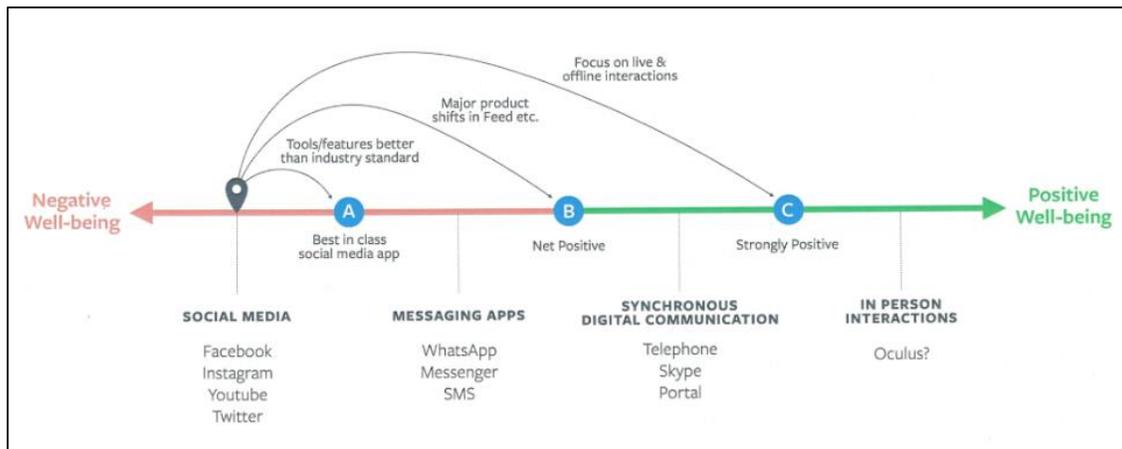
<sup>387</sup> Davis CG, Goldfield GS. Limiting social media use decreases depression, anxiety, and fear of missing out in youth with emotional distress: A randomized controlled trial. *Psychology of Popular Media*. 2025;14(1):1-11. doi:10.1037/ppm0000536.

**Figure 40: Effect of Reducing SMU on Symptoms of Depression**



Although a clinically meaningful difference in the CES-D10 has not been widely established, the observed effect size in this study (Cohen's  $d$ ) is .42 (.50 =medium). One important distinction of this study is that participants needed to be showing signs of distress at enrollment to be eligible for inclusion. The baseline average score on the CES-D10 was over 14 with above 10 being a clinically concerning score, and over 70% of subjects were above that threshold. As the authors indicate, it may be that social media usage is especially harmful to the mental health of individuals who are already showing signs of mental distress.

380. For its part, in 2019, Meta conceded that the average net effect of Facebook on well-being was slightly negative, a finding that was directly communicated to Mark Zuckerberg.<sup>388</sup> Its aspiration was to evolve into being on balance slightly net positive (a goal that was considered “really hard” to achieve).<sup>389</sup> Below is a figure presented at a meeting at which Sheryl Sandberg, COO of Facebook at the time, was allegedly present.



*Document 121: META3047MDL-003-00086233, -6243*

381. The largest publicly available experiment of Facebook effects is a National Bureau of Economic Research (NBER) working paper entitled “The Welfare Effects of Social Media,”<sup>390</sup> a study that Dr. Burke is aware of and says can be “taken seriously.”<sup>391</sup> In it, the authors recruited 2,743 users and elicited their willingness to accept payment to deactivate their accounts. Those randomized to the “treatment group” were paid \$102 to do so for four weeks (longer than the 1-

<sup>388</sup> Zuckerberg Dep. Tr. 263:17-265:20.

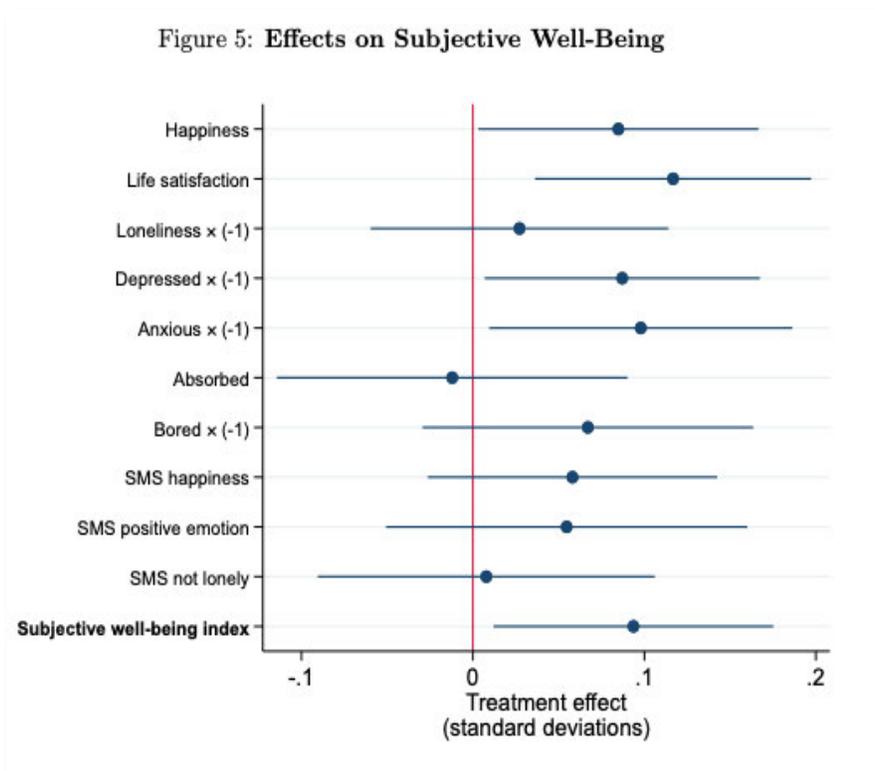
<sup>389</sup> Zuckerberg Dep. Tr. 261:7-11.

<sup>390</sup> Allcott H, Braghieri L, Eichmeyer S, Gentzkow M. The Welfare Effects of Social Media. *American Economic Review*. 2020;110(3):629–76. doi:10.1257/aer.20190658

<sup>391</sup> Moira Burke Dep. Tr. at 303:7-304:5

week threshold identified by Thrule as critical).<sup>392</sup> They measured a suite of outcomes using text message, surveys, and emails among other ways. Most relevant to this report are the items related to subjective well-being. As before, they looked at each item individually and then created a composite index of all of them.

**Figure 41: Effects on Subjective Well-Being**



382. Again, to benchmark the effect size of .09, the authors compared it to the effectiveness of psychological interventions including self-help therapy, group training, and individual therapy. For those they used a summary estimate derived from a metaanalysis of 39

<sup>392</sup> Thrul J, Devkota J, AlJuboori D, Regan T, Alomairah S, Vidal C. Social media reduction or abstinence interventions are providing mental health benefits – reanalysis of a published meta-analysis. *Psychology of Popular Media*. In press;

randomized controlled trials. They concluded that deactivating Facebook increases subjective well-being by about 25-40% as much as standard psychological treatments, a key finding Meta cites in a presentation related to the potential harm it is causing in 2019.<sup>393</sup>

383. For reasons that are not entirely clear, Mathew Gentzkow of Stanford (“an exceptionally acclaimed Stanford economist,” per Brent Harris<sup>394</sup>), an author on the Alcott et al “deprivation” experiment, was in contact with people at Meta prior to the publication of the first study alerting them to the findings.<sup>395</sup> Within Meta, there was concern about openly collaborating with him and his coauthors since there were, according to an email from John Hegeman, “risks about what [the study] might find and ... the substantial corresponding PR of those findings.”<sup>396</sup> David Ginsberg, who started as a Director of Research and eventually became Vice President of Global Communications and Public affairs,<sup>397</sup> replies that he is favor of the study in principle, but given the risk, he is “not convinced [they] need or should do it with external partners.”<sup>398</sup>

384. Rather than participate in this one, the suggestion is made to launch their own study fronted by Nielsen so as to mask their involvement in it. But Annie Franco, another Meta researcher, responds that while they should do their own study, if the Stanford groups is negative and Meta’s is positive they will have a hard time convincing people that their results are credible.<sup>399</sup> In spite of that concern, Meta decides to launch its own internal study called the “Nielsen thing.” In the meantime, the Alcott study was completed and shared with Meta researchers.

---

<sup>393</sup> Haugen\_00010114, -0123-24

<sup>394</sup> See META3047MDL-014-00048189, -8189

<sup>395</sup> META3047MDL-040-00586960 at -6965

<sup>396</sup> META3047MDL-040-00590304 at -0308

<sup>397</sup> Ginsberg Dep. at 53-55

<sup>398</sup> META3047MDL-040-00590304 at -0307

<sup>399</sup> META3047MDL-040-00590304 at -3087

385. Alex Leavitt, a lead quantitative UX researcher at Meta, was impressed with the Alcott study, responding to the abstract “we think the paper is great, and we’re excited to learn more from it and help our teams understand the implications.”<sup>400</sup> But they are worried that the press “will inflate the narrative.” Later, Meta tried to get them to recast their relative effect size to make it not seem as substantive.<sup>401</sup>

386. Ultimately, Gentskow gave leaders at Meta a heads up that their paper would be in the *New York Times*.<sup>402</sup> The press that Alcott’s study received further motivated Meta to advance the Nielsen project.<sup>403</sup> They ultimately decided to pilot the Nielsen study internally and then partner with academics “more fully afterwards.”<sup>404</sup> However, even before it launched a Meta memo contained the following caveat:

Publication / External communication: There are concerns about publishing our final report, and even talking about this study externally. There were even fears of running this study and what will we do if we find out negative things about Facebook (comparisons to cigarette studies). We had to get this study reviewed all the way up to Nick and Javi.

*Document 122: META3047MDL-019-00095647 at -9569*

387. In fact, the sensitivity to potential leaks of the findings or the study itself were great enough that being employed in journalism was an exclusion criteria for the study.<sup>405</sup> As the study ensues, there is this notable exchange amongst well-being scientists in which they discuss whether or not the findings will be made public:

---

<sup>400</sup> META3047MDL-040-00586960 at -6966

<sup>401</sup> META3047MDL-040-00586960 at -6964

<sup>402</sup> META3047MDL-014-00048189

<sup>403</sup> META3047MDL-003-00176564 at -6568

<sup>404</sup> META3047MDL-019-00095647 at -5648

<sup>405</sup> META3047MDL-020-00572432 at -2435

[REDACTED] (9/19/2019 15:43:54 PDT):  
>so if we get the nielsen results back, if they are good would we publish or is there no intention of publishing?

[REDACTED] (9/19/2019 15:45:09 PDT):  
we just aren't planning on it

[REDACTED] (9/19/2019 15:49:36 PDT):  
>ok. and just curious, do we think the results will be that negative? ive got to think that even if slightly negative they wont be THAT bad...i feel like all the well being research ive seen if it says FB isnt good for people it's always jsut like a smallllllll percentage over neutral

[REDACTED] (9/19/2019 15:49:52 PDT):  
>that is correct

[REDACTED] (9/19/2019 15:49:55 PDT):  
>im being naive but i feel like that this will bbe fine

[REDACTED] (9/19/2019 15:50:01 PDT):  
>i think we ll see some positive

[REDACTED] (9/19/2019 15:50:06 PDT):  
>some negative

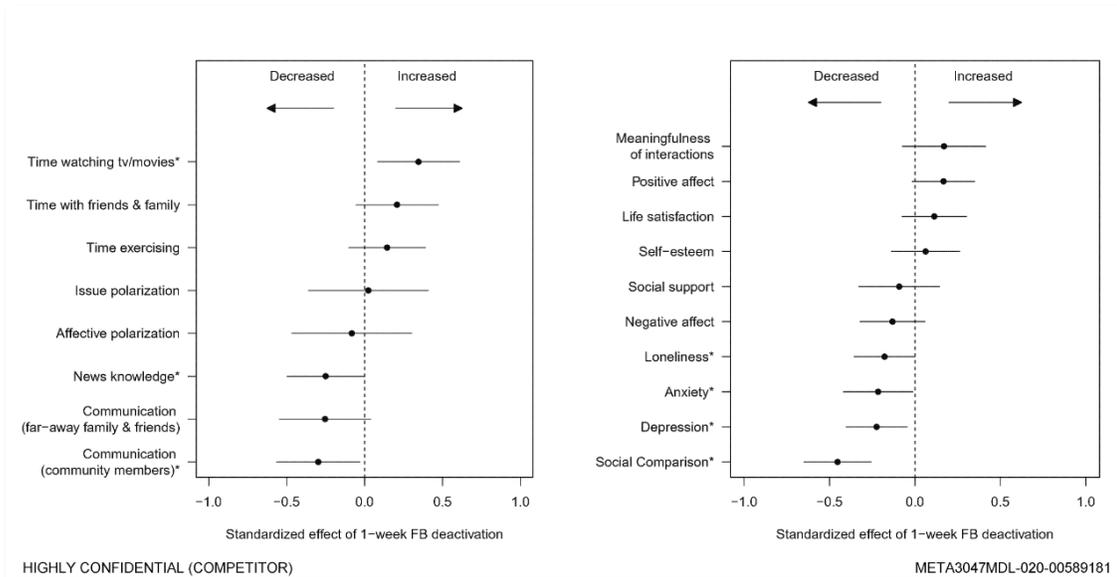
[REDACTED] (9/19/2019 15:50:06 PDT):  
>yeah i agree

[REDACTED] (9/19/2019 15:50:10 PDT):  
>and most neutral

*Document 123: META3047MDL-040-00031943 at -1943*

388. Later in this same exchange, [REDACTED] (News Feed and Algorithmic Transparency Communications Manager)<sup>406</sup> responds “someone just said to me if the results are bad and we don’t publish and they leak is it going to look like tobacco companies doing research and knowing cigs were bad and then keeping that info to themselves.” [REDACTED] (User Experience Researcher) attempts to reassure him “no, because we will do something about it.” The results of the “pilot” are presented below.

<sup>406</sup> <https://www.linkedin.com/in/james-raimo-64180b102/>



Document 124: META3047MDL-020-00589181 at -9181

389. After 1 week (which in prior studies was too soon to see effects and sometimes even found withdrawal effects), most of the findings are in the direction that is unfavorable for Meta, and some of the results including less social comparison, depression, anxiety and loneliness, were statistically significant. Their own internal summary was as follows:

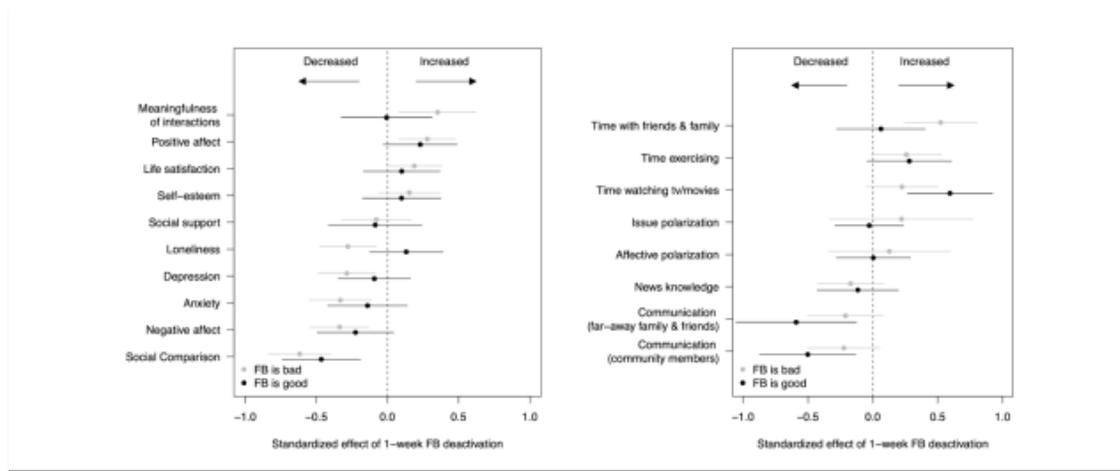
We found that those who stopped using Facebook for a week:

- Reported lower feelings of depression, anxiety, loneliness, and social comparison
- Reported fewer interactions with community members
- Reported more time watching tv/movies and spent less time on social media
- Had lower news knowledge

Document 125: META3047MDL-047-00058006 at Slide 2

390. These results are consistent with Alcott’s study as well as with numerous other longitudinal and experimental studies and the differences were evident after only one week. But the team decided subsequently to then stratify the results to see what role “expectations” played in

the observed results. To do this, they asked people at enrollment if they think Facebook is bad for their well-being and compared the results for those that believed it was and those that believed it did not. Here is what they found in that stratified analysis:



Document 126: META3047MDL-047-00058006 at Slide 7

391. Conditioning on whether users believed Facebook was bad for them considerably attenuated the effect sizes except for social comparison which remained the same. Notably, many of the effects are still in the same direction but are no longer statistically significant, and the fact that social comparison was unaffected was itself an important finding. Their read of the results, that expectations drove outcomes, presupposes that the findings were based entirely on people being primed to believe that Facebook worsened well-being. But of course, there is an alternative explanation which is that some people rightly deduced that for *them*, Facebook was harmful. That is, they had an accurate assessment of its effect on them. This is completely consistent with the differential susceptibility theory and “addiction theory.” Quitting drinking for example benefits those who think it’s a problem for them more than those who do not because people who think they have a drinking problem generally do.

392. Later, Paul Dow writes to Moira inquiring what happened with the study:

Paul Alexander Dow (5/04/2021 12:09:11 PDT):  
>hey. you might have told me this, but, if so, i forgot. what ever happened with the Nielsen deprivation study? I think I lost track when covid hit. I know there was one or more pilot studies completed, but did they ever do the full thing?

Moira Kathleen Ballantyne Burke (5/04/2021 12:12:07 PDT):  
>There were two one-wave pilot studies and the findings have been shared with leads internally, but afaik we're not planning to do a longer-term study. The findings were basically that people's expectations totally drove the results, and we weren't able to use the data to disentangle those expectation effects. Summary: <https://fb.quip.com/MkxAA3gfy84C>

Paul Alexander Dow (5/04/2021 12:29:55 PDT):  
>well, shit. that seems like a result worth publicizing on its own. not that anyone would want to hear it from us. maybe that is a study for FORT, what is the role of expectation effects on social media deprivation studies. i see Safaa has a deprivation study in her list of research projects for the Problematic Use stream.

*Document 127: META3047MDL-040-00213905 at -3905*

393. The fact that they could not disentangle the effects especially given that social comparison remained significant is hardly justification for not publishing (or trying to publish) the results. In fact, if they believed their results were robust, they should have subjected their assertion that the observed effects were driven by the lack of blinding, to peer review and then made them public. As ██████████ says in an email on October 23, 2020:

The **main findings** of the study are:

1. People who stopped using Facebook for a week reported an improvement in certain types of emotions (e.g., depression, anxiety, loneliness, and social comparison), but no change in other outcomes (e.g., political polarization, or meaningful interactions).
2. Most of the effects we found are linked to people's beliefs about Facebook. For example, the reductions in feelings of depression, anxiety, and loneliness associated with deactivation appeared mostly among participants who believed (prior to participating in the study) that Facebook was bad for them.

**The main implication of our findings:** External criticisms highlighting the negative impact of the company products on well-being related outcomes may be overblown. For example, if beliefs are the result of the existing media narrative around the company, the findings in the Nielsen study, and similar external work, will overestimate the true effect of Facebook (akin to placebo effects).

*Document 128: META3047MDL-165-00000428 at -0429*

394. Their interpretation of their findings is, scientifically speaking, what is “overblown”. A post hoc analysis showing that the effects were larger for those who believed that they had a problem at enrollment hardly invalidates the findings. If anything, it is consistent with

other studies that found that people with baseline mental health symptoms or problematic usage benefit more from diminishing or eliminating usage. He goes on to say that he is “working with Comms to determine the timing and framework under which we would make the study public if we chose to do so.” Given that they have not (yet) released it, one can deduce that they have buried it and perhaps have mimicked the tobacco companies as some of them feared they might.

395. Recently, Allcott, Gentzkow et. al. repeated this seminal SM deprivation study with some notable differences. First, among the 27 authors, there are 7 affiliated with Meta. Second, related to that collaboration, deprivation in this study was accomplished by central deactivation of the account by Meta (rather than voluntary abstinence). Third, the study included both Instagram and Facebook (more on that later) whereas the prior one was limited to Facebook alone. Fourth, they collected passive sensing data on a subset of participants which enabled them to track if people spent additional time offline or on other apps. Finally, they randomized participants to deactivate for 1 week vs. 6 weeks. This differed from the prior study in two important ways: it was longer in duration and better blinded since both groups had reason to believe that they were in the intervention arm.

396. The scale of the study was impressive and the design was incredibly robust. They recruited over 19,000 Facebook users and 15,000 Insta users making it by far the largest publicly available experiment of SMS use and they attempted to address demand effects. However, before delving into the findings, two important attributes of the study warrant mentioning. First, only 1% of people offered the opportunity to participate in the study ultimately did. This has implications for the generalizability of the findings which the authors acknowledge. It’s worth considering which way(s) participants might differ from non-participants. The most obvious one is that they are willing to “deactivate” their account which likely means they are less likely to be compulsive

or problematic users. By analogy, a study recruiting people to participate in dry January (or not) would assuredly under enroll problematic drinkers. Second, participants were free to use other apps as they wished and the findings suggest that in the Instagram group, they effectively replaced time on that app entirely with time on others (including Snap, YouTube, and TikTok) which would attenuate the findings insofar as those social media sites have many similar features and effects.

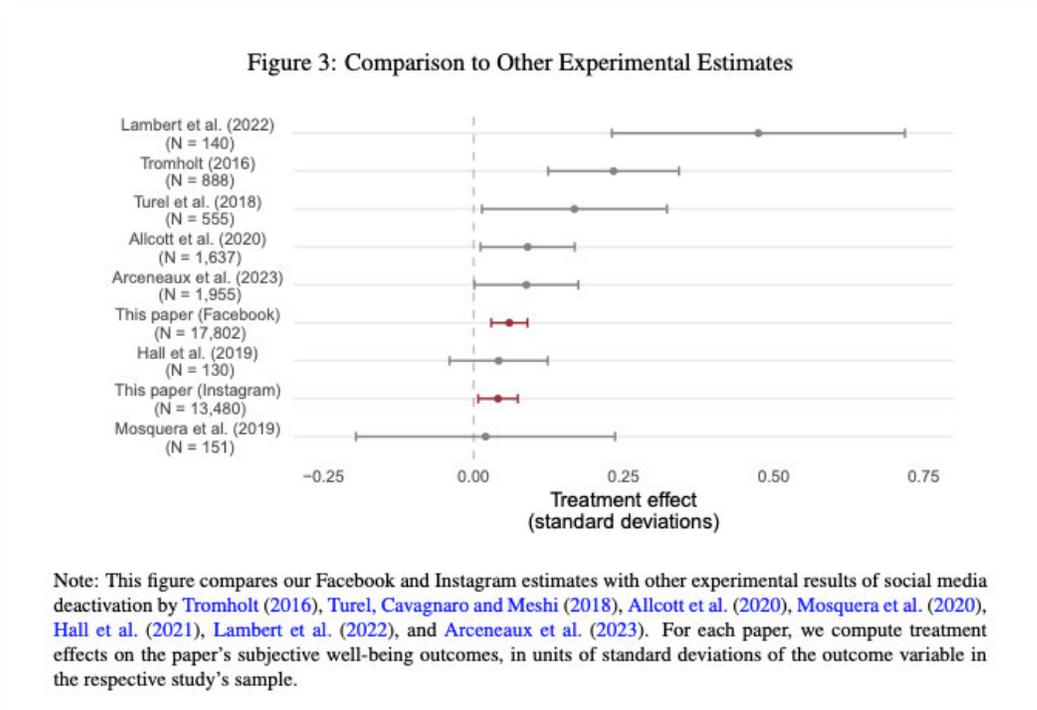
397. For the purposes of this report, I will summarize the Instagram results since they are most relevant to the plaintiffs in the case. Deactivation of Instagram resulted in a .041 standard deviation improvement in the emotional state index. In a sub-analysis, the effect was largest for women 18-24 with a .11 standard deviation improvement. The authors contextualize these effects by comparing them to those found in a meta-analysis of 419 randomized trials of nine types of psychological interventions which improve subjective well-being by .16 to .42 standard deviations.<sup>407</sup> Seen in this light, deactivation of Instagram is between 10 and 25% as effective as psychotherapy for all people and between 25 and 69% as effective for young women. However, from a public health standpoint, given that ready access to psychotherapy is both infeasible and unscalable, the overall effects of SM reduction on population health are likely considerably greater.

398. The authors of the paper go further and contextualize their findings compared to other experimental studies of SM deprivation with durations great than one week. Their summary represents an independent collection of relevant experiments as acknowledged by the Meta employees included as authors. Those results are presented in their Figure 3 below:

---

<sup>407</sup> Van Agteren J, Iasiello M, Lo L, et al. A systematic review and meta-analysis of psychological interventions to improve mental wellbeing. *Nature Human Behaviour*. 2021/05/01 2021;5(5):631-652. doi:10.1038/s41562-021-01093-w

**Figure 42: Summary of Relevant Experiments Externally Published and Known to Meta**

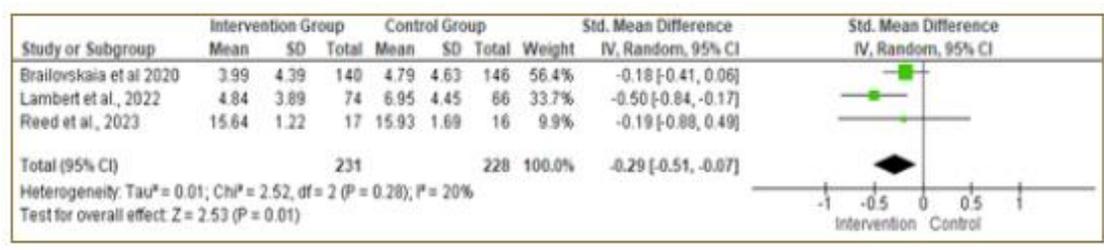


Notably, every study had positive effects and 7 of 9 were statistically significant. The magnitude of the effect in this study was smaller than most others.

399. One final point regarding this study warrants mentioning. It is error to claim that the only data that can establish causality must come from experimental studies. This, the largest and best experimental study done to date, highlights two salient problems with that contention. First, given low participation rates, it will be difficult to generalize from experiment participants to the overall population. Second, given substitution effects, deprivation studies of one site will measure the net effect of reducing use of one SM site while increasing the usage of others. In my opinion, both of these limitations make the observed effects conservative and argue for the consideration of all reliable studies in trying to determine causality.

400. A metaanalysis assessed the effects of experimental evaluations of “digital detox” on mental health outcomes. Again, these studies tested the putative benefits of a period of abstinence from social media sites.<sup>408</sup> There was considerable and statistically significant heterogeneity in both the approaches and the outcomes ( $p < .000001$ ). While the overall effect of “detox” on well-being was null (which is not surprising given the heterogeneity) the effects on depression was moderate and significant with an effect size of  $-.29$  (see below).

**Figure 43: Summary of Metaanalyses of Experimental Evaluations of ‘Digital Detox on Mental Health Outcomes**



This study also neglected to include the Allcott experiment for unclear reasons which would have further enhanced the results and made them more robust given its findings.<sup>409</sup>

401. Meta itself conducted a study of 6,000 Instagram users to test the effect of social comparisons on affect. They did so by randomly sequencing questions related to overall well-being and those related to negative social comparison. Below is a schematic recreated for image clarity from Haugen\_00000797, -0864.

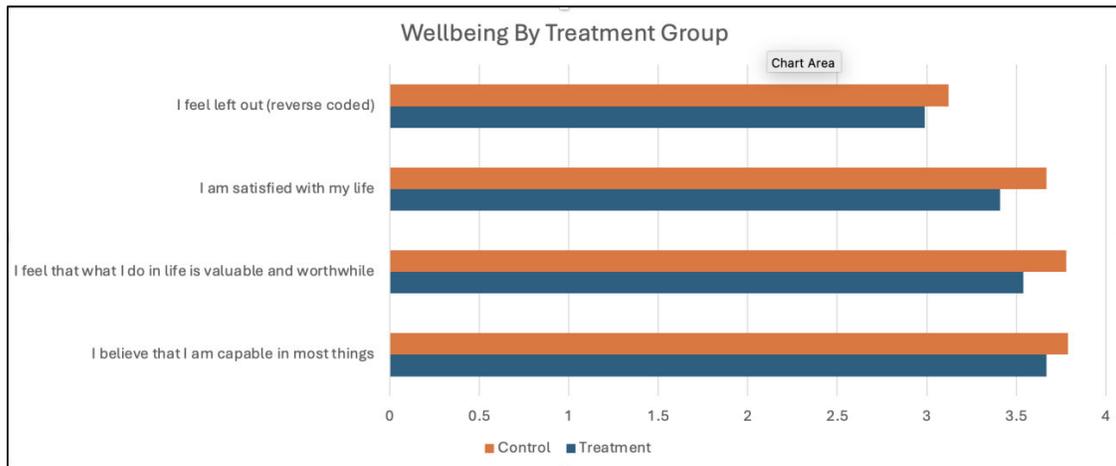
<sup>408</sup> Ramadhan RN, Rampengan DD, Yumnani DA, et al. Impacts of digital social media detox for mental health: A systematic review and meta-analysis. *Narra J.* Aug 2024;4(2):e786. doi:10.52225/narra.v4i2.786

<sup>409</sup> Marciano L, Schulz PJ, Camerini A-L. Cyberbullying Perpetration and Victimization in Youth: A Meta-Analysis of Longitudinal Studies. *Journal of Computer-Mediated Communication.* 2020;25(2):163-181. doi:10.1093/jcmc/zmz031



*Document 129: Meta Trial of Instagram "Priming" Replicated from Haugen\_00000864*

402. Both groups were asked the same questions but in a different order. In this way, the study tested the effect of priming of negative social comparisons on Instagram. Priming is a psychological phenomenon where exposure to a stimulus influences how a person responds to subsequent, related stimuli, often without conscious awareness. For example, seeing the word “bakery” makes someone recognize the word “pie” faster than an unrelated word like “telephone.” Below is a recreation (again for clarity) of a figure comparing responses to 4 of the 7 questions on the well-being assessment.



*Document 130: Wellbeing Treatment Group Chart Replicated from Haugen\_00000797, -0866*

403. For 4 of the 7 questions, the experimental group that was primed fared worse as a result of it; the other 3 were not statistically different from one another. Given that this was an experimental design, the authors appropriately conclude that these associations are “causal.”<sup>410</sup> What was experimentally manipulated here was thinking about negative social comparisons. While this was prompted in the context of the study design, in real life it presumably happens naturally when teens experience negative social comparisons on Instagram. Indeed, the majority (51%) of Instagram users report conducting social comparisons on the site.<sup>411</sup> For females in particular, these negative feelings of social comparison can lower self-esteem which, as reflected in in the conceptual model, Document 1 **p** , also leads to and exacerbates depression.

404. There are many explanations as to why social media use is linked to depression. Among the many mechanistic pathways (including loss of sleep, body dysmorphia, etc.) by which time spent on social media might adversely affect children’s mental health and well-being, one

<sup>410</sup> Haugen\_00000797, -0865

<sup>411</sup> Haugen\_00000797, -0797

must consider unwanted harmful online sexual experiences. This is addressed in detail in Section X.I below. Other pathways include features that lead to social comparison and negative body image and body image comparisons. Time spent on social media can replace otherwise positive mental health activities, including time with friends and family. Social media use can interfere with sleep, which is associated with depressive symptoms in children. Finally, rabbit holes can trap children in a depressive spiral.

405. For its part, TikTok’s own documents reveal a recognition that their highly effective algorithms also can create “rabbit holes” that can cause people with depression to spiral.

**Project 1: Dispersing Depression Videos** @杜婧璐 @沈思予

- \* **Depression**
  - o Goal: Reduce rabbit holes of depression content
  - o Gap1: Ability to detect depression content
  - o Gap2: Algo ability of dispersing depression content

*Document 131: TIKTOK3047MDL-002-00094384, -4399*

406. TikTok acknowledges the existence of “rabbit holes” created by their algorithms and the problems they can pose when they aggregate content:

## Trust & Safety: Why We Care

Rabbit holes raise trust and safety concerns for the following reasons:

1. Leads users to harmful content they would not have encountered otherwise.
2. There will always be grey-area policy cases. On a one-off basis they pose little risk of harm, but that's not the case if all grey-area cases of a particular type of content are concentrated on a user's feed.
3. Users could encounter violative mismoderated content in a high concentration
4. R2 policies (conspiracy theory, stereotypes, some sexualized content policies) are enforced at 12,000vvs. R2 violative low vv content could be concentrated on a user's feed before it hits moderation.

*Document 132: TIKTOK3047MDL-002-00064418, 4418-19*

407. In fact, other documents call out what they refer to as a “negative affect filter bubble” in which a user “regularly sees depressing, triggering, or otherwise inappropriate content.”<sup>412</sup> They estimate the “number of daily average users in such bubbles to be approximately

---

<sup>412</sup> TIKTOK3047MDL-002-00091521, -1524

1%.<sup>413</sup> Based on TikTok’s daily average user data from 2022, this represents 6,955,514 13–17-year-olds and 20,599,455 18-24-year-olds.<sup>414</sup> Once again, small numbers (and small effects) applied at scale can affect lots of people.

408. Later in the same exchange, it is noted that “The filter bubble problem on TikTok was publicly broadcasted by the *Wall Street Journal*, very negatively affecting our brand image.”<sup>415</sup> This last sentence epitomizes a pervasive attitude that is apparent in many internal company documents. A problem, “the negative affect filter bubble,” that is recognized internally as existing, is given added saliency (and resource) once it is publicized and impacts image/revenue. This is corroborated (with respect to Meta) by Ms. Jayakumar in her deposition when she is asked “In your experience, did the amount of external scrutiny from the press regulators and civil advocacy groups play into the priorities that Meta leadership set internally?” and she replies “Yeah, significantly so.”<sup>416</sup> Dr. Lee corroborates this in in an internal chat on 10/05/21 related to the Haugen leak, “Although I have real issues with how the whistleblower framed some of this work, I respect why she felt like this was the only way to push for change- if leadership heard this internally, it wouldn’t have come to this.”<sup>417</sup> Similarly, ██████████ (FB Production Engineer) says in a 2021 chat that, prior to the leaks, the “teen health problem” had not been treated with sufficient urgency:

---

<sup>413</sup> TIKTOK3047MDL-002-00091521, -1524

<sup>414</sup> TIKTOK3047MDL-002-00098058, -8060

<sup>415</sup> TIKTOK3047MDL-002-00091521, -1527

<sup>416</sup> Vaishnavi Jayakumar Dep. Tr. at 125:21-126:9

<sup>417</sup> Alison Lee Deposition Exhibit 13 at -1099



I wrote more about this today in more depth, but I think that the fact that the research referenced by the WSJ initially came out almost two years ago, and we have still not meaningfully moved the needle on the overall problem (as far as I can tell, I haven't seen any followups to the contrary) is a reasonable indication that we have not taken the teen health problem seriously.

Compare this to how we handle a problem we care more about. For example, there is currently a **SEV1 for Reels production growth being 3% lower than expected**. For this SEV, we have daily updates, posts, probably a war room going somewhere. We would need that same kind of concerted effort with dedicated product designers, engineers, and data scientists collaborating to measure and improve our products over a long period of time.

I don't doubt that we're making some progress on teen well-being, but it seems clear that there is a lot more we could be doing.

Document 133: META3047MDL-062-00000129, -0134

409. TikTok considered a solution for people “in extremis” who feel “trapped in a rabbit hole of ineffectively personalized content.”<sup>418</sup> It entailed giving them the opportunity to “reset” the algorithm but even for that allegedly “small” percentage of users (1%), the concern was raised about potential loss of ad revenue.

- "... [it would] help vulnerable users who need to distance themselves from their current viewing preferences. The research indicates that we could offer a reset option for those 'in extremis' and the majority would not misuse it. Focus group research: Providing Users With Choice Over Our Algorithm (focus groups across 5 EU countries)
- Algo proposed "non-personalized feed" in consideration of algo explainability and technical feasibility.

	Algo proposed plans	Pros	Cons
1	Non-personalized feed	• High feasibility	• Ads revenues

Document 134: TIKTOK3047MDL-002-00091625, -1625

<sup>418</sup> TIKTOK3047MDL-002-00091621, -1625

410. Consistent with corporate policy, this proposed change was subject to A/B testing.<sup>419</sup> Those test results found that “~32% of users turn off NP Feed within the week following the change” while “consumption metrics such as play duration, likes per uvv [Unique Video Views], comments per uvv, and shares per uvv decline.”<sup>420</sup> TikTok’s researchers go on to state that these results “confirms an idea that we already hold as true: that personalization is necessary in order to provide our users with engaging content.”<sup>421</sup> This is despite nearly 2/3 of users keeping the default non-personalized feed state and increases in “UVV and session count.”<sup>422</sup> Not only is this approach scientifically questionable, it also shows how TikTok’s researchers interpreted their data to support their company’s ultimate goal—increasing time spent on the platform.

411. In my opinion, to a reasonable degree of medical and scientific certainty, the totality of the evidence supports a causal relationship between social media use and depressive symptoms. There are numerous mechanistic pathways that explains this relationship. These include addictive design, features that increase negative social comparison, and algorithms that create problematic rabbit holes.

412. There is ample evidence within internal documents and depositions that support a causal relationship between SM and depression or depressive symptoms. Internal documents reflect a recognition that users were experiencing these symptoms as a result of the algorithms the defendants were developing and deploying. There is little evidence that that the companies undertook to meaningfully mitigate the risk of children developing depression. There is little

---

<sup>419</sup> TIKTOK3047MDL-002-00091621, -1630

<sup>420</sup> TIKTOK3047MDL-112-04262174, -2176

<sup>421</sup> TIKTOK3047MDL-112-04262174, -2176

<sup>422</sup> TIKTOK3047MDL-112-04262174, -2176

evidence that the company disclosed to parents, children, or the medical community the internally recognized risk of harm.

## **ii) Anxiety**

413. Like depression, anxiety can be both a clinical diagnosis and a constellation of problematic symptoms. The DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, 5th Edition) outlines specific criteria for diagnosing 6 different types of anxiety disorders including: Generalized Anxiety Disorder, Panic Disorder, Social Anxiety Disorder, Specific Phobia, Separation Anxiety, and Agoraphobia. Each disorder has its unique set of criteria, but all share the common feature of excessive fear or worry. Since the specific disorders are rarely clinically diagnosed in SM research studies, a full summary of each will not be provided but all share these general criteria. The fear, anxiety, or avoidance must:

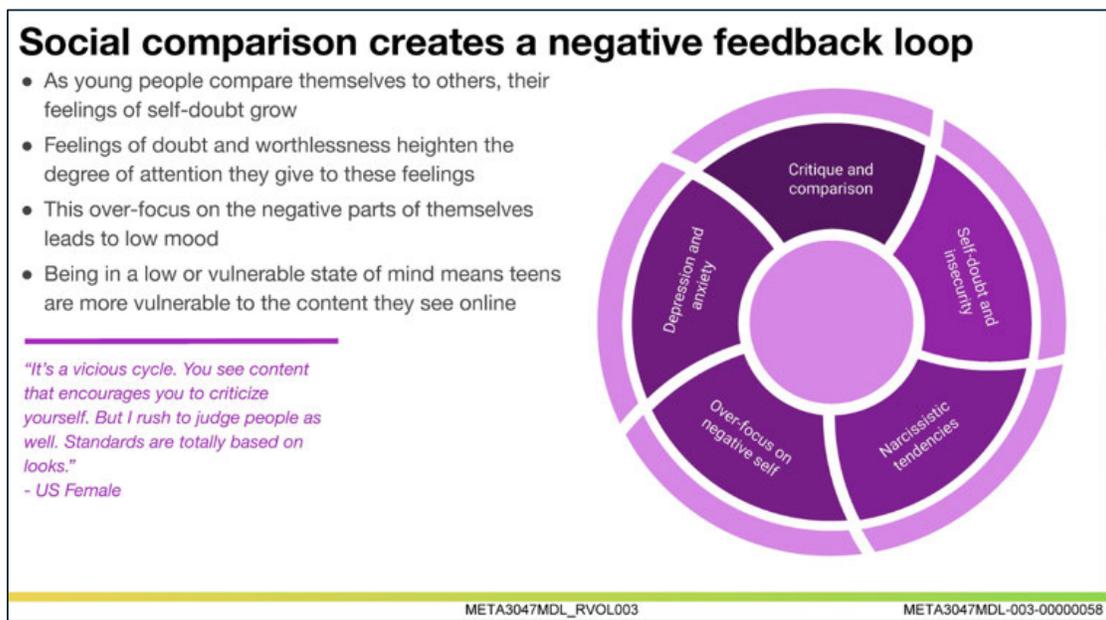
- Cause significant distress or impairment in functioning.
- Not be due to substances, medications, or a medical condition.
- Not be better explained by another mental health disorder.

414. From a mechanistic standpoint, both FoMo and social comparison can create anxiety. For example, several of the questions from Przybylski's FoMo scale are proxies for anxiousness:

1. I fear others have more rewarding experiences than me.
2. I fear my friends have more rewarding experiences than me.
3. I get worried when I find out my friends are having fun without me.
4. I get anxious when I don't know what my friends are up to.

415. As for social comparisons, research has shown that "upward" comparisons (comparing oneself to someone that one deems "superior") outnumber downward ones on social

media and that the effects of these comparisons is predominately negative.<sup>423</sup> Facebook’s own analysis reveals that social comparison feeds a negative feedback loop.



Document 135: META3047MDL-003-0000029, -0058

416. George Volichenko, a product engineer at Insta, in his deposition on December 6, 2024, even gives these a name, and a metric. He said “p-NAC is --so NAC stands for negative appearance comparison. P-NAC is basically a probability. So “p” is lowercase, which is often used in statistics as, like, probability. You know, you’ve heard of p-values also related to probability.”<sup>424</sup> Volichenko and his team proposed identifying high risk images for social comparisons (e.g. bikini model versus Christmas tree (*his words*)) and assigning a “probability” estimate to their likelihood of instigating negative comparisons so they could be down ranked and

<sup>423</sup> Midgley C, Thai S, Lockwood P, Kovacheff C, Page-Gould E. When every day is a high school reunion: Social media comparisons and self-esteem. *Journal of Personality and Social Psychology*, . 2021;121(2):285-307.

<sup>424</sup> George Volichenko Dep. Tr. at 40:18-24

less likely to appear in teens feeds. This would create a potential mechanism to mitigate the risk they posed to self-esteem. Volichenko asserts that when the mechanism was proposed to Adam Mosseri, head of Instagram, “he did not approve it.”<sup>425</sup>

417. Considerably fewer studies have examined anxiety symptoms than depression and fewer still an “anxiety diagnosis” as detailed above although many have examined both depressive and anxiety symptoms together as “internalizing” symptoms or overall “mental health.” Tang et al’s systematic review of longitudinal studies of SM and mental health in young people identified two studies that assessed social media use and subsequent psychological distress and both found significant correlations.<sup>426</sup> One of those, by Riehm et al, warrants a deeper dive as it was a large (6595 people) prospective study.<sup>427</sup> Their measure of “internalizing” symptoms included the following questions:

**Figure 44: Questions to Measure Whether a Person “Internalizes” Symptoms of Anxiety**

Internalizing Problems <sup>a</sup>	1. Feeling very trapped, lonely, sad, blue, depressed, or hopeless about the future?
	2. Sleep trouble, such as bad dreams, sleeping restlessly, or falling asleep during the day?
	3. Feeling very anxious, nervous, tense, scared, panicked, or like something bad was going to happen?
	4. Becoming very distressed and upset when something reminded you of the past?

418. They found that between 3 and 6 hours of social media use per day was associated with a 60% increased risk of internalizing problems even when adjusting for co-variates including

<sup>425</sup> George Volichenko Dep. Tr. at 44:9

<sup>426</sup> Tang S, Werner-Seidler A, Torok M, Mackinnon AJ, Christensen H. The relationship between screen time and mental health in young people: A systematic review of longitudinal studies. *Clin Psychol Rev.* Jun 2021;86:102021. doi:10.1016/j.cpr.2021.102021

<sup>427</sup> Riehm KE, Feder KA, Tormohlen KN, et al. Associations Between Time Spent Using Social Media and Internalizing and Externalizing Problems Among US Youth. *JAMA Psychiatry.* 2019;76(12):1266-1273. doi:10.1001/jamapsychiatry.2019.2325

baseline risk of internalizing problems. More than 6 hours per day was associated with a 78% increased risk. In addition, Bragheiri et al's quasi-experimental study of the Facebook rollout published after Tang's systematic review did specifically examine the effect size for having a diagnosis of an "anxiety disorder in the past year" and found a significant effect size of .07.<sup>428</sup> Finally, the Allcott et. al. experiment (summarized in the depression) did include "feeling anxious" as an outcome and found a significant treatment effect of .09 (larger than for depression) meaning that abstaining from Facebook for a period of one week resulted in a reduction in anxiety. In addition, the Davis SM reduction study in college students (detailed in the FoMo section) also found significant reduction in anxiety symptoms using the Generalized Anxiety Disorder 7 (GAD-7) of 2.35 points (Cohen's d .38).<sup>429</sup> Again, the experimental nature of the design of both of these studies allows for causal inferences to be made.

419. Meta conducted a "Teen Mental Health Deep Dive" in April 2020 that found the following:

---

<sup>428</sup> Tang S, Werner-Seidler A, Torok M, Mackinnon AJ, Christensen H. The relationship between screen time and mental health in young people: A systematic review of longitudinal studies. *Clin Psychol Rev*. Jun 2021;86:102021. doi:10.1016/j.cpr.2021.102021.

<sup>429</sup> Davis CG, Goldfield GS. Limiting social media use decreases depression, anxiety, and fear of missing out in youth with emotional distress: A randomized controlled trial. *Psychology of Popular Media*. 2025;14(1):1-11. doi:10.1037/ppm0000536.

## Teens blame Instagram for increases in the rates of anxiety and depression among teens

- This reaction was unprompted and consistent across all groups
- Constant comparison on Instagram is “the reason” why there are higher levels of anxiety and depression in young people
- Social comparison and perfectionism are nothing new, but young people are dealing with this on an unprecedented scale.
- The proliferation of new and different ways to compare themselves to others, combined with constant access to means that there is no way to escape social comparison on IG.
- For both boys and girls, this was called out as being the number one reason why IG is worse than other platforms for mental health. And, young people openly attribute their increased level of anxiety and depression to Instagram.

---

*“The reason why our generation is so messed up and has higher anxiety and depression than our parents is because we have to deal with social media. Everyone feels like they have to be perfect.”*  
- UK Female

*Document 136: META3047MDL-003-00109173, -9196*

420. I conclude to a reasonable degree of medical and scientific certainty that the evidence supports a causal relationship between social media use and anxiety symptoms. Furthermore, internal documents from Meta, Snap, Tik Tok and Google provide additional support that the use of their platforms and their features – including the algorithms they developed and deployed – increased anxiety in a portion of their users.<sup>430</sup> The actions they took, if any, were weighed against the impact they would have on their core metrics, and ultimately on their bottom line and were minimally effective by design.

### **E. Suicide, Suicidal Ideation, and Self-Harm**

421. Suicide and suicidal ideation are significant public health concerns that impact individuals across all demographics. Suicide, the act of intentionally ending one’s own life, is often

---

<sup>430</sup> See e.g., TIKTOK3047MDL-004-00137151, -7152; GOOG-3047MDL-00874191 at Slide 29; SNAP0933703, -3724

preceded by intense emotional pain, feelings of hopelessness, and mental health struggles, such as depression, anxiety, or trauma. According to the Centers for Disease Control, suicide is a leading cause of death among adolescents in the United States. In 2021, it was the second leading cause of death for individuals aged 10–14 and 20–24, and the third leading cause for those aged 15–19.

422. Suicidal ideation refers to thinking about, planning, or considering suicide, ranging from fleeting thoughts to detailed planning. These experiences can stem from a combination of factors, including biological, psychological, social, and environmental influences. Suicidal ideation as well as self-injurious actions (self-harm) frequently precedes suicide attempts and is a significant risk factor for them.<sup>431</sup> Because suicide itself remains relatively rare, most studies focus on the elevated odds of self-harm and suicidal ideation occurring as a proxy for suicide risk.

423. We have already identified social media as a risk factor for depression, anxiety, eating disorders, body image, and sleep disturbances all of which are independent risk factors for suicide. Here we will focus on the evidence linking social media usage to suicidal ideation, self-harm, and suicide itself irrespective of mechanism.

424. A recent *Morbidity and Mortality Weekly Report* from the Department of Health and Human Services analyzed cross sectional Youth Risk Behavior Survey data of U.S. high school students.<sup>432</sup> They dichotomized self-reported social media use as “frequent” if the respondent

---

<sup>431</sup> Franklin JC, Ribeiro JD, Fox KR, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol Bull.* Feb 2017;143(2):187-232.

doi:10.1037/bul0000084; Ribeiro JD, Franklin JC, Fox KR, et al. Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: a meta-analysis of longitudinal studies. *Psychol Med.* Jan 2016;46(2):225-36. doi:10.1017/S0033291715001804.

<sup>432</sup> Young E, McCain J, Mercado M, al. e. Frequent Social Media Use and Experiences with Bullying Victimization, Persistent Feelings of Sadness or Hopelessness, and Suicide Risk Among High School Students — Youth Risk Behavior Survey, United States, 2023. *MMWR Suppl* 2024;. 2024;73(4):23-30.

replied that they used them “several times a day” which 77% reported that they did. Their results are summarized below:

**Figure 45: Unadjusted Prevalence Estimates and Adjusted Prevalence Ratios for Bullying, Mental Health, and Suicide Risk Among High School Students**

Health behavior and experience (past 12 months)	Frequent social media use		PR <sup>†</sup> (95% CI)	aPR <sup>§</sup> (95% CI)
	Yes	No		
	% (95%CI)	% (95%CI)		
<b>Bullying victimization</b>				
Bullied at school	19.9 (18.3–21.4)	19.0 (12.9–27.1)	1.05 (0.72–1.52)	1.31 (1.12–1.53) <sup>¶</sup>
Electronically bullied	17.0 (15.7–18.4)	15.9 (8.1–28.7)	1.07 (0.57–2.02)	1.54 (1.26–1.88) <sup>¶</sup>
<b>Mental health</b>				
Persistent feelings of sadness or hopelessness	42.6 (40.4–44.8)	31.9 (25.3–39.3)	1.33 (1.07–1.65) <sup>¶</sup>	1.35 (1.23–1.47) <sup>¶</sup>
<b>Suicide risk</b>				
Seriously considered attempting suicide	20.2 (18.8–21.8)	18.7 (12.8–26.6)	1.08 (0.75–1.55)	1.21 (1.06–1.37) <sup>¶</sup>
Made a suicide plan	16.6 (15.1–18.2)	17.5 (10.3–27.9)	0.95 (0.58–1.55)	1.16 (1.00–1.35) <sup>¶</sup>
Attempted suicide	9.5 (8.4–10.8)	9.5 (6.6–13.5)	1.00 (0.70–1.43)	1.11 (0.89–1.39)

**Abbreviations:** aPR = adjusted prevalence ratio; PR = prevalence ratio.  
<sup>\*</sup> N = 20,103 respondents. The total number of students answering each question varied. Data might be missing because 1) the question did not appear in that student's questionnaire, 2) the student did not answer the question, or 3) the response was set to missing because of an out-of-range response or logical inconsistency. Percentages in each category are calculated on the known data. A total of 15,203 students responded to the social media item.  
<sup>†</sup> Logistic regression models estimated health behaviors and experiences between those who did and did not use social media at least several times a day.  
<sup>§</sup> Adjusted for age, race and ethnicity, sex, and sexual identity estimated health behaviors and experiences behaviors between those who did and did not use social media at least several times a day.  
<sup>¶</sup> Estimates were considered statistically significant if the 95% CIs did not include 1.0. Certain statistically significant aPRs have 95% CIs that include 1.0 because of rounding.

425. As can be seen in the above figure, “frequent” social media use was associated with a 35% increased risk of “persistent feeling of sadness,” a 21% increased risk of “seriously considering attempting suicide,” and a 16% increased risk of “making a suicide plan.” All of those associations were “statistically significant.” The authors acknowledge that these associations are cross-sectional and therefore causality cannot be established. It could credibly be asserted that the causality is reversed and that “persistent feelings of sadness” beget social media usage for example. Or more likely, that there is a dyadic, mutually reinforcing relationship where searching for self-harm videos (because one is considering it) leads to content that induces viewing more of it and increasing the likelihood of doing it.

426. There are at least two mechanisms by which social media usage can spur suicidal thoughts and actions: emulation and increased despair (they are not mutually exclusive). Suicide

as a contagious phenomenon has been reported for years. Niederkrotenthaler and colleagues meta-analyzed studies that examined the risk of suicide after a celebrity suicide was reported in the media.<sup>433</sup> They found that the risk of suicide in the intervening 1-8 days was increased by 13% and the risk of suicide using the method deployed by the celebrity was increased by 30%.<sup>434</sup> This provides strong evidence for “emulation” effects both because of the act itself as well as the methods chosen, but further, social media is assuredly one of the mechanisms by which the information was disseminated.

427. As for increased despair, the link between social media usage and depression was discussed in section X.D.ii but there are also pathways to despair via cybervictimization/cyberbullying and sextortion (discussed in later sections). A 2021 systematic review and meta-analysis by Nesi et al summarized 61 articles relating social media usage to self-injurious thoughts and behaviors (suicidal ideation and self-harm in the model).<sup>435</sup> Their results are summarized below:

---

<sup>433</sup> Niederkrotenthaler T, Braun M, Pirkis J, et al. Association between suicide reporting in the media and suicide: systematic review and meta-analysis. *BMJ*. 2020;368:m575. doi:10.1136/bmj.m575.

<sup>434</sup> Niederkrotenthaler T, Braun M, Pirkis J, et al. Association between suicide reporting in the media and suicide: systematic review and meta-analysis. *BMJ*. 2020;368:m575. doi:10.1136/bmj.m575.

<sup>435</sup> Nesi J, Burke TA, Bettis AH, et al. Social media use and self-injurious thoughts and behaviors: A systematic review and meta-analysis. *Clinical Psychology Review*. 2021/07/01/ 2021;87:102038. doi:<https://doi.org/10.1016/j.cpr.2021.102038>.

**Figure 46: Associations Between Social Media Use Variables and Self-Injurious Thoughts and Behaviors**

Associations between social media use variables and self-injurious thoughts and behaviors.					
	<i>k</i>	<i>N</i>	Effect size analyses		
			OR	95% CI	<i>p</i>
Cybervictimization					
Suicidal Ideation	45	135,424	2.93	2.43–3.54	<.001
Suicide Plans	10	40,760	3.07	2.18–4.34	<.001
Suicide Attempts	25	106,417	3.38	2.59–4.41	<.001
NSSI	3	532	4.36	2.32–8.20	<.001
Cyberbullying perpetration					
Suicidal Ideation	5	2444	1.89	1.54–2.32	<.001
Suicide Plans	1	650	1.87	1.41–2.48	<.001
Suicide Attempts	3	1890	1.65	1.25–2.18	<.001
SITB-related social media use					
Suicidal ideation	5	3871	2.79	1.85–4.21	<.001
Suicide plans	3	10,980	3.78	1.90–7.55	<.001
Suicide attempts	5	11,735	3.94	2.20–7.07	<.001
NSSI	2	245	2.98	1.46–6.11	.003
Frequency of social media use					
Suicidal ideation	6	2974	1.45	0.95–2.23	.089
Suicide plans	2	391	1.47	0.33–6.43	.612
VNSSI	3	570	2.03	0.79–5.21	.143
Problematic social media use					
Suicidal ideation	4	21,391	2.81	1.72–4.59	<.001
Sexting					
Suicidal Ideation	2	586	2.37	0.98–5.73	.057
Suicide Attempts	1	11,707	4.24	3.13–5.44	<.001
NSSI	2	6103	3.07	2.53–3.74	<.001
Importance of social media					
Suicidal ideation	3	858	1.05	0.96–1.15	.291
Suicide plans	2	391	1.02	0.71–1.49	.902
NSSI	2	391	1.25	1.06–1.47	.007

*k* = number of unique effects; CI = confidence interval; NSSI = non-suicidal self-injury; SITB = self-injurious thoughts and behaviors.  
 Note that only outcomes for which at least one effect was identified are listed for each social media predictor. Effect size estimates where *k* < 3 should be considered unstable and interpreted with a degree of caution.

428. Cybervictimization, cyberbullying perpetration, self-injurious thoughts and behavior related social media, and problematic social media use were all associated with increased odds of self-injurious thoughts and behaviors with odds ratios between 1.7 and 3.9. The vast majority of the included studies are cross-sectional and while the association is plausibly causal

based on the theoretical mechanisms discussed above, reverse causality once again cannot be excluded when studies are merely correlational.

429. However, there have been a few, well done longitudinal studies that have evaluated the risks posed by social media sites with respect to suicidal ideation. Arendt and colleagues conducted a prospective panel survey of 729 U.S. adults ages 18-29 years.<sup>436</sup> At baseline, participants were asked the following question: *Please think about the social networking site Instagram: How often, if ever, have you seen a post on Instagram showing someone who intentionally harms him- or herself, for example, by cutting? Was it more than once, just once, or never?*

430. Overall, 43% reported having seen it once or more than once. The outcome of interest at wave 2 (1 month after baseline) was the answer to the question: *Since the first survey, have you ever engaged in self-harming behavior, such as cutting your wrists.* The results of their regression analysis are presented below:

---

<sup>436</sup> Arendt F, Scherr S, Romer D. Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults. *New Media & Society*. 2019/11/01 2019;21(11-12):2422-2442. doi:10.1177/1461444819850106

**Figure 47: Regression Analysis of Effects of Exposure to Self-Harm on Social Media**

Outcome = self-harming behavior	Cross-sectional evidence (predicting wave 1 outcome)				Panel evidence (predicting wave 2 outcome)			
	B	SE	OR	p	B	SE	OR	p
	Step 2 (effect of Instagram exposure) $\Delta\chi^2_{(2)} = 20.30, \Delta R^2 = .044, p < .001$				Step 2 (effect of Instagram exposure) $\Delta\chi^2_{(2)} = 28.10, \Delta R^2 = .064, p < .001$			
Age	0.02	0.03	1.02	.515	-0.08	0.04	0.93	.046
Gender D1 (male)	-1.57	0.78	0.21	.045	1.47	1.17	4.37	.209
Gender D2 (female)	-1.00	0.75	0.37	.179	1.48	1.14	4.40	.195
Education	-0.13	0.06	0.88	.035	-0.08	0.07	0.92	.281
Race	0.26	0.22	1.30	.236	-0.56	0.25	0.57	.028
Self-harm exposure via other sources	0.16	0.05	1.18	.003	0.02	0.07	1.02	.787
Self-harming behavior (wave 1)	—	—	—	—	1.64	0.24	5.17	<.001
Instagram exposure D1 (just once)	0.68	0.25	1.97	.007	1.58	0.30	4.87	<.001
Instagram exposure D2 (more than once)	1.05	0.24	2.87	<.001	0.79	0.32	2.20	.013

SE: standard error; OR: odds ratio;  $R^2$ : Nagelkerke's  $R^2$ ; D: dummy variable; cross-sectional evidence full model:  $\Delta\chi^2_{(8)} = 55.52, R^2 = .124, p < .001$ ; panel evidence full model:  $\Delta\chi^2_{(9)} = 110.74, R^2 = .270, p < .001$ .  
This table only reports step 2 of the hierarchical model (i.e. full model). Both Instagram dummies were added in step 2. Therefore, the change in  $\chi^2$  of step 2 indicates whether Instagram exposure explains additional variance.

Focus on the wave 2 outcomes outlined in red which adjust for baseline characteristics and risk factors (wave 1) including exposure to others' sources of self-harm (e.g. newspapers, news reports etc.).

431. Exposure to self-harming behavior on Instagram was associated with an almost 5-fold increased odds of self-harm at 1 month follow-up. Further, the authors used the validated Eltz suicide risk prediction scale to assess suicidal risk.<sup>437</sup> Specifically, participants were asked to rate six items:

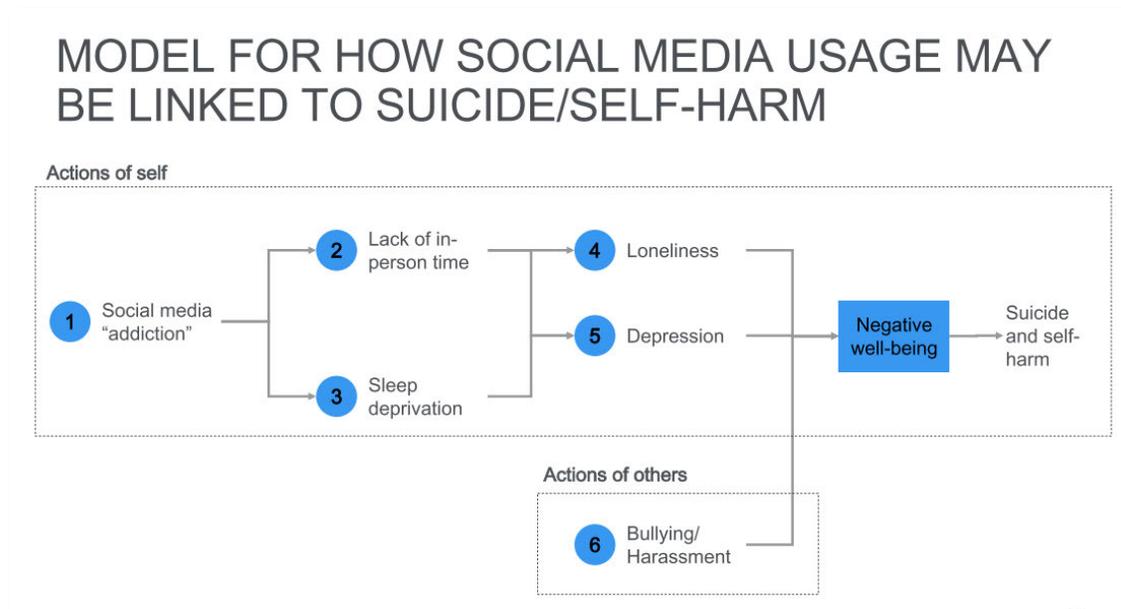
1. I think of suicide
2. I have thoughts about how to end my life

<sup>437</sup> Eltz M, Evans AS, Celio M, et al. Suicide probability scale and its utility with adolescent psychiatric patients. *Child Psychiatry Hum Dev.* Jun 2007;38(1):17-29. doi:10.1007/s10578-006-0040-7.

3. I feel it would be less painful to die than to keep living, given the way things are
4. I feel the world is not worth continuing to live in
5. I feel people would be better off if I were dead
6. In order to punish others, I think of suicide.

Again, adjusting for baseline risk factors, exposure to self-harm images on Instagram was associated with a significantly elevated score on the suicide risk score ( $p < .006$ ).

432. Meta has internally acknowledged that addictive use of Instagram plays a role in suicide and self-injury as is shown in an internal memo:

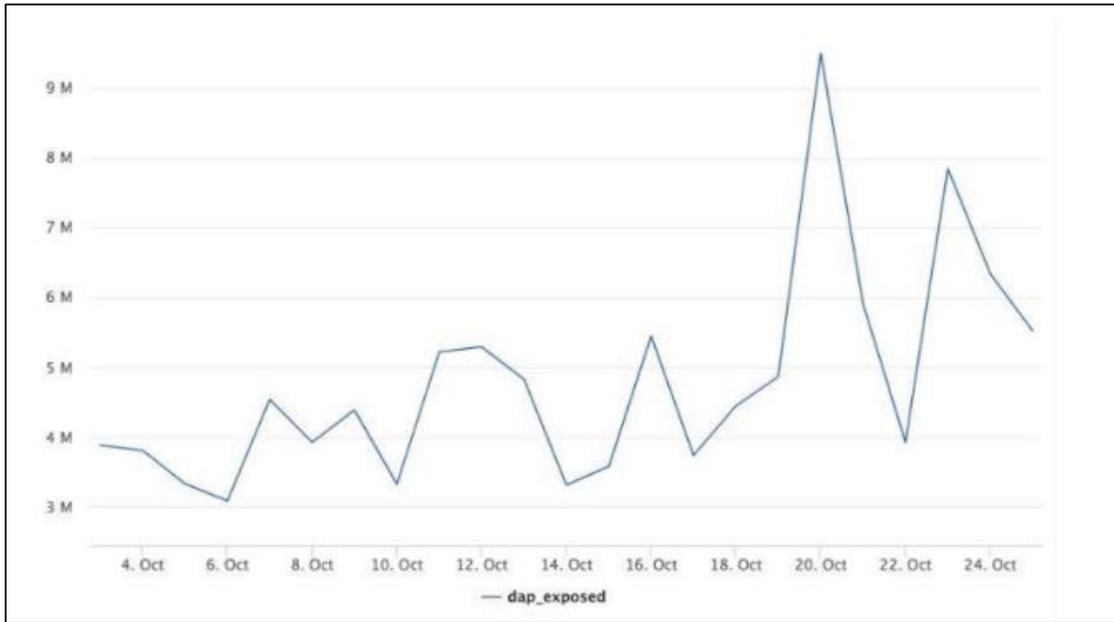


*Document 137: META3047MDL-037-00068917 at Slide 30*

433. Instagram’s viral promotion of suicide and self-injury-oriented content (including challenges glorifying and encouraging these harms) is a mediator along this causal pathway. Meta internal documents make clear (a) that such content exists on the platform, (b) that Instagram’s algorithm promotes it to teenagers (at higher rates than to non-teens), in part through viral

(contagious) algorithmic spreading of such content, (c) that teen viewing of this content sets up a cycle where teens are more likely to post such content, and (d) that the platform's exposing teens to such content is harmful. As such, this is not simply a question of "bad" content existing on the platform, it is a question of the platform driving that content to a specific set of vulnerable users knowing the attendant risks. Furthermore, Meta created this "causal" pathway. They subsequently acknowledged their role and purportedly intended to mitigate it. The problem is, in the end, they did nothing about it.

434. Let us examine in more depth the causal role that Meta may play in promoting SSI. First, there is the presence of the content itself. Meta's platform is replete with suicidal content with inadequate protections. An internal report notes that "5.1 million daily users are exposed to suicide or self-harm admission or promotion content." **PRIV** Furthermore, the report includes the graph for the three weeks in October 2023 and states that "the spikes are from very viral content (meme-like in quality) that have SSI content."



*Document 138: META3047MDL-040-00056476, -6481*

435. Second, there is the fact that the platform pushes the content. An internal audit revealed that, “after viewing or engaging with inappropriate content, it starts to show up more and more” and that this phenomenon applies to content that is “potentially inappropriate for early teens.” In other words, the algorithms preferentially foist content on the most vulnerable teenagers.

While 0.42% of content is policy violating, we have reason to believe that ~12% of content\* is potentially inappropriate for early teens.

And after viewing or engaging with inappropriate content, it starts to show up more and more on recommendation surfaces throughout the app.



\*We conducted a [labeling exercise](#) where we manually labeled a random sample of media that teens see. The approach we took is subjective, but provides directional evidence that there is far more age-inappropriate content on our recommendation surfaces than our prevalence numbers may suggest. This [link](#) provides a sample of media that teens saw on Explore.

We look an early teen lens here. It is likely that as teens get older, there is far less content that we may believe is inappropriate for them.

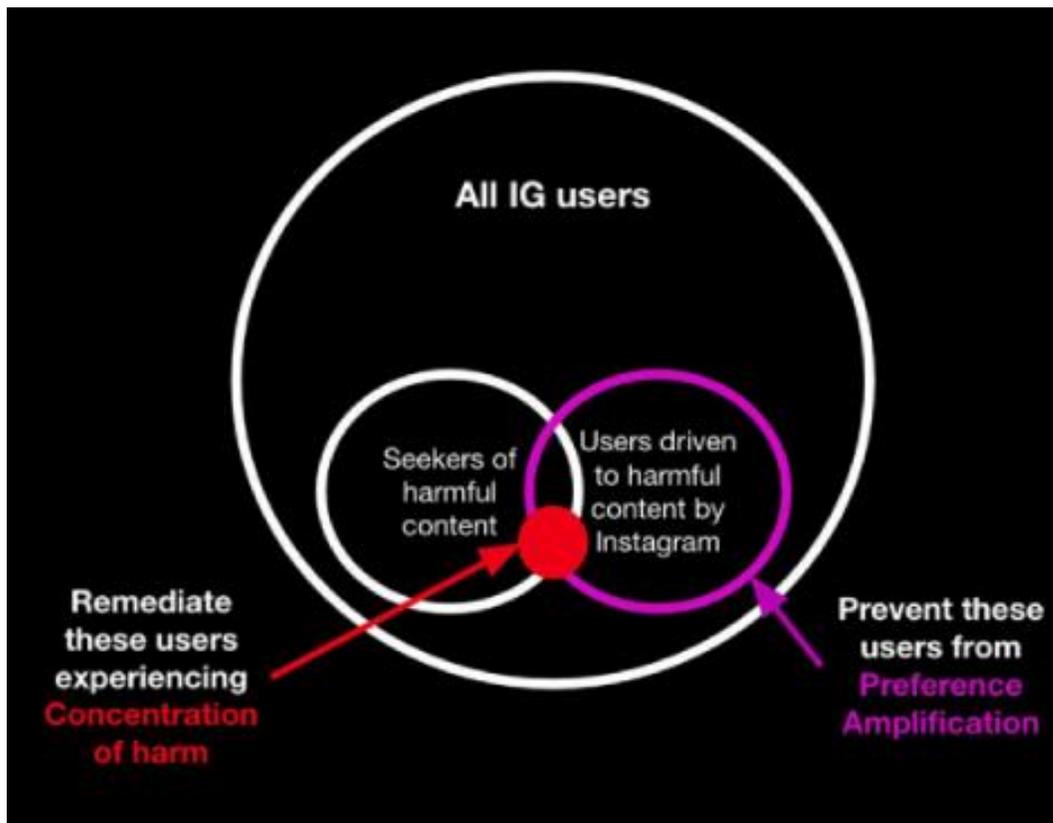
*Document 139: META3047MDL-035-00002761 at Slide 46*

436. As Miki Rothschild explained in an internal email from May 28, 2020, viewing “unsafe” content leads to more such content:

However, I think there might be a simpler and more effective way to tackle the rabbit holing on safety problems (vs the time spent one). Today the primary way we combat rabbit holing you into bad stuff is by trying to detect and filter out bad stuff. This means we treat every recommendation the same even though in practice once you engage with something unsafe we’re so much more likely to show you more unsafe recommendations, even though we should actually be better at detecting clusters of recommendations than individual ones. So theoretically we could reduce the bar for what we consider unsafe on subsequent recommendations to avoid rabbit holing on those. Let me check with the team to see if there’s more targeted approach we can consider...

*Document 140: META3047MDL-003-00064697, -4697*

437. Further, a 2022 “Concentration of Harm” presentation at Instagram refers to “preference amplification” for a subset of Instagram users who are not “seekers” of harmful content but who are “driven” to it by Instagram:



*Document 141: META3047MDL-054-00000061, -0063*

According to the Oxford English dictionary, the second definition of the verb “drive,” after the one pertaining to operating a vehicle, is “propel or carry along by force in a specified direction.” That does seem like an apt description of how the algorithms work and by definition assigns culpability to them as “propellers.”

438. Yet another Meta internal memo explained the “significant risk of contagion” of SSI related content on Instagram:

- **Instagram is a popular platform for communicating about and finding support for suicide and self-injury (SSI) especially among teens.** SSI involves 1/4th as many reporters as B+H and has 1/5th as many content producers as B+H. But, there is a significant risk of contagion, especially during times of heightened awareness (e.g., high-profile suicide). Indeed, the odds of posting SSI content on Instagram are 7x higher for people who see a suicide admission than for people who do not.

Document 142: META3047MDL-031-00048769, -8769

In fact, viral SSI challenges are noted to be an increasing phenomenon in a 2021 memo from Instagram.<sup>438</sup>

We have noticed an increasing rate of SSI viral challenges on the platform.

- 2016 - 2020 H1: 2 challenges
  - **Blue Whale:** A 'group administrator' assigns 'daily tasks' to members, which they have to complete for 50 days e.g. self-harming, watching horror movies and waking up at unusual hours. These gradually get more extreme; on the 50th day, the controlling manipulators behind the game reportedly instruct the members to commit suicide
  - **Momo:** You have to do whatever "Momo" — or the person sending you the Momo image — says, or else Momo will come for you and your family and curse you. The final step in the process is killing yourself and filming it
- 2020 H2: 3 challenges
  - **Dog Face** ([ HYPERLINK "https://www.internalfb.com/intern/tasks/?t=76771922" \h ]): Accounts encouraged others to engage in "challenges," which solicited photo proof of inflicted self harm, otherwise threatening to reveal users' private information, or threatening to kill users and/or their family members
  - **Sinner** ([ HYPERLINK "https://www.internalfb.com/intern/tasks/?t=70769369" \h ], [ HYPERLINK "https://www.internalfb.com/intern/tasks/?t=70697558" \h ]): Variant on Blue Whale Challenge; a stranger messages a person on Instagram, giving them "challenges" or "games" to participate in, collecting personal & sensitive information from them, and ultimately threatening with releasing the sensitive information if they don't comply with the challenges
- 2021 H1 to date:
  - **White Rat** ([ HYPERLINK "https://www.internalfb.com/intern/tasks/?t=81915313" \h ]): An iteration of the Blue Whale Challenge

<sup>438</sup> META3047MDL-111-00204020, -4020 (“We have noticed an increasing rate of SSI viral challenges on the platform.”)

- o [March 3](https://www.internalfb.com/intern/tasks/?t=85816542): A trend/viral challenge on TikTok encouraging teens to commit suicide on 3 March, that we were concerned might spill over to Instagram
- o [Apostle of God](https://www.internalfb.com/intern/tasks/?t=83873869): An anonymous figure on Instagram making people draw a "pact of fidelity", followed by numerous tests of loyalty, increasingly pressing and dangerous requests. A 15 year old girl in Italy nearly threw herself off the roof of a building to acquiesce to the Apostle's demands
- o [Resurgence of Dog Face](https://www.internalfb.com/intern/tasks/?t=76819409)

*Document 143: META3047MDL-111-00204020, -4020-21 (emphasis in original)*

439. As indicated in this document, a 15-year-old girl in Italy “nearly threw herself off the roof of a building” because of the challenge.<sup>439</sup> Given this example, it is particularly astonishing that the memo goes on to say:

- We've historically taken a reactive approach to dealing with viral events (i.e. we deal with it once it's become so viral that media outlets catch wind of it). It's very difficult to predict/surface trends before they become viral, so it's tough to proactively ID them.

*Document 144: META3047MDL-111-00204020, -4022*

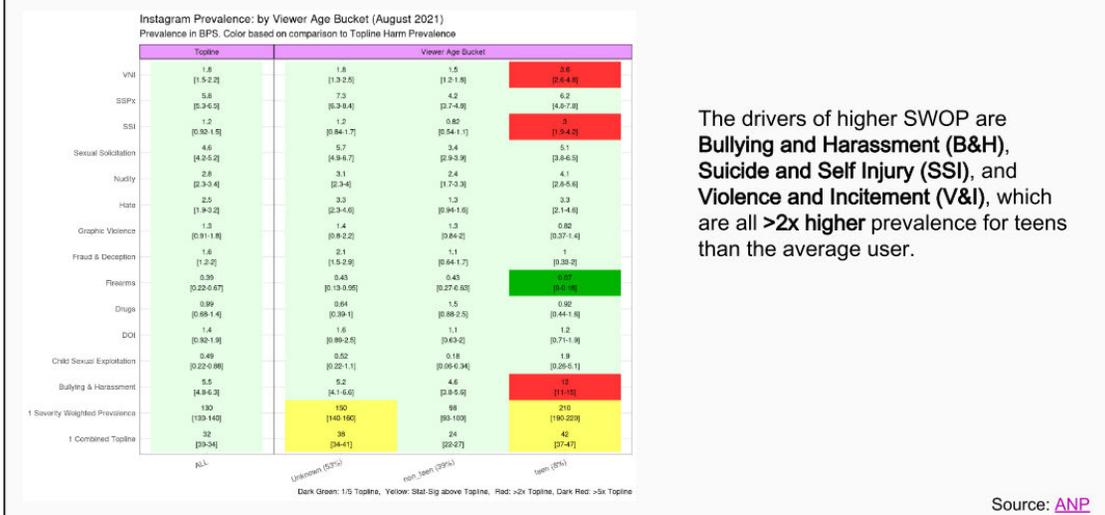
Meta well understood the potentially fatal consequences of its “reactive approach” (dealing with viral suicide challenges after “media outlets catch wind of it”).

440. Third, the promotion of suicide-oriented content specifically to teens. Instagram’s own documents reveal that teens see more than two times the prevalence of suicide and self-injury compared to non-teens:

---

<sup>439</sup> META3047MDL-111-00204020, -4021

Teens see >2x prevalence for B&H, VNI, and SSI.



The drivers of higher SWOP are **Bullying and Harassment (B&H)**, **Suicide and Self Injury (SSI)**, and **Violence and Incitement (V&I)**, which are all >2x higher prevalence for teens than the average user.

Document 145: META3047MDL-035-00002761 at Slide 44

I am not sure what “prevalence” means in this context. One meaning is the percent of users who see each of these types of content in a given time period. If that is the case, it is not only concerning that teens see more than twice the amount of Suicide and Self Injury Content but that 3% do at all.

441. Further, Meta knew through internal research that viewing of SSI is significantly associated with posting it (see below):

Logistic Regression: Actioned SSI Post	Odds ratio (unadjusted model)	Odds ratio (adjusted for confounders)
Suicide admission	15.72 (12.05, 20.49)	7.16 (5.42, 9.46)
Suicide promotion	1.64 (1.09, 2.46)	1.29 (0.86, 1.93)
Self-harm admission	3.92 (2.86, 5.38)	2.88 (2.10, 3.96)
Self-harm promotion	1.51 (0.94, 2.44)	1.30 (0.80, 2.11)

(Source: People Exposed to Suicide / Self Injury Content on Instagram)

Document 146: META3047MDL-014-00298174, -8196

442. Finally, there is Meta’s awareness of the harm from all this. Meta conducted an internal study where they created suicide/self-injury “personas” and then fielded test accounts for them.

## SSI Personas Test



- We studied SSI personas ongoing research done by Hitomi Hayashi-Branson, Product Researcher, CI
- We created test accounts and replicated the online behavior of the Personas
- We observed content surfacing to the Personas and IG intervention
- Hypothesis: An SSI persona through their natural online behavior will encounter more harmful suicide and self-injury content (through explore, related, follower suggestions, etc)

16

CONFIDENTIAL (COMPETITOR)      META3047MDL\_RVOL003      META3047MDL-003-00068878

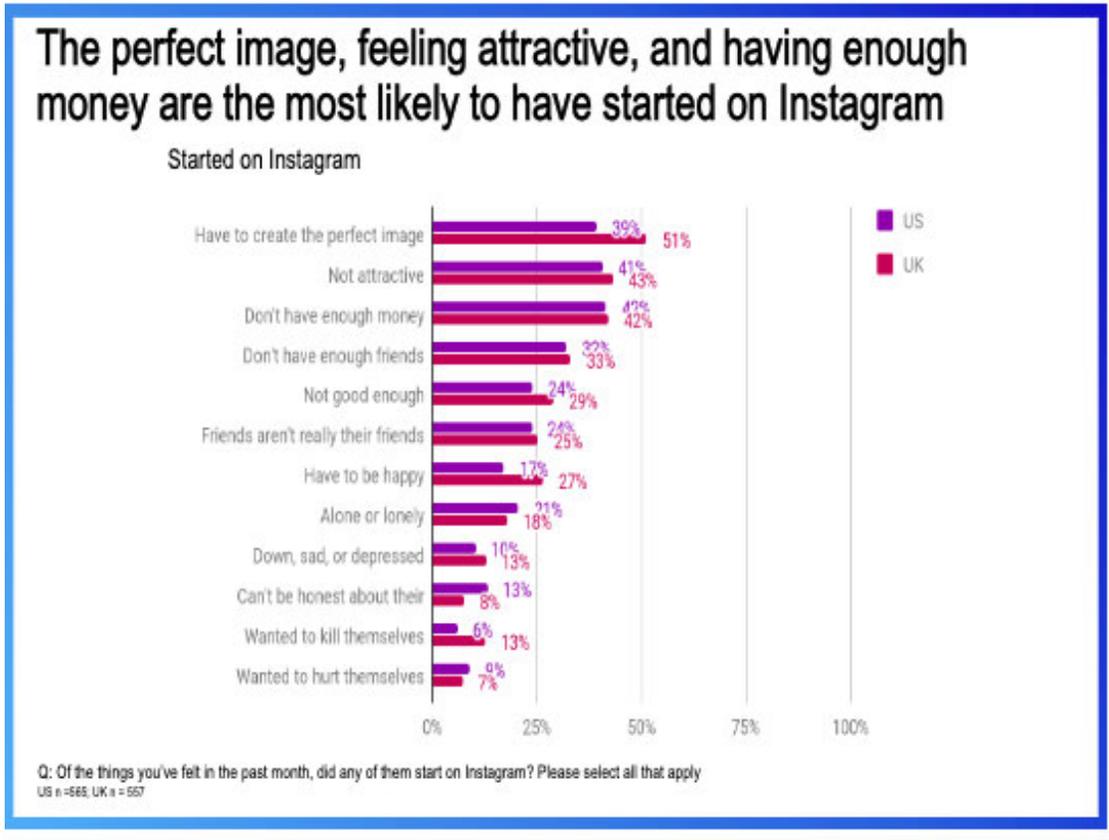
*Document 147: META3047MDL-003-00043617, -3632*

This test revealed that Instagram Explore “serves content based on user’s distressed online behavior”<sup>440</sup> and that numerous hashtags leading to suicide-promoting content were allowed, such as #suicidal, #killmyselfnow, and #suicidalthoughts. (The misspellings are intentional.)

443. In 2019, Meta did its own survey of 2503 teen users in the US and UK and asked the question: “*Of the things you’ve felt in the past month, did any of them start on Instagram? Check all that apply*”:

---

<sup>440</sup> META3047MDL-003-00043617, -3637



Document 148: META3047MDL-035-00002796 at Slide 16

The slide and its attendant text “calls out” the top three responses which include image, attractiveness and money, but does neither makes specific mention of the fact that 6 and 13% of US and UK teens respectively wanted to kill themselves and 9 and 7% wanted to hurt themselves and those thoughts “started on Instagram”<sup>441</sup>

444. Those statistics are tragically illustrated by the story of 14-year-old Molly Russell, a British girl who while battling depression, was inundated with graphic self-harm images, videos, and messages on social media (including Instagram) and took her life in 2017. After a five-year

<sup>441</sup> META3047MDL-035-00002796 at Slide 16

investigation, including a proceeding at which Meta witnesses were summoned to testify, the coroner concluded:

Molly had become depressed, a common condition affecting children of this age. This then worsened into a depressive illness. Molly subscribed to a number of online sites. At the time that these sites were viewed by Molly some of these sites were not safe as they allowed access to adult content that should not have been available for a 14-year-old child to see. The way that the platforms operated meant that Molly had access to images, video clips and text concerning or concerned with self-harm, suicide or that were otherwise negative or depressing in nature. The platform operated in such a way using algorithms as to result in some circumstances of binge periods of images, video clips and text, some of which were selected and provided without Molly requesting them.<sup>442</sup>

In 2019, a full two years after Molly's tragic death, Instagram reported that "All graphic images of self-harm will be removed."<sup>443</sup>

445. Despite that claim, limitations in Instagram's ability or willingness to limit access to SSI content are acknowledged in a Sept 2020 document:

- **IG:**
  - lack classifiers to detect and take action on problematic aggregated content on surfaces like Explore, hashtag pages, Reels
  - spaces like IGD being used to encourage or promote suicide e.g., group suicides in Norway that occurred last year and we don't scan there

*Document 149: META3047MDL-004-00027423, -7437*

Further, when Ms. Jayakumar was asked at her deposition (in 2025), "Does Instagram's algorithm sometimes push content that promotes suicide?" she responded, "The algorithm does sometimes push content that, in aggregate, could be seen as encouraging suicide but not at the individual level."<sup>444</sup>

---

<sup>442</sup> Vaishnavi Jayakumar Dep. Exhibit 20 at p. 2

<sup>443</sup> See Vaishnavi Jayakumar Dep. Exhibit 21

<sup>444</sup> Vaishnavi Jayakumar Dep. Tr. at 221:9-11

446. In March of 2020 Instagram held focus groups with both news-informed parents and young people (18-24) about what they viewed Instagram’s responsibility to be related to SSI content. ██████████ (IG Program Policy Manager) stated that the “biggest take away” from the groups was it “absolutely confounded them that we could be working on an issue so long, yet we are where we are. Rather than it seeming responsible, they took it as ‘so you've known about this all this time and you still haven't fixed it.’”<sup>445</sup> Reflecting on those same focus groups during her deposition, Jayakumar recalls:

This session was specifically about the messaging around SSI and the types of messages or narratives that parents and young people would find compelling and make them feel more favorable about Instagram. It wasn't about specific product interventions or policy interventions that we could take.<sup>446</sup>

These focus groups “focused” if you will on Instagram’s public messaging about SSI and were convened 3 years after Molly’s death and one year after Mosseri’s promise to take action.

447. Ultimately, Meta did create a “strike” policy for posting SSI promotional materials (see below):

```
Vaishnavi Jayakumar (5/20/2020 17:42:57 PDT):
>Hi - bumping this! To narrow it down a little more, trying to confirm our SSI promotion strikes policy.
My understanding is:
>
>IG has strikes for SSI promotion (both suicide and self-injury). The strike logic is:
>
>2 strikes => 1 day of blocking live videos
>3 strikes => 3 days of blocking live videos
>4 strikes => 7 days of blocking live videos
>5, 6, 7, 8 or 9 strikes => 30 days of blocking live videos
>>20 strikes in 90 days => Disable account
```

*Document 150: META3047MDL-020-00270857, -0857*

By most measures, the policy could be deemed “lenient” or “forgiving.” It takes several strikes to be blocked for any reasonable period of time. In the meantime, the multiple contagions have been released multiple times. And rather than just blocking live videos, why was a policy of parental

---

<sup>445</sup> META3047MDL-040-00544758, -4758

<sup>446</sup> Vaishnavi Jayakumar Dep. Tr. at 227:13-17

notification not created or some other similar mechanism to ensure that frequent posters were given help? This is a good example of the general and pervasive phenomenon Dr. Lee described in her deposition: “I think that the integrity guardrails that they put in place were not sufficient to seriously consider the integrity impacts and that there were other alternative strategies that could have been used.”<sup>447</sup>

448. What’s more, accounts promoting suicide and self-injury remained searchable. As ██████████ (Policy Communications Manager) stated in an email September 11, 2020, “We currently block inherently violating hashtags so that no results appear when you search for them. These include things like #proana and #thinspo. BUT — we don’t block any results for accounts. So when you search for these terms, there are no results under the ‘hashtag’ tab, but there are endless results under the ‘Top Accounts’ and ‘Accounts’ tab, and almost all are violating.”<sup>448</sup>

449. Importantly, and as confirmed in Jayakumar’s deposition, despite these critical gaps, there were no substantive changes in Meta’s suicide and self-injury posting policy between 2020 and 2025.<sup>449</sup> That policy has consistently stated, “we do not allow people to intentionally or unintentionally celebrate or promote suicide, self-injury or eating disorders” and “[w]e remove any content that encourages suicide, self-injury or eating disorders,” even though it is apparent (and Jayakumar confirmed) these statements are inaccurate—as she acknowledged, “there’s a disconnect between the suicide and self-injury policy and the aggregated content that actually gets recommended by Instagram’s algorithm.”<sup>450</sup> Put another way, there was no real way for any parent

---

<sup>447</sup> Alison Lee Dep. Tr. at 53:17-21

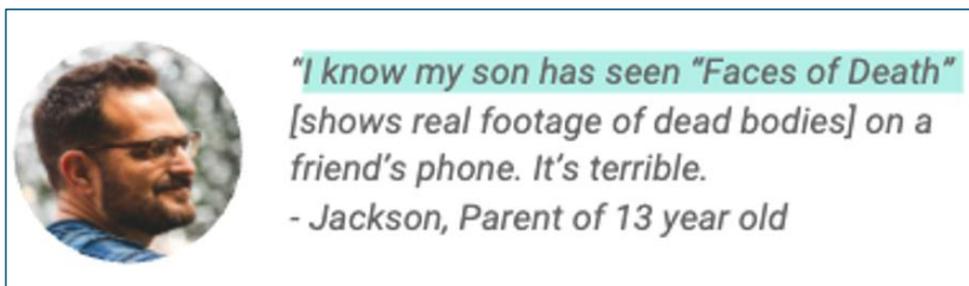
<sup>448</sup> META3047MDL-031-00246746, -6761

<sup>449</sup> Vaishnavi Jayakumar Dep. Exhibits 22 & 23; *See also* Vaishnavi Jayakumar Dep. Tr. at 217:1-21

<sup>450</sup> Vaishnavi Jayakumar Dep. Tr. at 222:6-12

to understand the risks that suicide and self-injury promoting content would be pushed to their kids—Meta’s stated policy indicated that it would *not*.

450. For its part, YouTube labels content as “gray” when it falls into a questionable area that although it does not technically violate its guidelines could still be problematic for some people. Their internal analysis reveals that in spite of “high parent and expert concern, <1% of watch time is gray content WT on YT.”<sup>451</sup> It is unclear what the denominator for “watch time” is. Total time everyone in the world watches YouTube? Total time teens watch? It likely is not total time teens at risk for suicide watch because aggregation algorithms will concentrate content on select individuals. For this reason, using a denominator of total watch time is not appropriate. A YouTube report has the following insert:



*Document 151: GOOG-3047MDL-00236723 at Slide 12 (emphasis in original)*

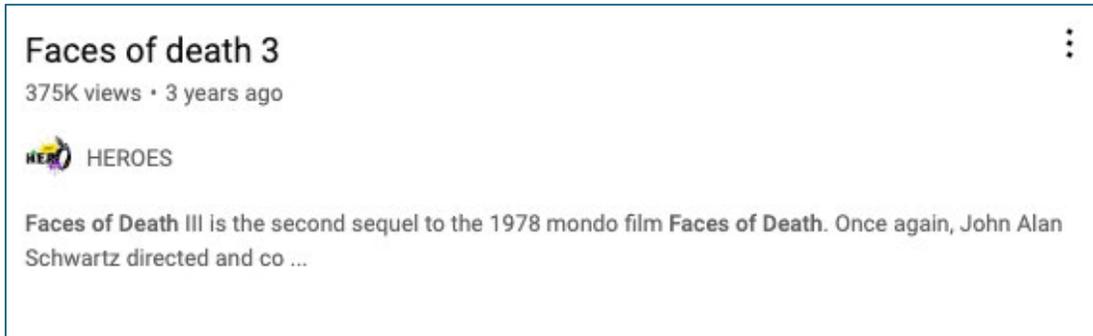
451. Deemed “The Ultimate Taboo” by *Entertainment Weekly*, *Faces of Death*, according to the box description is “Possibly one of the most talked about series of all-time [and] examines the many guises of death in the extreme close-up. Sure to shock, horrify and even repulse, these brutal films are not meant for the faint of heart.” Unrated in the US, banned in Finland and parts of Canada, and heavily age-restricted in many other countries, the movie is surely not appropriate for a 13-year-old regardless of their underlying mental health challenges. Having

---

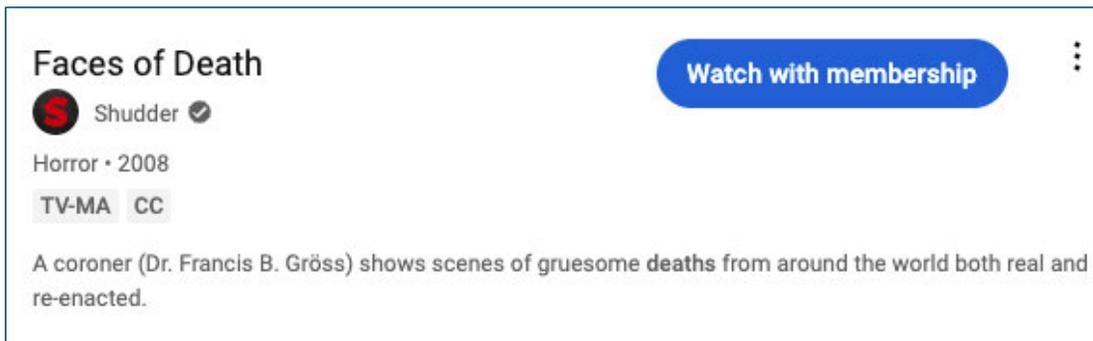
<sup>451</sup> GOOG-3047MDL-00236723 at Slide 12

viewed it, either because it was chanced upon or suggested via algorithmic curation, YouTube will feed a teen more such content. When I opened an incognito chrome page (so as to have no history) on February 15, 2025, and searched for “Faces of Death” on YouTube, I found the following:

**Figure 48: Sample YouTube Search Result for “Faces of Death”**



**Figure 49: Sample YouTube Search Result for “Faces of Death”**



452. The “Shudder” video link required a subscription and the “Heroes” link (Figure 45) asked for age verification. It is unclear when that gating was added or how easy it is to circumvent, but the link below also instantly popped up and autoplayed for me a “disturbing” clip from the movie which I did not watch in its entirety.

**Figure 50: Sample YouTube Search Result for “Faces of Death”**



453. What’s more, the same search also provided links to several other videos including:

**Figure 51: Sample YouTube Search Result for “Faces of Death”**



454. I am imagining that my feed would soon be populated with multiple such videos defaulted to autoplay. YouTube does appear to recognize that it is failing its teen users by providing suggestions that could be triggering, not reducing the concentration of views, and not allowing them to turn off SSH content.

## Current state and gaps

### (1) Explicit SSH intents:

- (a) Robust coverage of crisis resources for SSH intents globally, but without teen focus
- (b) Gap: No current mechanism to slow users and shift them off-topic

### (2) Implicit SSH intents

- (a) Gap: no current mechanism to reduce concentration of views

### (3) All SSH intents:

- (a) Gap: no current mechanism for expressing a preference to not see activating SSH content

Document 152: GOOG-3047MDL-00864164 at Slide 25

455. YouTube was aware that its algorithm was suggesting suicide and self-harm content alongside of “helpful” advice (see below).

#### **Suicide / Self-Harm / 'Nihilist' Content**

Features have been put in place by a number of platforms used by teenagers in particular e.g. Instagram, in relation to Suicide / Self-Harm / Nihilist content - not to remove all of the content - but to provide prompts and potential support for users searching for such content. This does not appear to be a feature on the YouTube platform. It appears that a large % of videos appearing through common searches in this area are offering prevention / helpful advice, but problematic content is easily discovered through related videos against these. E.g.,

- Search "how to kill yourself"
- Click into 2nd video "How to kill yourself. properly" 1.3M views. Appears to be a comedy/drama type video of a person trying unsuccessfully to kill themselves. Age gated. Comments troubling.
- Recommended videos: "Easy Steps to tie the perfect hangmans noose". Screen shot of recommendation (ranked 6th in list).
- @Dina, would you be ok to partner with us to understand the S&D implications in this space?

Separately, there is a genre of 'depressing edits' videos ([example](#)) - contains a number of triggering elements which could inspire those thinking of self harming, or those struggling with the issue. The comments section appears to contain a lot of these individuals. @Shadie, does a tear sheet already exist on this topic?

Document 153: GOOG-3047MDL-00488901, -8905

The document acknowledges that as of 2019 (at least) they were behind their competition in providing warnings about suicidal content.

456. TikTok, for its part, uses the following true case study as illustrative of how its algorithms can do harm to people with underlying mental health conditions.

A young person with undiagnosed obsessive compulsive disorder started watching videos of people doing checking behaviours (tidying things, straightening rows, making beds with neat edges etc). None of it was violative or problematic for the majority of viewers. The more he watched, the more he got served. He got caught in a loop and, following his diagnosis, had to come off TikTok. The young person wants to come back on to TikTok and asks for a reset button so he can leave his problematic viewing behind.

*Document 154: TIKTOK3047MDL-002-00075240, -5240*

Although they used obsessive compulsive disorder in this example, SSI would have functioned in much the same way.

457. Proactively identifying which videos will go viral may indeed be challenging, although for a company founded on predictive analytics it seems highly achievable. In fact, as part of their 2021 “wellness goals,” TikTok had the following:

- **【SSH】 Suicide Challenge**
  - Goal: Reduce the exposure of Suicide Challenge Video, Text promoting suicide in Video, Livestream, Comment
  - Gap: CV model detection of suicide characters and gestures; Proactive detection of suicide text variations

*Document 155: TIKTOK3047MDL-002-00094384, -4419*

458. Reducing exposure to “Suicide Challenge Videos” is a commendable goal, but anything short of 100% effectiveness poses real risks at scale to severely depressed teens. What’s more, elsewhere, they explicitly acknowledge the “contagion” effect such videos can have:

◦ Real contagion effect to some violence --> suicidal ideation and behavior and the way that it is talked about on the platform (chatsafe guidelines --> talk responsibly online about suicide/self-harm); to what extent can there be self curation of our own feed

*Document 156: TIKTOK3047MDL-004-00144763, -4763*

459. It is my opinion to a reasonable degree of medical and scientific certainty that a causal relationship exists between social media use and suicidal ideation, suicide, and self-harm. Furthermore, there is ample support in the internal documents and depositions that the Defendant platforms because of algorithms developed and deployed and addictive design increased the risk of suicidal ideation and self-harm. Based upon the internal documents, Defendant actions to mitigate these risks were minimal, and any action was weighed against the impact they would have on their core metrics.

#### **F. Attentional Capacity**

460. There are at least two mechanisms whereby digital media usage in general, and social media use in particular may diminish attentional capacity. One is direct, and the other indirect. The direct pathway is mediated through the “scan and shift” hypothesis which states that the frequent diversions afforded by digital media can condition the brain to expect high levels of input—to seek distraction—impeding its ability to focus on a specific task at length.<sup>452</sup> The second pathway is mediated through the displacement of other activities. Simply put, time spent on digital devices comes at the expense of other activities (e.g. physical activity and sleep) which can impact attentional capacity.

---

<sup>452</sup> Jensen PS, Mrazek D, Knapp PK, et al. Evolution and revolution in child psychiatry: ADHD as a disorder of adaptation. *J Am Acad Child Adolesc Psychiatry*. Dec 1997;36(12):1672-9; discussion 1679-81. doi:10.1097/00004583-199712000-00015v.

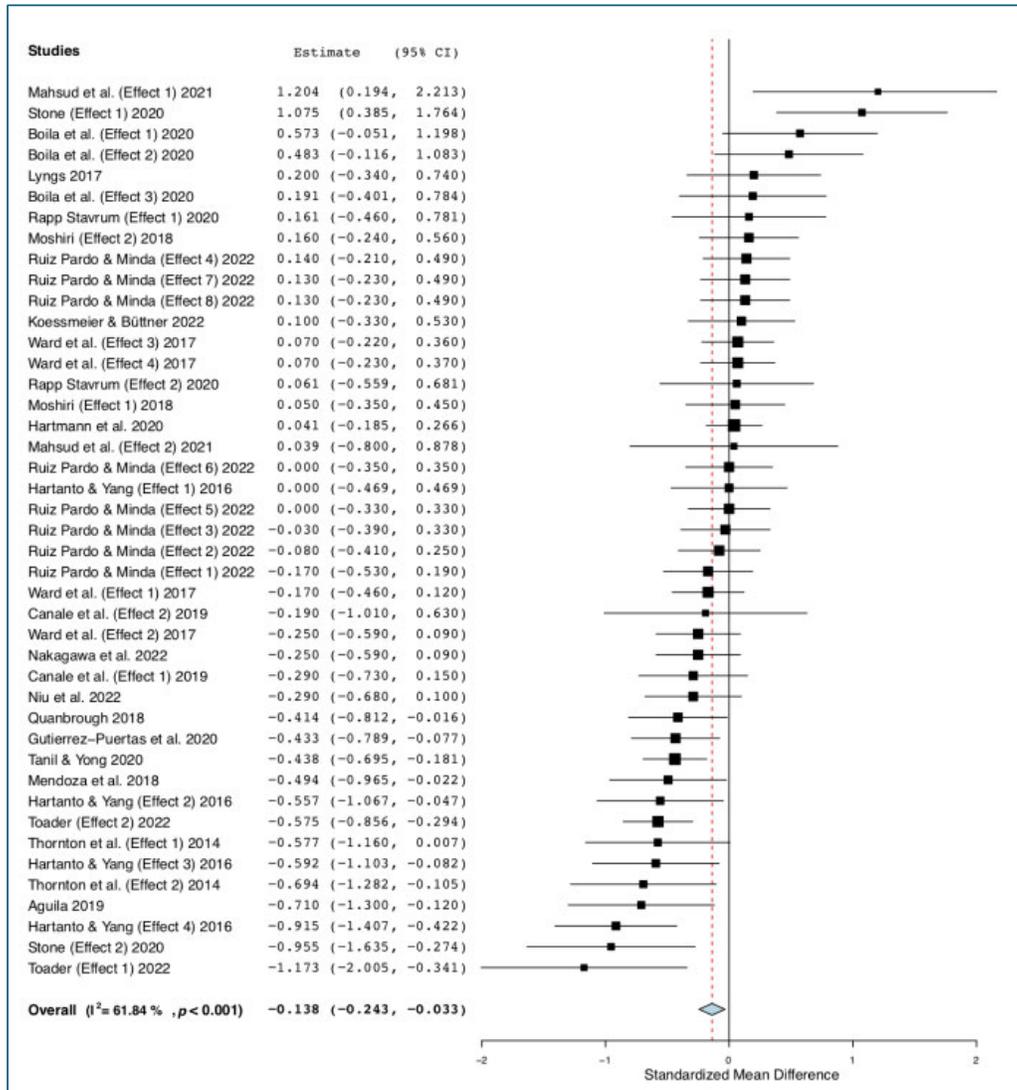
461. Some have thought that the mere presence of a phone is distracting and negatively impacts learning and attention, leading to what they call “brain drain.”<sup>453</sup> This phenomenon recapitulates what we discussed in the sleep section (Section X.C) where a metaanalysis found that the presence of a phone in the bedroom is associated with sleep disorders. Bottger and colleagues conducted a focused metaanalysis of “brain drain” by synthesizing 22 studies that assessed smartphone presence and cognitive function.<sup>454</sup> There was significant heterogeneity in the studies. Nevertheless, applying a random effects approach, the summary results are as follows:

---

<sup>453</sup> Ward A, Duke K, Gneezy A, Bos M, Drain B. The mere presence of smartphones reduces cognitive capacity. *J Assoc Consum Res.* 2017;2:140-54.

<sup>454</sup> Böttger T, Poschik M, Zierer K. Does the Brain Drain Effect Really Exist? A Meta-Analysis. *Behav Sci (Basel).* Sep 11 2023;13(9)doi:10.3390/bs13090751.

**Figure 52: Summary of Metanalyses of Relationship Between Smartphones and Cognitive Function<sup>455</sup>**



<sup>455</sup> Böttger T, Poschik M, Zierer K. Does the Brain Drain Effect Really Exist? A Meta-Analysis. *Behav Sci (Basel)*. Sep 11 2023;13(9)doi:10.3390/bs13090751.

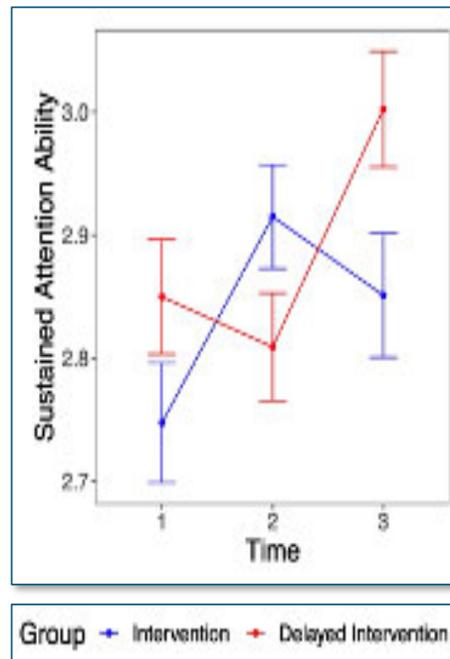
462. Overall, the presence of a smartphone was associated with a small but statistically significant negative effect (-0.138) on cognition. Sub-analyses reported that the effects were larger for memory and lower for attention and general cognitive performance.<sup>456</sup>

463. The Castelo et. al. study of making smartphones “dumb” detailed in Section X.D.i. also assessed attentional capacity using both self-reported attention and the “gold standard” approach of continuous performance task (CPT). Briefly, the CPT measures how well people can focus on a tedious task over time by having them stare at a blank screen and push buttons when prompted to do so based on the appearance of a specific symbol. It is based on a procedure used to screen cadets for service as radar operators during WWII, a task for which focused attention was of paramount importance to alert commanders to incoming aircraft. The results are presented below.

---

<sup>456</sup> Böttger T, Poschik M, Zierer K. Does the Brain Drain Effect Really Exist? A Meta-Analysis. *Behav Sci (Basel)*. Sep 11 2023;13(9)doi:10.3390/bs13090751.

**Figure 53: Sustained Attention Ability<sup>457</sup>**



The effect size was .24 ( $p < .008$ ). To put that into perspective, the authors report that it is the same magnitude of a 10-year age related decline in cognition and about  $\frac{1}{4}$  of the magnitude of difference between healthy adults and those with ADHD.

464. Ra and colleagues conducted a large longitudinal study.<sup>458</sup> His team surveyed 2587 15–16-year-old adolescents in 10 Los Angeles high schools without ADHD symptoms at baseline over the course of 2 years. High-frequency digital media use (defined as past-week use of 14 different media activities, such as checking social media, liking or commenting on others’ posts, online browsing, or streaming videos) many times a day was associated with an 11% increased

<sup>457</sup> Castelo N, Kushlev K, Ward AF, Esterman M, Reiner PB. Blocking mobile internet on smartphones improves sustained attention, mental health, and subjective well-being. *PNAS Nexus*. 2025;4(2)doi:10.1093/pnasnexus/pgaf017.

<sup>458</sup> Ra CK, Cho J, Stone MD, et al. Association of Digital Media Use With Subsequent Symptoms of Attention-Deficit/Hyperactivity Disorder Among Adolescents. *JAMA*. 2018;320(3):255-263. doi:10.1001/jama.2018.8931

risk of attentional problems at 12 and 24 months of follow-up ( $p < .001$ ). Briefly, an effect size of .11, if true, has significant public health implications at scale and a p value of less than .001 means that there is less than a 1/1000 chance that the association *is a false positive*. The major limitations of this study were that it relied on self-report of media usage and that it was observational meaning there could still be residual confounding, although the authors adjusted for multiple possible confounders including depression, delinquent behavior, socioeconomic status, substance use etc.

465. I next used the web of science database to see what studies have cited this high profile one since its publication. That strategy yielded several relevant studies as well as a systematic review “Longitudinal associations between digital media use and ADHD symptoms in children and adolescents.”<sup>459</sup> Because that systematic review includes all the relevant studies I identified except one that was published after it was, I will present the review’s findings and then the subsequent study briefly.

466. Thorell and colleagues’ systematic review identified 25 relevant studies related to use of digital devices and subsequent ADHD symptoms. The studies were too heterogenous (varying ages, follow-up periods, outcome measures, digital media usage measures etc.) to create a meta-analytic summary estimate so they were summarized in narrative format. Also notable was that some studies assessed the role of problematic screen use either at baseline or at follow-up. They also explored the bidirectionality of the relationship meaning they looked both at screen use predicting ADHD symptoms and ADHD symptoms predicting screen use. Below is a summary figure from their paper. The most robust analyses are those that are within subject (meaning they

---

<sup>459</sup> Thorell LB, Buren J, Strom Wiman J, Sandberg D, Nutley SB. Longitudinal associations between digital media use and ADHD symptoms in children and adolescents: a systematic literature review. *Eur Child Adolesc Psychiatry*. Aug 2024;33(8):2503-2526. doi:10.1007/s00787-022-02130-3.

looked at changes to individuals over time) or those that are between subjects but controlled for baseline levels (highlighted in red).

**Figure 54: Summary of the Number of Studies Showing Significant Associations Between ADHD Symptom Levels and Digital Media (DM)<sup>460</sup>**

**Fig. 2** Summary of the results displaying the number of studies showing significant associations between ADHD symptom levels and digital media (DM)

References DM → ADHD	Within-subject effects		References ADHD → DM
McNamee et al. [45]	Screen time	1 of 4 1 of 3	Beyens et al. [44]
Boer et al. [46]	Problem use	1 of 2 0 of 2	
References DM → ADHD	Between-subject effects with control for baseline levels		References ADHD → DM
Allen et al. [52] Barlett et al. [53] Baumgart... et al. [47] Beyens et al. [44] Gentile et al. [54] Parkes et al. [55] Poulain et al. [56] Ra et al. [57]	Screen time	8 of 14 3 of 9	Beyens et al. [44] Gentile et al. [54] Stenseng et al. [40]
Boer et al. [46] Hygen et al. [42] Wartberg et al. [61]	Problem use	3 of 4 7 of 10	Chen et al. [63] Ferguson & Cera... [62] Hirota et al. [68] Jeong et al. [65] Peeters et al. [66] Wartberg et al. [61] Wartberg et al. [67]
References DM → ADHD	Between-subject effects without control for baseline levels		References ADHD → DM
Barlett et al. [53] Beyens et al. [44] Boer et al. [46] Gentile et al. [54] Hetherington [48] Liu et al. [49] Ra et al. [57] Stenseng et al. [40]	Screen time	8 of 9 7 of 9	Beyens et al. [44] Boer et al. [46] Gentile et al. [54] Männikkö et al. [51] Rydell & Brocki [50] Stenseng et al. [40] Yang et al. [64]
Boer et al. [46] Hygen et al. [42] Wartberg et al. [61]	Problem use	3 of 3 6 of 6	Boer et al. [46] Hygen et al. [42] Jeong et al. [65] Peeters et al. [66] Wartberg et al. [61] Wichstrøm et al. [41]

<sup>460</sup> Thorell LB, Buren J, Strom Wiman J, Sandberg D, Nutley SB. Longitudinal associations between digital media use and ADHD symptoms in children and adolescents: a systematic literature review. *Eur Child Adolesc Psychiatry*. Aug 2024;33(8):2503-2526. doi:10.1007/s00787-022-02130-3.

Overall, the authors (rightly in my opinion) conclude that the majority (74%) of studies found significant associations between digital media and subsequent ADHD symptoms. The effect sizes were on in general less than .30. Although these are not experimental studies, they are longitudinal and within subject which supports causality versus merely cross-sectional observational study.

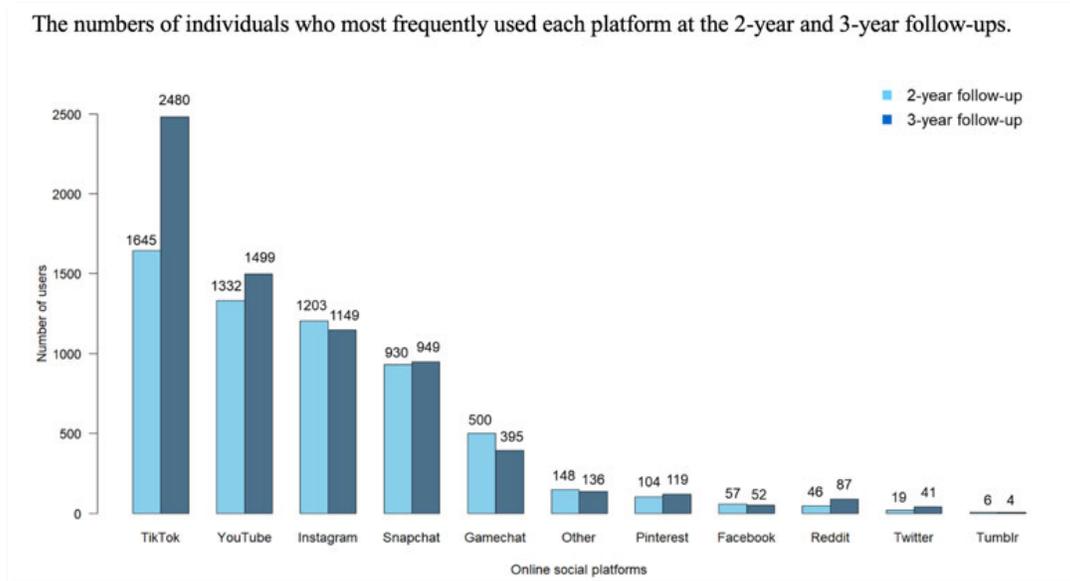
467. Again, the primary limitations of the studies were reliance on self-reported media use and the observational nature of them all even though they were longitudinal. As before, experimental studies in this space, with sufficiently long follow up periods would be exceedingly difficult to deploy. Further, relevant to this case, digital media usage included all forms although as expected the predominant forms were social media sites as detailed above.

468. Deng and colleagues used the over 11,000 individuals from the ABCD study, a large, national US cohort of early adolescents followed longitudinally with extensive and comprehensive data collection to look at online social activity *specifically* and subsequent ADHD symptoms over a 3-year follow-up period adjusting for baseline characteristics.<sup>461</sup> The social media sites they examined are presented below. Overall, the most frequently used platforms were TikTok (36%), YouTube (22%), Instagram (17%) and Snapchat (14%). The rest comprised ~11% of usage. They found a small effect but highly statistically significant effect size of .07 ( $p < .001$ ). Again, this overall effect may be more pronounced in specific sub populations at greater risk of ADHD.

---

<sup>461</sup> Deng H, Song K, Geng X, et al. Online social activity time predicts ADHD problems in youth from late childhood to early adolescence in the ABCD study. *European Child & Adolescent Psychiatry*. 2024/12/26 2024;doi:10.1007/s00787-024-02620-6.

**Figure 55: Chart Summarizing the Number of Individuals Who Most Frequently Used Each Platform 2-year and 3-year Follow-ups<sup>462</sup>**



469. This phenomenon is reflected in Defendants’ internal documents. For example, in a survey of clinicians regarding “the role of Social Media in Mental Health,” some participants reported that social media “exacerbates some symptoms in particular that ADHD already causes, such as irritability and anxiety” and in children can lead them to be “more hyperactive, [and] more impulsive.”<sup>463</sup> Likewise, YouTube’s documents support a similar conclusion finding that “[t]he short form viewing experience exacerbates concerns of addiction and ADHD.”<sup>464</sup> The same is true for TikTok, whose documents reported that an analysis of 28 studies “showed a bi-directional relationship between usage and [ADHD] symptoms” and that “children with ADHD symptoms

<sup>462</sup> Deng H, Song K, Geng X, et al. Online social activity time predicts ADHD problems in youth from late childhood to early adolescence in the ABCD study. *European Child & Adolescent Psychiatry*. 2024/12/26 2024;doi:10.1007/s00787-024-02620-6.

<sup>463</sup> META3047MDL-019-00099847 at Slide 37.

<sup>464</sup> GOOG-3047MDL-00793501 at Slide 290.

were at higher risk for excessive use, and excessive use was related to more ADHD symptom development.”<sup>465</sup>

470. It is my opinion to a reasonable degree of medical and scientific certainty, that social media usage causes cognitive impairment and ADHD symptomology. This observation is well-supported in the literature and occurs through both direct and indirect mechanisms. There is additional support within the companies’ internal documents for a causal relationship between their platforms and ADHD symptoms.

### **G. School Performance**

471. There are concerns about cell phone usage during school hours which can detract from learning, distract teachers, increase bullying, and reduce opportunities for physical play and in person interaction. Reliable data using passive sensing during school time are limited. In the context of the same study reference above, Common Sense Media assessed it and reported that children spend a median of 43 minutes per day on phones during school hours.<sup>466</sup> I have also published on this issue. Our study using both Android and iOS (also referenced above) found that they spend an average of 1.5 hours (95% CI, 1.31-1.73) on smartphones during the 6.5 hours of school, accounting for approximately 27% of average 24-hour phone use of 5.59 hours daily. Furthermore, in our sample, 25% of adolescents spent more than 2 hours on their phone during school.<sup>467</sup> Results are presented below.

---

<sup>465</sup> TIKTOK3047MDL-062-01192752, -2752

<sup>466</sup> Radesky J, Weeks HM, Schaller A, Robb M, Mann S, Lenhart ACCAWitLoaYPsSUSF, CA: Common Sense. Constant Companion: A Week in the Life of a Young Person's Smartphone Use. 2023.

<sup>467</sup> Christakis DA, Mathew GM, Reichenberger DA, Rodriguez IR, Ren B, Hale L. Adolescent Smartphone Use During School Hours. *JAMA Pediatrics*. 2025;doi:10.1001/jamapediatrics.2024.6627

**Figure 56: Summary of Adolescent Daily Smartphone Use<sup>468</sup>**

Characteristic	No. of participants <sup>a</sup>	Minimum	25th Percentile	Median	75th Percentile	Maximum	Participants, % <sup>b</sup>
Smartphone use, h							
24-h Smartphone use	117	0.16	3.62	5.49	7.04	19.18	5.59 (5.06-6.12)
24-h Social media use	117	0.00	0.47	2.00	3.13	6.31	2.14 (1.84-2.45)
School-day smartphone use	117	0.04	0.68	1.11	2.16	5.84	1.52 (1.31-1.73)
School-day social media use	117	0.00	0.20	0.39	0.76	2.46	0.60 (0.49-0.71)
Smartphone application use during school day, min							
Messaging and chat	88	1.00	2.69	5.82	26.92	186.96	19.46 (13.39-25.53)
Instagram	82	1.18	8.64	13.32	26.42	269.42	24.61 (17.91-31.31)
TV, movie, or video streaming	68	1.22	4.02	7.86	21.67	132.49	17.19 (10.99-23.39)
Music, media, and podcasts	58	1.01	1.81	3.21	4.82	99.49	4.48 (2.79-6.17)
Email	53	1.02	1.33	2.74	5.21	15.63	3.92 (2.94-4.89)
TikTok	48	1.17	6.05	9.23	35.75	71.99	18.88 (13.23-24.54)
Shopping and entertainment	47	1.01	1.28	2.22	4.45	39.02	5.21 (3.01-7.41)
Games	44	1.01	3.11	9.68	18.88	100.29	13.43 (8.84-18.02)
Facebook	40	1.18	2.97	7.33	20.27	87.39	19.88 (11.05-28.71)
Utilities	39	1.03	1.30	2.38	5.51	21.40	4.27 (2.82-5.71)

By any measure, more than 2 hours of phone usage during school negatively impacts the time and attention available for learning and engaging with fellow students and teachers.

472. By number of users, the top 5 most frequented apps or categories (excluding internet browsers) were messaging, Instagram, video streaming, audio, and email. This study found that teens spend approximately 25% of the school day on their phones either in or out of class but in either case it diminishes engagement in human face to face interactions that are foundational to socio-emotional and cognitive learning. Studies in this area continue.

473. A highly publicized and recently completed cross-sectional study in England assessed the association between school phone policies, phone use during school hours, and mental well-being. They found that although restrictive policies result in less phone use during school

<sup>468</sup> Christakis DA, Mathew GM, Reichenberger DA, Rodriguez IR, Ren B, Hale L. Adolescent Smartphone Use During School Hours. *JAMA Pediatrics*. 2025;doi:10.1001/jamapediatrics.2024.6627.

hours, there was no significant differences in mental well-being.<sup>469</sup> The median cell phone usage in restrictive schools was about 45 minutes less than non-restrictive ones. However, the 10<sup>th</sup>-90<sup>th</sup> percentiles were the same. They also reported that social media time *during* school was associated with small increases in anxiety (.06) and depression (0.11). And *overall* social media usage per week was cross-sectionally associated with decreased well-being, and increased anxiety and depression consistent with many other similar studies discussed in greater detail elsewhere in the report. This study suggests that phone policies alone—or failure to enforce them better-- are not effective at fully addressing children’s mental health.

474. Finally, it should be noted that all the above mental health harms – increased depression, anxiety, negative social comparison, FoMo, sleep disruption, body image issues and eating disorders – do not dissipate in the classroom. These conditions instead adversely affect the learning environment. A recent National Education Association report documented marked increases in student mental health issues at school, an inability of students to concentrate, and negative impacts such as cyberbullying and underdeveloped social skills that carry over into the classroom.<sup>470</sup>

475. Children spend more time in school than in any other place but their homes. Accordingly, all untoward outcomes of social media ultimately impact their school experience and their school performance. For that reason, the “School” section of this report is situated near the end: all paths ultimately lead to it.<sup>471</sup> The problems students bring to school, they also bring into

---

<sup>469</sup> Goodyear VA, Randhawa A, Adab P, et al. School phone policies and their association with mental wellbeing, phone use, and social media use (SMART Schools): a cross-sectional observational study. *The Lancet Regional Health – Europe*. doi:10.1016/j.lanpe.2025.101211.

<sup>470</sup> <https://www.nea.org/resource-library/impact-social-media-and-personal-devices-mental-health>

<sup>471</sup> It should also be noted that internal documents reflect that at least Snap recognized school bans could change the trajectory of daytime use. *See e.g.*, SNAP7428962.

the classroom where teachers—tasked with educational responsibilities—must also address or attempt to overcome them. Suffice it to say that sleep-deprived, distracted, anxious, depressed, bullied students are more challenging to teach.

476. The literature on social media’s direct effects on school performance is minimal. To wit, the review criteria entered for systematic reviews about social media and school only yielded results that were mediated by the other outcomes discussed above (e.g. Sleep, Anxiety etc). As these reviews largely recapitulated and overlapped with the prior ones already reviewed except insofar as they sometimes narrowly focused on “grades,” I will not revisit them here but will focus instead on what we know (so far) about the other pathways that might impact school performance.

#### **H. Cyberbullying and Risky Behaviors**

477. Many studies connect social media use to an increase in cyberbullying and risky behaviors. These behaviors impact the school environment. “Risky behaviors” can include drug or alcohol use, sexting, high risk sexual encounters, weapon carrying, or violence perpetration. Cyberbullying includes sending hurtful messages, spreading gossip about others, and getting others to disclose private information that is then shared online.<sup>472</sup> It is recognized in the medical and scientific community that these behaviors tend to occur among peers from school. This necessitates the need for schools to address cyberbullying, notwithstanding the difficulties in doing so.<sup>473</sup>

478. The mechanism driving an increased risk of engaging in risky behaviors because of social media use can be explained by examining Social Cognitive Theory, or SCT. Dr Albert Bandura (1925-2021) was a world-renowned professor of social science and psychology at

---

<sup>472</sup> Handbook of Children and Screens at p. 433.

<sup>473</sup> Handbook of children and Screens at p. 434-436.

Stanford University who originated SCT, a psychological theory that explains how people learn by observing others, emphasizing the dynamic interaction between a person's thoughts, behaviors, and their environment, with a key concept being "reciprocal determinism" where these factors continuously influence each other. Essentially, SCT posits that people learn not just through direct experience but also by watching and imitating others, with their own cognitive processes playing a significant role in how they interpret and apply what they observe.

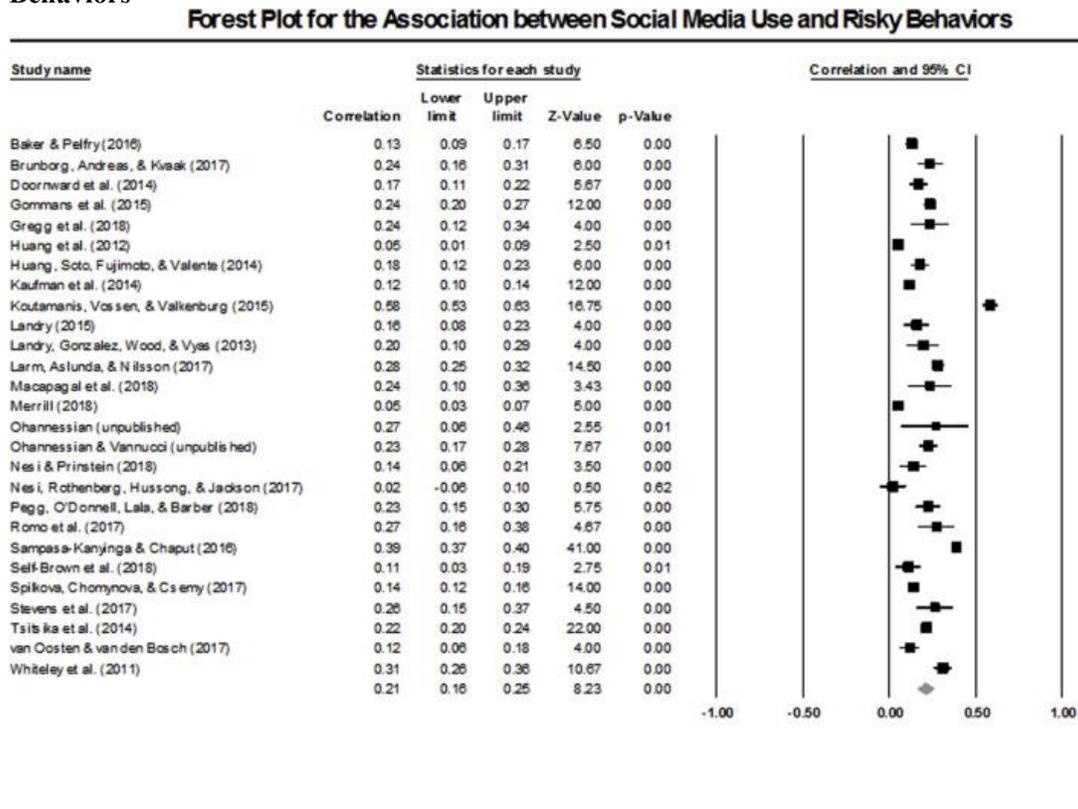
479. In his early, pioneering work, Dr. Bandura demonstrated through a series of experiments that children would imitate novel behaviors—even unusual or violent ones—if they saw them modeled and they were even more likely to imitate them if they were performed by people or characters (including cartoons) the children respected for some reason.

480. SCT informs behaviors at all ages but younger people, especially those without fully developed pre-frontal cortices are more prone to its effects. Adolescents are especially developmentally susceptible to behavioral suggestions as they seek to differentiate their emerging “adult” identity from their childhood one. They are poised to adopt behaviors that they perceive as “cool,” popular, or desirable, but lack the self-control to limit the impulse to take undue risk. SM provides copious opportunities for pre-teens, teens, and emerging adults to emulate behaviors – good and bad. Notably, SM companies are eager to take credit for the spread of virtuous behaviors via their platforms (e.g. the Ice Bucket Challenge for ALS research) but equally eager to distance themselves from the pernicious ones (e.g. Blue Whale Challenge as a suicide pact). But the mechanism is the same for both; they are, if you will, two sides of the same Bandurian SCT coin.

481. A meta-analysis of 27 cross-sectional studies examining the association between self-reported risky behaviors and self-reported social media use found a small to medium effect size (.21) (see below for summary table) The modest effect size should be interpreted as conservative

because the predictor variable was *overall* social media use and the key driver would be the content viewed (e.g. risky behavior) which was not collected.<sup>474</sup> Further, the effect is likely to be greater among adolescent boys given their predisposition to take risks, but the data were not stratified by gender.

**Figure 57: Forest Plot for the Association Between Social Media Use and Risky Behaviors<sup>475</sup>**



<sup>474</sup> Vannucci, Anna & Simpson, Emily & Gagnon, Sonja & Ohannessian, Christine. (2020). Social media use and risky behaviors in adolescents: A meta-analysis. *Journal of Adolescence*. 79. 258-274. 10.1016/j.adolescence.2020.01.014.

<sup>475</sup> Vannucci, Anna & Simpson, Emily & Gagnon, Sonja & Ohannessian, Christine. (2020). Social media use and risky behaviors in adolescents: A meta-analysis. *Journal of Adolescence*. 79. 258-274. 10.1016/j.adolescence.2020.01.014.

“Risky behaviors” in the context of this study were defined as drug or alcohol use, sexting, high risk sexual encounters, weapon carrying, or violence perpetration. These are the most common adolescent risk behaviors but the mechanism underlying undertaking them is the same as that for other less common “risky behaviors” (i.e. subway surfing).

482. This metanalytic summary is not without limitations, most notably that the studies were all cross-sectional so causality cannot be proven; it could be that risk taking adolescents also seek risk taking posts online. There was also significant heterogeneity among studies, making combining the results questionable and conservative.

483. Longitudinal studies of SM exposure to risky behavior and subsequent risk taking are limited but there are ample studies of other media exposures (e.g. tobacco and alcohol use and in movies and advertising) and subsequent drinking and smoking.<sup>476</sup> The Bridge *et al* study detailed before also examined the impact of viewing suicide relevant content to risk of committing suicide.<sup>477</sup>

484. Finally, a metanalytic summary of 29 studies examined the effects of exposure to tobacco on social networking sites and subsequent smoking.<sup>478</sup> In the figure below which includes both cross-sectional and longitudinal studies, exposure to tobacco use on SMS was associated with

---

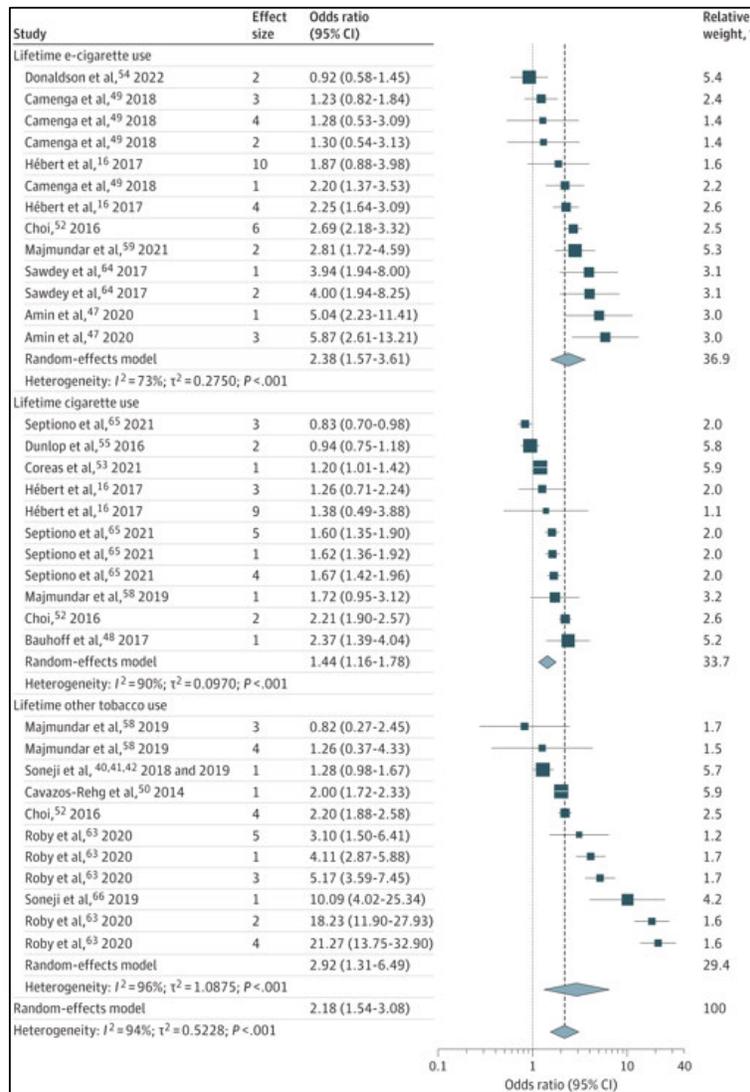
<sup>476</sup> Tang S, Werner-Seidler A, Torok M, Mackinnon AJ, Christensen H. The relationship between screen time and mental health in young people: A systematic review of longitudinal studies. *Clin Psychol Rev.* Jun 2021;86:102021. doi:10.1016/j.cpr.2021.102021; Franklin JC, Ribeiro JD, Fox KR, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol Bull.* Feb 2017;143(2):187-232. doi:10.1037/bul0000084

<sup>477</sup> Bridge JA, Greenhouse JB, Ruch D, et al. Association Between the Release of Netflix’s 13 Reasons Why and Suicide Rates in the United States: An Interrupted Time Series Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2020/02/01/2020;59(2):236-243. doi:<https://doi.org/10.1016/j.jaac.2019.04.020>

<sup>478</sup> Donaldson SI, Dormanesh A, Perez C, Majmundar A, Allem JP. Association Between Exposure to Tobacco Content on Social Media and Tobacco Use: A Systematic Review and Meta-analysis. *JAMA Pediatr.* 2022 Sep 1;176(9):878-885. doi: 10.1001/jamapediatrics.2022.2223. PMID: 35816331; PMCID: PMC9274450.

a 1.5 to 2.4 increased risk of smoking (depending on how it was defined) and even when the analysis was restricted to longitudinal studies the risk remained 1.5 times higher (data not shown).

**Figure 58: Forest Plot of 3-Level Meta-analysis for Exposure to Tobacco Content on Social Media and Lifetime Tobacco Use, Including e-Cigarettes, Cigarettes, and Other Tobacco Use<sup>479</sup>**



<sup>479</sup> Donaldson SI, Dormanesh A, Perez C, Majmundar A, Allem JP. Association Between Exposure to Tobacco Content on Social Media and Tobacco Use: A Systematic Review and Meta-analysis. *JAMA Pediatr.* 2022 Sep 1;176(9):878-885. doi: 10.1001/jamapediatrics.2022.2223. PMID: 35816331; PMCID: PMC9274450.

485. TikTok for its part appears to be aware of the risk posed by “harmful” and “dangerous” challenges. Below are two slides from TikTok presentations (Red boxes added):

**Minor Safety Risks**

**4 Pillars of Harm for Minors**

<b>Exploitation</b>	Predators offer teens attention with motives to exploit them for sexual gain or monetary profit.
<b>Developmental</b>	Developmentally inappropriate themes can lead to teens emulating adult activities without fully comprehending the ramifications.
<b>Psychological</b>	Upsetting themes can trigger psychological malaise in teens, and solidify permanent neural pathways for anxiety, depression, etc.
<b>Physical</b>	Teens' increased willingness to take risks, with minimal consideration for consequences, leads to imitation of physically dangerous behaviors.

*Document 157: TIKTOK3047MDL-001-00060986, -1000*

## Minor Safety Risks

### Each harm has different triggers & vectors

<u>Harm</u>	<u>Trigger</u>	<u>Vector</u>
<b>Exploitation</b>	Posting self-sexualized content. Posting nude or partially nude images. Posting requests for love, attention, romance, affection, etc.	<b>CREATE</b>
<b>Developmental</b>	Exposure to hypersexualized adults and youth. Exposure to sexual acts and erotic nudity. Exposure to substance misuse, abuse, and addiction.	<b>CONSUME</b>
<b>Psychological</b>	Exposure to diet trends, body-size valuations, etc. Exposure to overt gore, physical violence, real-world trauma. Exposure to themes of suicide, self-harm.	
<b>Physical</b>	Exposure to dangerous challenges and other physical dares. Exposure to unhealthy diet and exercise tips. Exposure to glorified experiences using drugs and alcohol.	

*Document 158: TIKTOK3047MDL-001-00060986, -1001*

These documents confirm that internally, such videos were viewed as posing risks to teens, enough so that tracking them and removing them was a goal although it appears to have been an under resourced one:

Tiktok Challenge	4	<ul style="list-style-type: none"> <li>* <b>#1: Lug nut Challenge</b>, (Old, 2 escalation this week, 2 in history) <ul style="list-style-type: none"> <li>o Users call it the "lug nut challenge" and are loosening or removing lugs from car tires in hopes of gaining likes online.</li> <li>o No related cases were located by IM/Algo. Related bits ticket [ <a href="https://bits.byteoversea.net/rocket/story?projectId=2648&amp;type=list&amp;viewType=require&amp;viewId=2415&amp;rid=1074548&amp;storyTab=detail">HYPERLINK</a> ]</li> </ul> </li> <li>* <b>#2: Dryscoping Challenge</b>, (Old, 1 escalation this week, 1 in history) <ul style="list-style-type: none"> <li>o The challenge was dimmed violative in Jun and sweeping was done. We still find some dry scooping challenge related videos on the platform. New request for sweeping has been submitted.</li> <li>o Related bits ticket [ <a href="https://bits.byteoversea.net/rocket/story?projectId=2571&amp;type=list&amp;viewType=require&amp;viewId=2502&amp;storyTab=detail&amp;rid=1049385">HYPERLINK</a> ]</li> </ul> </li> <li>* <b>#3: Slap Teacher Challenge</b>, (Old, 1 escalation this week, 1 in history) <ul style="list-style-type: none"> <li>o A video showing students punching teacher was shared offline, uploaded to Instagram first then uploaded to Tiktok (all removed now). So far haven't find neither much cases nor evidence of students being inspired by Tiktok contents.</li> <li>o Related bits ticket [ <a href="https://bits.byteoversea.net/rocket/story?projectId=2648&amp;type=list&amp;viewType=require&amp;viewId=2415&amp;rid=1074697&amp;storyTab=detail">HYPERLINK</a> ]</li> </ul> </li> </ul>
------------------	---	---

Document 159: TIKTOK3047MDL-002-00094384, -4385

**Harmful challenges (Dangerous challenges, illegal challenges, ANSA challenges) detection**

Harmful challenges are (anecdotally) the most common issue we have for minors that causes real world harm or media escalations. We want to try to detect these challenges proactively, before they are in the media, or before minors are harmed in some way.

**2) Do we have enough engineers on Live Safety?**

*Document 160: TIKTOK3047MDL-004-00290064, -0064*

486. In summary, it is my opinion that to a reasonable degree of medical and scientific certainty, social media via the algorithms it developed and deployed and addictive design increases the risk of both risky behaviors and cyberbullying. These effects impact the school environment, as these issues tend to arise between classmates. A review of internal documents provides additional support that social media use increases the risk of these two behaviors.

**I. Unwanted Interactions from Adults**

487. As discussed extensively in Section VI above, children are uniquely vulnerable by virtue of both brain and social development, and social media is a largely unregulated environment that can pose unique dangers that exploit these vulnerabilities. One particularly disturbing safety risk is the Defendant platforms' facilitation of child predators contacting and grooming children—a phenomenon that is well-documented in Defendants' internal documents, that Defendants did not successfully address despite notice of the same, and that no Defendant appears to have warned children or parents about. When children are being constantly pulled back online by the addictive features of social media, it's more likely that they will experience unwanted interactions with adults especially if the algorithms have already begun to target them. This is even more harmful when one considers that children are being targeted in a forum in which their ability to access crucial in-person support is limited.

488. The scope of the problem was known within Meta since at least 2020. A memo from Sayed Otaru (IG Safety Product Manager) to ██████████ (FB Safety Product Manager) on June 24<sup>th</sup>, 2020, related to Inappropriate Interactions with Children (IIC) states it as follows:

**Problem Summary**

Messenger implements a classifier that predicts adult accounts that engage in IIC with underage accounts. Messenger subsequently restricts those flagged adult accounts from being able to make friend requests to all non-adult accounts. Messenger does not de-platform flagged adult accounts and believes the restriction is a sufficient deterrent. Instagram (IG) has historically not been able to classify non-adult accounts, until recently. The CI team has concluded a study (the numbers have not yet been formalized) and currently show that **500,000 IG underage accounts receive IIC on a weekly basis**. The IG prevalence number is 3x Messenger's.

*Document 161: META3047MDL-003-00028214, -8218*

The actual figure according to Malia Andrus, who joins the discussion is “**500k victims per DAY in ENGLISH markets ONLY**” (*emphasis in original*).<sup>480</sup>

489. The persistence of the problem stems in part from the lack of classifiers according to Michael Rothschild, Senior Director of Product Management at Meta:

---

<sup>480</sup> META3047MDL-003-00028214, -8216

Michael Yehuda Rothschild (3/29/2020 14:34:33 PDT):  
>Btw, according to Manoli we don't have CEI/IIC classifiers integrated to filtering in recommendations. The only protection we have is tiering - ie 1 strike will filter that account out for 90 days. Does that seem right from what you know?

[REDACTED] (3/29/2020 14:39:20 PDT):  
>Yep, those classifiers are somewhere between non-existent and shitty.

Michael Yehuda Rothschild (3/29/2020 14:39:36 PDT):  
>I see. Bummer

[REDACTED] (3/29/2020 14:39:41 PDT):  
>The list of signals we have available is here:  
[https://our.internmc.facebook.com/intern/wiki/Instagram\\_Content\\_Signal\\_Service/#system-map](https://our.internmc.facebook.com/intern/wiki/Instagram_Content_Signal_Service/#system-map)

Michael Yehuda Rothschild (3/29/2020 14:40:15 PDT):  
>So the only protection is tiering which relies on strikes which are significantly lower volume due to CO shutdown

[REDACTED] (3/29/2020 14:40:21 PDT):  
>Correct.

[REDACTED] (3/29/2020 14:40:31 PDT):  
>but those low-prevalence problems are in a shitty state across all of Facebook, Inc.

[REDACTED] (3/29/2020 14:41:01 PDT):  
>People seem to refuse to believe it, sadly, across teh family of apps. Everyone wants those classifiers to be better, but they suck, and they're not improving fast.

*Document 162: META3047MDL-014-00349418, -9418*

490. [REDACTED] goes on to say, “Most of our policies I don’t believe will have really good classifiers for years, if ever.”<sup>481</sup> In an email a few months after this chat, with apparent indifferent resignation, [REDACTED] states on June 26, 2020, “Child Safety is explicitly called out as a non-goal in our H2 plans. So if we can do something here, cool. But if we can do nothing at all, that’s fine too.”<sup>482</sup> Rothschild writes, “I agree that Rothschild writes, “I agree that [child safety] is a non-goal but also agree that these numbers [500k IIC’s per day] are quite alarming.”<sup>483</sup>

491. In November 2020, Meta summarized what it called its “vulnerabilities” in this space in a slide entitled “Child Sexual Exploitation: State of Play”:

<sup>481</sup> META3047MDL-014-00349418, -9419

<sup>482</sup> META3047MDL-014-00350154, -0155

<sup>483</sup> META3047MDL-014-00350154, -0155

## Vulnerabilities

- IG: minimal child safety protections
- Classifiers: CSAM low accuracy prevents automation; grooming lacks global efficacy
- Data: under resourced, just unlocking valuable insights
- Review: capacity constraints
- User education: minimal
- Growth: not resourced to address growth concerns with valuable mitigations like forward and group size limitations
- Groups: need stronger policies and review tools
- Minor sexualization: policies and classifier immature
- Trafficking: minimal focus, no classifier
- Interop: increases discoverability and reachability of minors across apps, need stronger protections; particularly for IG
- Live, video chat and remote presence products: introduce new vector for abuse, lack reporting with content capture
- Ephemerality: impacts efficacy of reporting

*Document 163: META3047MDL-004-00027515, -1518*

The long list of “vulnerabilities,” many of which were known for years, could be read as a “to-do” list (one that lacked completion). That Meta failed to address these issues is amply demonstrated by a shareholder proposal submitted three years later to Meta’s Board of Directors (which the board recommended voting against):

#### Meta Platforms — Child Safety — 2023

The internet was not developed with children in mind. Social media impacts children's brains differently than adult brains.<sup>1</sup> It also poses physical and psychological risks that many children and teens are unprepared for, including sextortion and grooming, hate group recruitment, human trafficking (for any means), cyberbullying and harassment, exposure to sexual or violent content, invasion of privacy, self-harm content, and financial scams, among others.

Meta is the world's largest social media company with billions of children and teen users. Meta's platforms, including Facebook, Instagram, Messenger and WhatsApp, have been linked to numerous child safety impacts and social policy challenges, including:

##### Mental Health:

Meta's own company research showed Instagram's negative impacts on teens' self-image, increased rates of depression and anxiety, and a link to suicidal thoughts. The *Wall St. Journal* concluded that these Instagram documents revealed "Facebook has made minimal efforts to address these issues and plays them down in public."<sup>2</sup>

##### Sexual Exploitation:

In 2021, nearly 29 million cases of online child sexual abuse material were reported; nearly 27 million of those (92 percent) stemmed from Meta platforms— including Facebook, WhatsApp, Messenger and Instagram.<sup>3</sup> A *Forbes* report on Instagram pedophiles described Instagram as "a marketplace for sexualized images of children."<sup>4</sup>

##### Cyberbullying:

*Time Magazine* reported that "By one estimate, nearly 80% of teens are on Instagram and more than half of those users have been bullied on the platform."<sup>5</sup> A UK study found that Instagram accounted for 42 percent of online bullying, followed by Facebook with 39 percent.<sup>6</sup>

*Document 164: META3047MDL-050-00343376, -3465*

492. "Discoverability" and "reachability" of minors by unknown adults and in particular predators was a longstanding and long ignored issue at Instagram in particular. Instagram allows people to find each other and while there can be additional privacy settings that restrict who can see one's content, the default setting for teens for many years was set to "public." When set that way, anyone can direct message (DM) someone they find on Instagram.

493. In July 2019, a researcher at Instagram, Hitomi Hayashi-Branson, recommended that this setting be changed and that "smart defaults" be applied "to make it easier for users to leverage and benefit from existing privacy and security tools so they can feel safe on Instagram."<sup>484</sup> The researcher recommended applying new privacy defaults because "most users don't know about the tools we have ...And are not likely to go into settings to find them."<sup>485</sup> The researcher explained how, "In a research study on US Tweens and their social media use tweens were concerned with their safety online and were particularly aware of stranger-danger and bullying

---

<sup>484</sup> Diego Castaneda Dep. Ex. 14 at -8266.

<sup>485</sup> Diego Castaneda Dep. Ex. 14 at -8248

online. And while they did not seek out trouble, they were aware that trouble could find them. In this study, parents had similar concerns around privacy and bad actors.”<sup>486</sup>

494. Other internal documents at Instagram make the case for why teen account should be set to “private,” including that these defaults would be “in-line” with “teen user expectations,” “parent expectations,” and “safety experts’ expectations.”

**Why should we default teens to private accounts?**

- 1. Will Increase Teen Safety:** Defaulting teens to private will help prevent high severity actions such as child grooming and inappropriate contact with minors. As a result of defaulting to private, private account adoption increases by 5.5x for teens, making it harder for bad actors to discover and message them.
- 2. In-line With Teen User Expectations:** Parents and teachers often warn teens to avoid “creepy older people on the internet.” Because of that, most teens prefer private accounts and wish to see privacy controls during onboarding.” ([First 14-days research from Yujian](#))
- 3. In-line With Parent Expectations:** Parents are worried about the security and privacy of information and who can contact them/their teens. While parent usage of Instagram is relatively low, they are aware of private accounts and said it limits the number of strangers that can reach out ([Talking about Safety Research from Wendy](#)).
- 4. In-line With Regulator Expectations:** Providing additional resources and protections—such as private-by-default—for teens will strengthen many of our legal and policy positions across multiple jurisdictions as they relate to the processing of the data of minors, and may soon be required in some countries.
- 5. In-line With Safety Experts' Expectations:** Safety partners and civil society groups regularly flag for us that defaulting teens to private accounts is necessary to give them a safe onboarding experience to our platform.
- 6. Previously Announced Externally:** We announced this as part of our age collection moment in December 2019. The inclusion of the safety rationale for collecting age helped us land this favorable in press and avoid criticism that this was a move solely intended to help with ads targeting. We’ve also discussed with third-party policy partners (e.g., Anti-Bullying Association in UK, Unilever 5Rights Foundation)
- 7. Decreased Follows from Growth Hackers:** By defaulting teens to private, we decreased # teens with inbound follows from Growth Hackers by 24% and # growth hacker follows by 25% during their first 7 days on the platform.

Document 165: Darius Kilstein Dep. Exhibit 57 at -6982.

[REDACTED] (7/09/2020 11:42:58 PDT):  
>what specifically are we doing for child grooming (something I just heard about that is happening a lot on TikTok)?

[REDACTED] (7/09/2020 11:43:36 PDT):  
>Somewhere between zero and negligible. Child safety is an explicit non-goal this half. I'd argue we're making it worse with Interop, but's that's a can of worms.

Document 166: META3047MDL-003-00069904, -9908

<sup>486</sup> Diego Castaneda Dep. Ex. 14 at -8251

495. Unfortunately, it took Meta several years (and thousands of lawsuits) to act on these findings. The reasons for the delay are made clear by the documents. In a 2019 exchange between Darius Kilstein and ██████████, design lead for growth at Instagram, Mr. Kilstein predicted that private by default would “smash engagement, DAP, MAP, etc.”<sup>487</sup>

```
Darius Kilstein (8/29/2019 14:20:03 PDT):
>holy shit

Darius Kilstein (8/29/2019 14:20:21 PDT):
>what's the rationale for the push?

Darius Kilstein (8/29/2019 14:20:41 PDT):
>This will likely smash engagement, DAP, MAP etc

██████████ (8/29/2019 14:20:53 PDT):
>current climate? policy pressure? potentially contributing to teen suicides?

Darius Kilstein (8/29/2019 14:21:24 PDT):
>What policy pressure?

██████████ (8/29/2019 14:21:36 PDT):
>what are we doing to protect minors

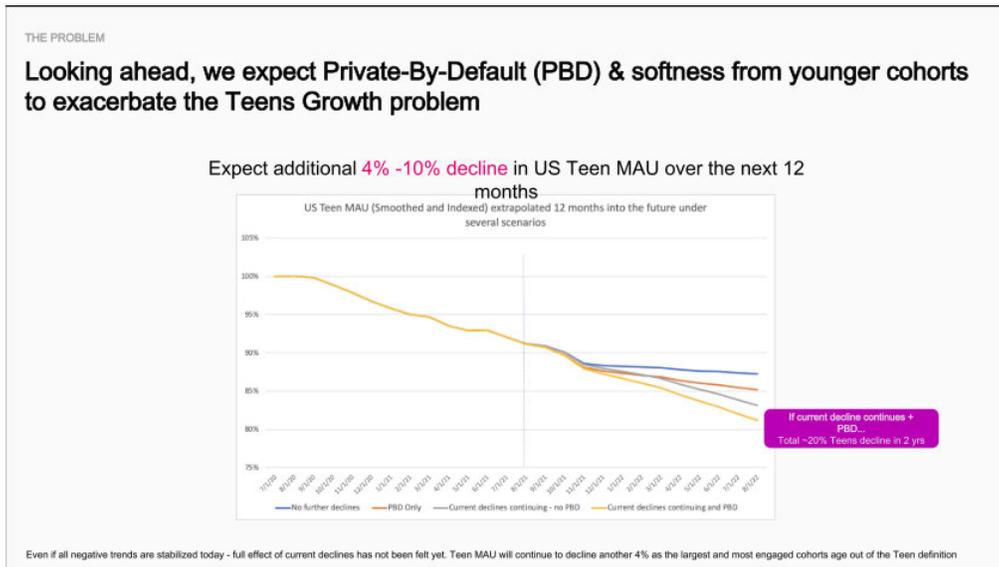
Darius Kilstein (8/29/2019 14:21:40 PDT):
>i see
```

*Document 167: META3047MDL-003-00005175, -5175*

496. The concern that “engagement” might be “smashed” came at a time when Instagram was already concerned about losing young teen market share as the below graphic illustrates.

---

<sup>487</sup> Darius Kilstein Dep. Ex. 29 at -5175



Document 168: META3047MDL-035-00002761 at Slide 32

Hence, despite considering a change, in March 2020 Meta leadership wrote researchers to “explicitly say that we will not be defaulting anyone to new interaction settings.”<sup>488</sup> The documents I have reviewed leave little room for doubt that growth considerations motivated this decision. Dr. Castaneda acknowledged in a private message that the project was shelved due to “a potentially untenable problem with engagement and growth”:

Dr. Diego Emiliano Castaneda (3/12/2020 01:09:16 PDT):  
 >clearly, we are in a bit of a churn (maybe have been for awhile). got to ride it out bro. Listen, I've seen and experienced worse.. my last team. for this big redesign project---the designers and engineers were at dagger heads for like 2 months due to philosophical differences. research came in and negotiated a sit down and turned it around. That was interesting.  
 >So, I'm not going to let you go to bed miserable. 1) Your data pretty succinctly shows that taking away unwanted interactions via private default settings is likely to lead to a potentially untenable problem with engagement and growth. 2) This is supported by qualitative data that teens don't want to miss out on discovering, connectin, and engaging. 3) The data also shows that we would fix a lot of the unwanted interaction problem by making DM's private. At what cost. to paraphrase Sam, how do we increase safety without decreasing value?  
 >incidentally, his original wording "what are the other ways we could increase safety without decreasing value?" is indicative of him wanting to keep the defaults.

Document 169: Diego Castaneda Dep. Exhibit 12 at 5

Ms. Jayakumar communicated to Ms. [REDACTED] (IG Privacy Lead) that the proposed launch of “teen private by default was scrapped due to growth concerns.”<sup>489</sup>

<sup>488</sup>Castaneda Dep. Ex. 12

<sup>489</sup> Vaishnavi Jayakumar Dep. Tr. at 69:6-10

497. This was particularly worrisome given that the potential problem posed by predators DM'ing minors on Instagram was recognized as it prepared to launch "Reels" in order to compete with TikTok, which had become the dominant platform in short order.

- **Launch of Reels v2:** Reels v2 launched on 6/24 in Germany, France, and Brazil. Ahead of the launch, we reviewed the curation guidelines for Featured Reels and developed a set of reactive talking points for IG Policy around our top safety considerations. Reels surfaces unconnected users and content to Instagram's audience, which changes the degree of virality and visibility an otherwise unknown account would get. This is great for discovering new creators, but can also come with more unwanted interactions and unsafe experiences on the platform. It is tentatively scheduled to launch in the rest of the world in late August / early September.
  - IG: Megan, Palak | SP: Vaishnavi

*Document 170: META3047MDL-020-00271173*

Jayakumar expressed concerns that "viral" reels created by teens would expose them to a wide array of adults including potential predators who could then DM them, leading to grooming and sextortion. She proposed not making Reels public or at least restricting access once they reached a certain view threshold, but that recommendation was likewise not followed.<sup>490</sup> A presentation by Lee summarized her team's concerns about Reels:

---

<sup>490</sup> Vaishnavi Jayakumar Dep. Tr. at 47:1-48:13

I reviewed the related body of Reels, integrity, and well-being (equity, teens) research ([link](#)) to identify key priority areas for Reels.

**TLDR: The Reels ecosystem makes it more vulnerable to integrity challenges.**

- 1 There are **multiple** Reels teams launching a **high volume of product improvements**, leading to increases in sexually suggestive and objectionable content.  
Reels product teams should **coordinate** to release launches strategically, and **consider integrity impacts** proactively for new product launches
- 2 The focus on **entertainment / mimicry**, and the **teen audience** leaves Reels more open to 1) sexually suggestive and offensive content and 2) vulnerability of younger people.  
**Reels teams should build user-centered tools** (audience controls, recommendation feedback) **and preventive mechanisms** (i.e., tuning ranking models, protections for minors) to keep users safe.
- 3 **Teens, women, LGBT people, and communities of color** are more vulnerable to harm and marginalization on IG. Early research suggests these inequitable patterns are reproduced on Reels.  
Reels teams should use a **targeted universalism approach** to reduce inequities of harm and improve socio-culturally diverse content, resulting in **wins in integrity, engagement, and public perception**.

*Document 171: Alison Lee Deposition Exhibit 4 at Slide 2*

498. The introduction of “Reels” at Meta in August of 2020 did increase the prevalence of sexual content featuring minors.<sup>491</sup> In fact, “Sexually suggestive content featuring minors is 2-3 times higher on Reels than Explore.”<sup>492</sup> Lee further stated:

I think that content that was being distributed on reels that were reaching audiences that users did not intent as original audience. And we also conducted research that showed that young people had a preferred audience that they wished to reach.

And there was also research that suggested that particularly young women who were producing content on Reels were experiencing unwanted interactions as a result of their content being shown to people that they did not wish to have it be shown to.<sup>493</sup>

In response she was asked was that one example where the algorithmic distribution could endanger young users?” Her reply, “Yes, it could endanger young users as a result of that distribution.”<sup>494</sup>

<sup>491</sup> Alison Lee Dep. Tr. at 49:16-21

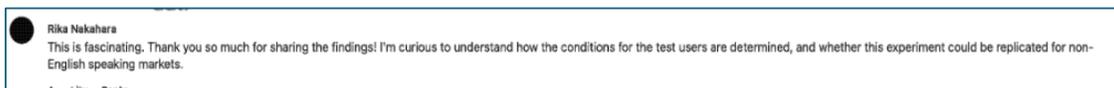
<sup>492</sup> Alison Lee Dep. Tr. at 98:6-10

<sup>493</sup> Alison Lee Dep. Tr. at 56:6-18

<sup>494</sup> Alison Lee Dep. Tr. at 58:24-25

499. What is more, Exhibit 6 in the Lee deposition has the following quote, “New approaches to protect children and teens (i.e. Preventing visibility between known IIC [Inappropriate Interactions with Children] adults should be ported to Reels.”<sup>495</sup> Lee confirmed in her deposition that there are accounts that have been “flagged” as having “had previous inappropriate interactions with children.”<sup>496</sup> She later confirmed that said adults were being limited in their ability to contact teens on other surfaces but not (yet) on Reels.<sup>497</sup> One might rightly ask why such individuals are allowed on the platform at all given the “adjudicated risk” they pose. Or at least, why parents were not informed that the site did not, by policy or practice, ban people they suspected might be child predators.

500. To their credit, for reasons that are not clear in the documents, in June of 2020, [REDACTED] on the Instagram investigation team performed a test in which she created a fake user and followed 70 accounts that were either sexualizing minors or talking about them inappropriately. In response, within 24 hours, the Instagram algorithms filled her Feed, Explore, and Recommended accounts with “almost exclusively minor sexualization content and accounts” including CEI.<sup>498</sup> When she reported this on a group chat, one of the first responses was:



*Document 172: META3047MDL-144-00000324, -0325*

501. The precise verbiage warrants repeating. “This is *fascinating*. (emphasis added) Thank you so much for sharing the findings!” Fascination and gratitude are not the first reactions

<sup>495</sup> Alison Lee Dep. Exhibit 6 at -8942

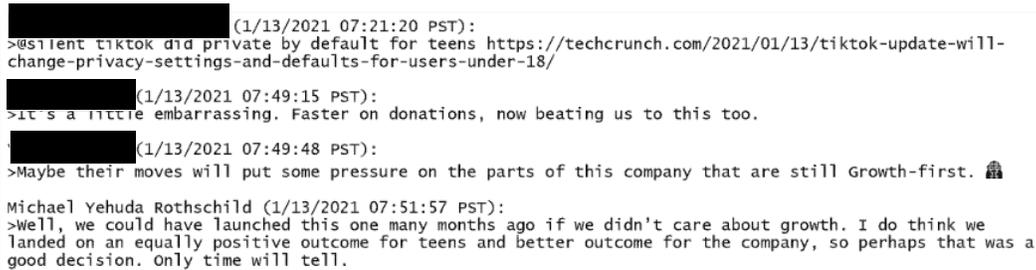
<sup>496</sup> Alison Lee Dep. Tr. at 106:9-12

<sup>497</sup> Alison Lee Dep. Tr. at 106:17-107:11

<sup>498</sup> Vaishnavi Jayakumar Dep. 31 at -0324

most people—certainly most parents—would have to this revelation. Shock and outrage would be more appropriate. The tepid response may explain why in spite of there being 38 times as many victims (of IIC) on Instagram as on Facebook, CEI deletes on Instagram were 21 times lower and IIC auto-enforcing on Instagram was 90 times lower compared to Facebook.<sup>499</sup>

502. In January 2021, [REDACTED] (IG Well-Being Research Lead), [REDACTED] (IG Well-Being Engineering Lead), and Miki Rothschild (IG Well-Being Product Management Lead) discussed how the delay in launching private by default close to a year earlier resulted in Instagram getting beat by TikTok on this exact safety feature:



[REDACTED] (1/13/2021 07:21:20 PST):  
>@slient tiktok did private by default for teens <https://techcrunch.com/2021/01/13/tiktok-update-will-change-privacy-settings-and-defaults-for-users-under-18/>

[REDACTED] (1/13/2021 07:49:15 PST):  
>it's a little embarrassing. Faster on donations, now beating us to this too.

[REDACTED] (1/13/2021 07:49:48 PST):  
>Maybe their moves will put some pressure on the parts of this company that are still Growth-first. 🙄

Michael Yehuda Rothschild (1/13/2021 07:51:57 PST):  
>Well, we could have launched this one many months ago if we didn't care about growth. I do think we landed on an equally positive outcome for teens and better outcome for the company, so perhaps that was a good decision. Only time will tell.

*Document 173: META3047MDL-014-00351807, -1807*

As Mr. Rothschild acknowledged, “Well, we could have launched this one many months ago if we didn’t care about growth.”

503. In the above exchange, Mr. Rothschild speculates that the company “landed on an equally positive outcome for teens” by not defaulting teens to private accounts. But the documents tell a different story. An internal Instagram document quantified multiple undesirable experiences in the past 7 days for users ages 13-24 and reported the following:

---

<sup>499</sup> META3047MDL-031-00192305, -2307

	I know them offline/ in real life	I know them, but only online	I don't know them
Negative comparison	15.5%	19.6%	64.9%
Bullying target	14.6%	16.0%	69.4%
Impersonation	9.9%	21.1%	69.0%
Drugs	9.7%	12.3%	78.0%
Self harm	9.5%	19.2%	71.3%
Political posts	8.9%	16.2%	75.0%
Bullying witness	6.4%	17.6%	76.0%
Violence	5.1%	13.7%	81.2%
Hate witness	4.6%	14.3%	81.1%
Misinfo	4.5%	9.7%	85.8%
Nudity	3.7%	9.1%	87.1%
Unwanted sexual advances	2.1%	4.2%	93.8%

*Document 174: Alison Lee Deposition Exhibit 25 at Slide 13.*

The above columns show that the vast majority of bad experiences come from complete strangers. In his deposition, when asked if setting the default to private would prevent teens from being “groomed” by a sexual predator for abuse, Kilstein responded, “Well, I don’t know if it would prevent all the cases, but it would prevent some cases, yes.”<sup>500</sup> But that understates the issue; the data above indicate that fully 94% of unwanted sexual advances could have been prevented through this feature change. Kilstein himself acknowledged (in internal documents) that this was the point of the private by default concept:

---

<sup>500</sup> Darius Kilstein Dep. Tr. at 382:25-383:2

```
Darius Kilstein (8/21/2020 11:04:12 PDT):  
>What they really want is to fix direct  
  
Darius Kilstein (8/21/2020 11:04:16 PDT):  
>because strangers reach out  
  
Darius Kilstein (8/21/2020 11:04:20 PDT):  
>so they should just do that  
  
Darius Kilstein (8/21/2020 11:04:24 PDT):  
>thats what im trying to tell them  
  
Darius Kilstein (8/21/2020 11:04:32 PDT):  
>stop creeps reaching out to kids in direct  
  
Darius Kilstein (8/21/2020 11:04:33 PDT):  
>its simple
```

*Document 175: Darius Kilstein Dep. Exhibit 31 at -8210.*

Again, “strangers reaching out” is no idle concern, especially given that fully half of message requests to teen users are from non-teens.<sup>501</sup> While Facebook messenger was “locked down” for minors, DM’ing at Instagram was not,<sup>502</sup> which resulted in 38 times as many inappropriate or illicit contact with children on Instagram as on Facebook.<sup>503</sup>

504. A 2020 study conducted by Thorn and produced (for me) via discovery surveyed 1,000 minors ages 9-17 about their online sexual interactions across multiple online platforms and reported the following:

---

<sup>501</sup> See Diego Castaneda Dep. Tr. at 280 (“50 percent of message requests to predicted teen DAU are from predicted non-teens”).

<sup>502</sup> Vaishnavi Jayakumar Dep. Tr. at 453:1-5

<sup>503</sup> META3047MDL-031-00192307



	Someone I believed was 18 or older... [Left columns indicated in white]						Someone I believed was younger than 18... [Right columns indicated in grey]											
Potentially harmful online experience (any selection)	38%	36%	38%	40%	35%	31%	36%	37%	34%	38%	37%	35%	39%	36%	41%	41%	32%	28%
Online sexual interaction (any selection)	25%	23%	26%	24%	22%	21%	19%	19%	16%	19%	21%	19%	29%	27%	34%	28%	23%	23%
...asked me to send a nude photo or video	12%	11%	15%	15%	7%	7%	8%	8%	7%	10%	8%	7%	15%	14%	22%	19%	7%	6%
...asked me to go 'on cam' with a nude or sexually explicit stream	9%	8%	10%	9%	6%	6%	5%	6%	6%	8%	3%	3%	12%	9%	13%	9%	9%	8%
...shared a nude photo or video of themselves with me	12%	10%	14%	10%	9%	10%	8%	7%	7%	6%	8%	8%	15%	13%	20%	13%	9%	12%
...shared a nude photo or video of another kid with me	6%	7%	6%	7%	6%	6%	5%	5%	3%	8%	6%	3%	7%	8%	8%	7%	5%	8%
...sent me sexual messages	16%	13%	19%	16%	11%	9%	10%	9%	10%	10%	9%	6%	20%	17%	26%	20%	12%	11%
...bullied me	15%	17%	15%	22%	15%	11%	16%	19%	15%	25%	18%	14%	15%	15%	15%	19%	13%	8%
...made me feel uncomfortable	19%	17%	23%	19%	13%	13%	14%	14%	14%	11%	15%	17%	22%	18%	30%	26%	11%	10%
[Q19] Please indicate if you have ever experienced any of the following on a platform? Please select all that apply and remember that your answers are anonymous.	ALL MINORS		GIRLS		BOYS		ALL 9-12 YEAR OLDS		GIRLS		BOYS		ALL 13-17 YEAR OLDS		GIRLS		BOYS	
	ALL AGES						AGES 9-12						AGES 13-17					

Document 177: META3047MDL-003-00186841, -6850

In addition, the Thorn report found that LGBTQ+ youth were at considerably increased risk for all such unwanted interactions: 47% overall vs. 35% for non-LGBTQ+.

507. More concerning still is how few minors reported that they turned to anyone for support. Overall, only 6% of 9–17-year-olds turned to a parent/caregiver/trusted adult after receiving a nude by someone they thought was an adult. When asked why they did not seek help, the reasons reported were:

<b>% said reason why they turned to 'no one' for support</b>	...a potentially harmful online experience (any selection)	...an online sexual interaction (any selection)*
You felt this was not a big deal	62%	49%
You worried your report would not be anonymous	24%	23%
You felt embarrassed and worried of being judged	23%	17%
You worried about being in trouble with your family	19%	26%
You worried about police or law enforcement getting involved	15%	19%
You worried about not being allowed to use a platform	15%	19%
You felt like you were to blame or partly to blame for the situation	15%	12%
You worried about losing friends	8%	9%
You worried about being in trouble with your school	5%	6%
You worried about getting bullied at school	5%	2%

[Q34] Which of the following describe why you did not look for additional support following experiences you have had on a platform?

\*Base size is < 100; Columns will total more than 100 because question was select multiple

*Document 178: META3047MDL-003-00186841, -6858*

508. The low percentage of children who sought help and the reasons why they fail to do so highlight the importance of there being structural, anonymous, and effective protective features within the platforms to empower them to be safe since 5-24% of them feel reporting the encounter to others could cause them some source of distress. These features would not be substitutes for a mechanism to inform responsible adults, but they are clearly needed since changing attitudes to make children feel less judged or worry less about getting in trouble (among other things), will prove challenging. In fact, one of the “key insights” of the Thorn report is that “minors are more than twice as likely to use online safety tools to combat potentially harmful online sexual interactions than they are to use offline support systems such as caregivers and parents.”

509. Unfortunately, reporting and blocking, the two most accessible and deployed tools for unwanted contact appear to be minimally effective. Most minors who have either reported or blocked someone say they have been recontacted by that person. Over half of all participants (54%) who had blocked someone they only knew online said they were recontacted by that same person. For those who reported a user they only knew online to the platform, the rate of recontacts was

only slightly lower (51%). But 70% of 9–12-year-old boys and 47% of 9–12-year-old girls who blocked someone they only knew online were recontacted by that person.<sup>504</sup> Not surprisingly then, 41% of minors said they thought “nothing happens if you try and report an inappropriate photo or video to an online platformer app,” and “63% of minors that have shared their own nudes said the same.”<sup>505</sup> I found no evidence that these failures were disclosed to parents or users.

510. Meta’s lackadaisical approach to receiving reports of, and removing, child sexualization content was tested during the COVID pandemic. Maria Lanz, a safety policy representative in sub-Saharan Africa sent an urgent email on March 28, 2020 asking “Is anyone else receiving child abuse reports on platform? In the last 24 hours I’ve receive a ton 😞 I’ve seen terrible things, we’re not detecting it:(((( ”.<sup>506</sup> In response to these reports, Meta’s platform generated a pop up that read the following:

We couldn't review your report. We have fewer people available to review the reports because of the coronavirus (COVID-19) pandemic, so we're only able to review content with the most potential for harm. If you don't want to see amyloves0916 on Instagram, you can unfollow, mute or block them to hide their posts and comments from your feed. Reports like yours are an important part of making Instagram a safe and welcoming place for everyone.<sup>507</sup>

511. According to Jayakumar’s deposition, what was reported was determined by an algorithmic classifier and not the purported victim or a reporter of it.<sup>508</sup> When pressed as to why humans didn’t have the agency to directly report child exploitation or abuse, Ms. ██████████ (IG Product Policy Manager) responded that she’d be “concerned that people would abuse the reporting option to report anything they want reviewed and they would need to review it because

---

<sup>504</sup> META3047MDL-003-00186841, -6881

<sup>505</sup> META3047MDL-003-00186841, -6847

<sup>506</sup> Vaishnavi Jayakumar Dep. Exhibit 29 at -1113

<sup>507</sup> Vaishnavi Jayakumar Dep. Exhibit 30 at -5734

<sup>508</sup> Vaishnavi Jayakumar Dep. Tr. at 271:6-8

it's being reported as CEI [Child Exploitation Imagery]."<sup>509</sup> Putting aside for a moment that abusing the system in this way might be expected to be a rare occurrence, the approach Instagram opted for would be akin to the 911 system not answering phones since sometimes people use it for non-emergencies (locked out of house, cat in tree, long line at drive thru etc.). It's a further indication that child safety was not the number one priority in spite of attestations made to the contrary. As of 2020, the backlog was indeed sizeable. Karina Lynn Newton (Head of IG Public Policy) on April 1 texted:

Karina Lynn Newton (4/01/2020 09:41:42 PDT):  
>Hi all, wanted to take your temperature on if we would ever consider turning off recommended accounts as a break glass/short term measure. It's killing us in this German CEI issue. The underlying issue is we have a 26k backlog of CEI reports, we have 4 FTE reviewing CEI reports right now, and the CI classifiers have very low precision/recall on IG.

*Document 179: META3047MDL-014-00349432, -9432*

512. Compounding Meta's inability to facilitate and act on reports of Child Exploitation Imagery, its algorithms continued to affirmatively push that content onto users. In 2021, Miki Rothschild (IG Well-Being Product Management Lead) opines:

Hey folks, with the recent PR fire TikTok has seen around inappropriate content for teens (see [here](#)), we were chatting within Youth XFN about our own deficits when it comes to discoverability, namely: (1) it's still very easy to find borderline and sometimes violating content/accounts in search, e.g. try typing "drug" or "love" (2) if you start following borderline accounts or interacting with such content, our recommendations algorithms will start pushing you down a rabbit hole of more egregious content.

*Document 180: META3047MDL-003-00077939, -7939*

Rothschild acknowledges that a search beginning with something as potentially innocuous as "love" or "drug" can trigger Meta's algorithms towards showing teens more "egregious" content. In fact, in a notorious, and apparently viral story of Meta's autocomplete logic in March 2018,

---

<sup>509</sup> Vaishnavi Jayakumar Dep. Tr. at 275:6-10

typing in “video of” suggested completing with: “little girl suck,” “giving oral,” “suck dick,” and “minor sex,” among others.<sup>510</sup>

513. In March 2021, close to two years after Hitomi Hayashi-Branson had recommended doing so, Meta issued a press release stating it would “restrict[] DMs between teens and adults they don’t follow” and “encourage[e] teens to make their accounts private.”<sup>511</sup> While a positive step on paper, in execution these measures were incomplete at best. First, this version of private “by default” only applied to new accounts, not existing ones (which did not get the feature until the end of 2024). Second, what was deployed is not what would be considered “private by default” (which entails requiring someone to take active steps to change their account to “public”). Rather, on signing up teens were given the option to select private or public. Jayakumar explained in her deposition, “Generally speaking, most users, not just within Instagram, don't change their default settings until they are in a moment of crisis.”<sup>512</sup> Later, “[s]ix out of ten teens said that they weren't changing their settings. They were just going to go with whatever the app suggested.”<sup>513</sup>

514. Further, Instagram’s March 2021 promise to “restrict[] DMs between teens and adults they don’t follow” was riddled with loopholes that the company does not appear to have warned the public about. As of March 2022, it remained the case that Instagram still allowed senders with stated age 18-20, or no stated age at all, to send DM requests to teens.<sup>514</sup> It also remained the case that Instagram allowed senders outside the U.S. who claimed to be teens (but weren’t) to send such DM requests.<sup>515</sup> As a consequence of these gaps, teens were 50% more likely

---

<sup>510</sup> META3047MDL-014-00346869, -6869

<sup>511</sup> Diego Castaneda Dep. Ex. 18 at 2, 6.

<sup>512</sup> Vaishnavi Jayakumar Dep. Tr. at 72:14-18

<sup>513</sup> Vaishnavi Jayakumar Dep. Tr. at 74:15-19

<sup>514</sup> Vaishnavi Jayakumar Dep. Ex. 71 at 15

<sup>515</sup> Vaishnavi Jayakumar Dep. Ex. 71 at 15

than non-teens to receive a DM request and twice as likely than non-teens to receive an *unwanted* DM request—causing one researcher to conclude “previous effort[s] to block non-teens from sending DM requests to teens haven’t solved teen problems.”<sup>516</sup> Another internal report put the matter even more bluntly: “Teens receiving DM requests from unconnected adults [is] breaking the public commitment that we made.”<sup>517</sup> Meta does not appear to have closed all these loopholes until September 2024, with the launch of Instagram Teen Accounts—over five years after smart defaults were first proposed.<sup>518</sup> Despite the argument that restricting access to teen accounts might even reduce suicides, since unrestricted access facilitates sextortion, cyberbullying, etc., Meta elected to allow teen users to remain reachable by many adult strangers until it launched “Teen Accounts.”<sup>519</sup>

515. The foregoing history shows how Meta has taken a reactive rather than proactive approach to child safety issues, to the detriment of its users. Indeed, internal Meta documents acknowledge as much, such as this slide deck from 2021:

---

<sup>516</sup> Vaishnavi Jayakumar Dep. Ex. 71 at 11-12

<sup>517</sup> META3047MDL-046-00495408, 5408

<sup>518</sup> Diego Castaneda Dep. Tr. at 301:14-304:24.

<sup>519</sup> *Introducing Instagram Teen Accounts: Built-In Protections for Teens, Peace of Mind for Parents*, INSTAGRAM (Sep. 17, 2024), <https://about.instagram.com/blog/announcements/instagram-teen-accounts> (“Teens will be placed in the strictest messaging settings, so they can only be messaged by people they follow or are already connected to.”).

## We rely on ‘after the fact’ enforcement using message content, and have historically prioritised FB & MSGR over IG

1. Our current child safety playbook is optimized for where our technology is strongest, and for meeting our legal obligations - when we find violating content or behaviour, we enforce against the violator and escalate to NCMEC (where required). This is ‘after-the-fact’ enforcement, where we take action once we have confirmed a minor has been harmed (either directly or through revictimisation).

*Document 181: META3047MDL-003-00029989, -996*

Another internal memo similarly stated that “leadership was unwilling to prioritize mitigations” that would reduce unwanted CEI sharing.<sup>520</sup> Repeatedly throughout this memo, the need for additional resources to support safety efforts is called out (e.g. p. 29, 30, 31, 32, 34 among others).

516. A chat from the same year between Alison Lee and [REDACTED] (IG User Experience Researcher), excerpted below, explains the dynamics further. (Meta has a priority system with p0 being emergent, and p1, p2 following in terms of urgency.)

```
Alison Lee (10/08/2021 09:31:20 PDT):
>my team is actually using this framework already! we have our p0/p1/p2s with the expectations that anything past p0s are "like to have"
goals. the challenge with that is that our p0s are often solving the "worst of the worst" foundational problems, and our p1s and p2s are
often things that actually IMPROVE our systems (i.e., creating more precise classifiers to reduce over-enforcement). we often don't have
bandwidth to do more ambitious work that could make our systems better, and in some cases we've had to straight up say no to even putting
that work on our workstream because our PM had to put her foot down to protect our team (like this work on comments integrity
https://fb.workplace.com/notes/314770766810626)

Alison Lee (10/08/2021 09:32:22 PDT):
>not to mention there's always a sense of urgency/importance to work in integrity/wellbeing because our p1s aren't "improve X engagement"
(not to knock engagement/prod teams), but are often things like "examine and mitigate the fairness issues in the hate speech classifier"

Alison Lee (10/08/2021 09:32:47 PDT):
>and our p1s and p2s are what are the subject of public scrutiny/critique

Alison Lee (10/08/2021 09:35:30 PDT):
>i think rethinking the goaling strategy can help mitigate burnout on the teams but doesn't solve the larger questions around under-
resourcing / under-prioritization of this work by the larger org/leadership

[REDACTED] (10/08/2021 10:13:19 PDT):
>we were pushed as a team to consider p0 as must meet and everything else as aspirational. Sam Parker really encouraged us to keep this
number low (i.e.1) and even went so far as to encourage us to consider sometimes having no p0 at all, if there really were no imperatives.
```

*Document 182: Alison Lee Deposition Exhibit 14, -2150*

517. Meta’s “reactive” approach—waiting for harm to occur, rather than trying to prevent it (by leaving as “‘like to have’ goals” the “things that actually IMPROVE our systems”) is akin

<sup>520</sup> META3047MDL-004-00027423, -7425

to summer camps not doing background checks on counselors and waiting instead to see how they act at camp. Slide 14 in the 2021 slide deck excerpted above states the following, “Instead of whack-a-moling [sic] abuse when we find it, we need to pivot our resources to stopping it happening in the first place.” Later in the same presentation it states:

3. We’re underinvested in minor sexualisation on IG, notably on sexualised comments on content posted by minors. Not only is this a terrible experience for creators and bystanders, it’s also a vector for bad actors to identify and connect with one another.

*Document 183: META3047MDL-003-00029989, -9996*

518. The fact that Meta was still adopting a “whack a mole” approach in 2021 is a bit surprising given Sheryl Sandberg’s “Urgent” email in the wake of a story in the Sunday Times about 30 families claiming social media killed their children. In it, she says the following:

If these are the right things to do, we should not need an article to push us to do them. We have reviewed our policies and enforcement in areas like this over and over with this in mind—and then we always find more to do. I am really alarmed by this. We absolutely have to solve this problem. In obvious areas of concern (guns, opiots [sic], harm, etc) we should be ahead of all of this.<sup>521</sup>

The truth appears to be closer to what others lower in Meta’s ranks acknowledged privately.

In Kilstein’s opinion, “[It] was never about increasing safety. It was all for the PR wins.”<sup>522</sup>

This is reiterated in a subsequent exchange between Michale Kane, a data engineer, and James Holland, a data scientist (both are at Instagram):

```
(7/29/2021 15:22:45 PDT):  
> Thanks both. I'm still kinda confused about the goal. We want to make sure teens don't accidentally share more than they want, and avoid creepers, but then some of them totally want that and then if we boost them, we are inviting people to see more teens in their suggestions. Is it about decoupling clear privacy control from integrity issues, or is it about saying we want to reduce integrity problems, but not wanting the consequence of less interactions and follows?
```

*Document 184: Darius Kilstein Dep. Exhibit 37 at -6693*

<sup>521</sup> META3047MDL-004-00025094, -5094

<sup>522</sup> Darius Kilstein Dep. Exhibit 30 at -8138

To which James Holland Replies:

>It is about looking good to regulators so that they don't block our under 13 year old IG version we are working on. Thats it. It has a terrible impact on teen engagement and retention and no detectable benefit on integrity metrics

*Document 185: Darius Kilstein Dep. Exhibit 37 at -6694*

In the end, it is not clear how serious Meta's efforts were when it came to keeping children safe from adult predators on Instagram in particular.

519. I turn now, briefly, to similar issues as they surfaced on other defendant platforms. For its part, TikTok documents recognize that DMs to children are an important part of the strategy child predators deploy:

- Typical behaviour path between a predator and a minor: comment to praise the author → sending follow invite → DM the minor several times → minor responded → build up connection
- The key part is DM. Predators have much higher frequency to DM to minors(over 10x, especially to l1) than normal users, and the entry point is different - Predator tend to send DM through message notification page (65%) while normal users are more like finding videos from hot homepage and message users then.

*Document 186: TIKTOK3047MDL-002-00094384, -4400*

By their own estimates, these "key" DMs are common:

- There are 10K~20K DM report from minors per day globally
  - 60% of DM reports from minors are to non-minors, 40% minors reports are to minors.
- Need Help:

*Document 187: TIKTOK3047MDL-002-00094384, -4392*

520. As noted above, in January 2021, TikTok did turn its under 17-year-old accounts to private by default.<sup>523</sup> This was done as part of a concerted public relations campaign. Complete with queued up complementary quotes from the president of the National PTA, and the CEO of the Family Online Safety Institute, it was touted as an “industry leading initiative, and as a “proactive change rather than as a result of regulatory or media pressure.”<sup>524</sup> Importantly, however, this change was made in the context of TikTok continuing to allow people to self-report age (TikTok’s lax age verification is discussed in more detail below, at Section XII.B.(ii).) The consequences were predictable. Mathew Tenenbaum (Senior Product Manager) writing to James Cummings (Senior Product Manager) on May 5, 2023 said, “The weird data point is that 85% of ‘teen’ users are age gated 18+.”<sup>525</sup> In other words, private by default impacted fewer than 15% of teen users.

521. Despite launching a version of private by default before Instagram, by its own internal assessment, TikTok lagged behind the industry on minor safety issues.

---

<sup>523</sup> TIKTOK3047MDL-002-00119426, -9426

<sup>524</sup> TIKTOK3047MDL-002-00119426, -9426

<sup>525</sup> TIKTOK3047MDL-067-LARK-01022641, -2641

- **US**
- When comparing the perceived performance of TikTok and competitor platforms on **protecting minor safety**, TikTok is severely underperforming by 26% points.
- A possible source for poor user perception can be attributed to **user-to-user interactions** (DM, comments etc) whereby 4 in 10 users claimed to have encountered inappropriate content related to minor safety.
- Beyond that, TikTok also performed worse than competitor platforms on (i) spamming, impersonation and misinformation (-20% points), (ii) ANSA (-12% points), (iii) harassment and bullying (-10% points), (iv) hate speech (-6% points) and (v) dangerous individual/org (- 4% points).

*Document 188: TIKTOK3047MDL-002-00102517, -2527*

522. Finally, I have seen indication in internal company documents that YouTube has had difficulty successfully addressing child exploitation issues on its platform. As of an August 2021 YouTube presentation by [REDACTED] (Software Engineering Manager), “**20%** of YouTube users uploading shorts are unsupervised minors (**1.8M users/week**).”<sup>526</sup> This poses real safety risks, as YouTube’s “Trust and Safety” Division’s internal investigation revealed:

---

<sup>526</sup> GOOG-3047MDL-01262144 at Slide 3 (emphasis in original)

## Abuse is not decreasing

We have taken down more than 2M videos since the start of Greenlight, but kids continue to upload videos of themselves in risky situations

We've restricted live access to 73K accounts but enforcement actions are still holding steady w/w. Recent live data shows account circumvention is happening (10% of actions from accounts created after 12/22)

This tells us our current deterrence systems are not working.

Confidential & Proprietary

Trust & Safety

Document 189: GOOG-3047MDL-00246776 at Slide 15

## Bad actors continue to evade detection

Predatory Account: **robertalexander ramosmedina**  
 38yo Male, joined June 7, 2018  
 1 TOU Strike - Account suspended Sept 14 after investigation

- Uploaded 1 video on June 28, 2018 - removed on July 10, 2018 for child safety and issued 1 TOU strike
- Commented 24 times on 13 videos in 2 months asking to exchange videos - 3 flagged as Minor Sexualization and Vulgar/Lewd from CSAI Comments queue
- 1 Unlisted playlist titled "Naked Children" of 539 videos of young children created June 7 and last updated on Sept. 14.

Chasing entities one by one allows these users to abuse the platform for longer periods of time

Confidential & Proprietary

Trust & Safety

Document 190: GOOG-3047MDL-00246776 at Slide 14

523. In acknowledging that “bad actors continue to evade detection, YouTube gave the following case study of “robertalexander.”<sup>527</sup> It illustrates what appears to be a “three strikes” protocol whereby individuals are given two warning before an action is taken. The posted videos are clearly intended to sexualize children. What is more, this predatory account blatantly offers to exchange videos on the site.

524. I have used YouTube both personally and as an academic; I have worked with the American Academy of Pediatrics on media guidelines and counseled parents regarding use of YouTube (and other SM) with their children. Despite this prior experience, I was unaware that YouTube was being used to exchange sexualized images of children. Nor have I seen any evidence that the pediatric medical community or parents and families we serve are alerted to this danger.

525. Perhaps the reason these risks to children are not recognized is in slide 25 of this slide deck which says that educating families about responsible behavior risks giving the public the “perception that YT is unsafe.”<sup>528</sup> Underage users are an important market driver per YouTube as “these kids are our future user base so we should not isolate them.”<sup>529</sup> This acknowledgement is in direct tension that isolating kids to kids spaces may very well be in their best interest.

526. Furthermore, internal YouTube documents recognize that they have “150 million tweens on YouTube creating via adult accounts with very little prevention-based responsibility measures (e.g. any creep can engage with them in comments and punitive policies that tween don’t understand today.”<sup>530</sup> And while YouTube is terminating some of these accounts, it’s recognized

---

<sup>527</sup> GOOG-3047MDL-00246776 at Slide 14.

<sup>528</sup> GOOG-347MDL-00246776 at Slide 25.

<sup>529</sup> GOOG-347MDL-00246776 at Slide 17.

<sup>530</sup> GOOG-3047MDL-04366467 Slide 1.

that the “current strategy of terminating these accounts (~300K/week) will lead to more brand favorability in this user group.”<sup>531</sup>

527. The net result of the Defendants platforms’ failure to act, failure to alert, and failure to impose structural barriers to using social media platforms as vehicles for exploitation of children is as tragic as it is predictable. A recent metanalysis of online child sexual exploitation and abuse (OCSEA) synthesized data from 123 studies. It defined OCSEA as unwanted, forced, or non-consensual exposure to technology-facilitated abuse and reported the following global prevalences:

**Figure 59: Effect Estimates from Meta-Analysis of the OCSEA Prevalence and Heterogeneity**<sup>532</sup>

	Prevalence estimates				Heterogeneity		
	Studies (n)	Incidents (n)	Observations (n)	Pooled prevalence (95% CI)	$\tau^2$	95% prediction interval	$I^2$
<b>Past year recall</b>							
Online solicitation	50	13 364	112 852	12.5 (10.5-14.7)	0.48	3.4-36.6	99.0%
Non-consensual taking, sharing, and exposure to sexual images and videos	73	17 717	146 868	12.6 (9.7-16.2)	1.64	1.1-65.2	99.5%
Online sexual exploitation	15	1277	21 155	4.7 (2.9-7.3)	0.87	0.6-28.1	93.7%
Sexual extortion	12	652	12 552	3.5 (1.9-6.4)	1.21	0.3-31.9	94.6%
<b>Lifetime recall (childhood)</b>							
Online solicitation	22	25 732	136 331	11.5 (7.2-18.0)	1.55	0.9-64.9	99.6%
Non-consensual taking, sharing, and exposure to sexual images and videos	20	2873	78 819	4.0 (2.3-6.9)	1.78	0.2-42.2	99.4%
Online sexual exploitation	3	480	5894	7.3 (2.2-21.8)	1.23	0.0-100.0	99.1%
Sexual extortion	5	612	11 862	5.1 (4.0-7.2)	0.16	1.3-17.8	93.1%

OCSEA=Online Child Sexual Exploitation and Abuse.

**Table 2: Effect estimates from meta-analysis of the OCSEA prevalence and heterogeneity, by subtype and recall period**

528. Not surprisingly, there is considerable heterogeneity to the results given the global scope, the varied platforms and sampling methods, the recall frames, and the ages of the victims,

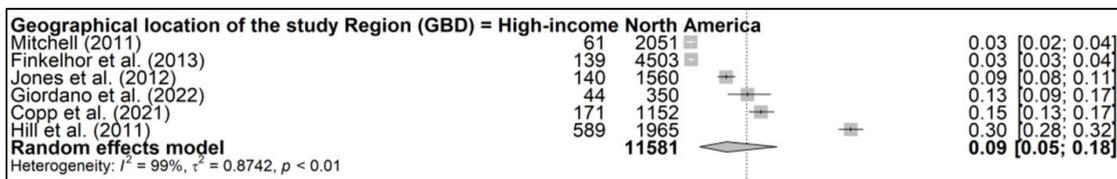
<sup>531</sup> GOOG-3047MDL-04366467 at Slide 1.

<sup>532</sup> Fry D, Krzeczowska A, Ren J, et al. Prevalence estimates and nature of online child sexual exploitation and abuse: a systematic review and meta-analysis. *The Lancet Child & Adolescent Health*. doi:10.1016/S2352-4642(24)00329-8.

so the “pooled” estimate may not be entirely accurate. But no matter, the 95% confidence interval accurately and conservatively represents the range of “true” values and even at the “low” end, the estimates range from 1.9% (sexual extortion) to 10.5% (online solicitation) of teens receiving OSCEA in the past year.

529. Focusing for a moment on North America alone where many of the victims in these cases or their families live, the following is Figure 10 from the appendix to the article.

**Figure 60: Past year experience of online solicitation in all respondents<sup>533</sup>**



The best estimate, derived from six studies, is that 9% of underage children in North American report unwanted online solicitation in the past year. That prevalence merits restating in a different format: **1 in 11 children in North America have experienced unwanted online solicitation annually.** In comparison, the most common chronic disease of childhood is asthma which affects about 1 in 12 children.

530. Again, the platforms were aware of this problem. In closing, consider a 2018 Meta document containing well-being “Highlights”: “Highly sexual content (including N/P, solicitation, CEI/IIC) is a big problem on IG both in terms of reach and intensity. Solicitation has the third highest reach among violating content types in terms of DAP exposed and the fifth highest report

<sup>533</sup> Fry D, Krzeczowska A, Ren J, et al. Prevalence estimates and nature of online child sexual exploitation and abuse: a systematic review and meta-analysis. *The Lancet Child & Adolescent Health*. doi:10.1016/S2352-4642(24)00329-8.

rate,” and teens are at greater risk than non-teens.<sup>534</sup> It goes on to say, “3% of searches on Instagram result in a violating entity.”<sup>535</sup> These are troubling and avoidable statistics.

## **XII. Selected High Profile Dissenting Studies**

531. As discussed in the sections above, this report prioritizes systematic reviews and meta-analyses (the top of the pyramid in Figure 1) over individual studies and experimental or longitudinal studies over cross-sectional ones. It does so because of the sheer volume of studies, the heterogeneity of methods and populations, and the at times conflicting results. That said, several individual studies because of their size, scope, or novelty have garnered considerable attention in both the press and in academic circles and present arguments that countervail my findings. I will selectively review a few of them here and place them into the larger context of the report.

### **J. Orben and Przybylski (2019)<sup>536</sup>**

532. Briefly, this paper used three large existing publicly available data sets, Monitoring the Future (MTF), Youth Risk and Behavior Survey (YRBS) and the Millenium Cohort Study (MCS) to look at the associations between a number of variables including “technology use” and wellbeing. Technology use was defined based on existing variables within each database but included summations of “TV use,” “mobile phone use,” “electronic device use,” “computer use,” and “internet use,” among others. They found that although there were small, negative associations between “technology use” and diminished wellbeing the size of the association was smaller than others including substance use, bullying, sleep, fruit consumption, and about the same size as

---

<sup>534</sup> META3047MDL-031-00048769, -8769

<sup>535</sup> META3047MDL-031-00048769, -8769

<sup>536</sup> Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. *Nature Human Behaviour*. 2019/02/01 2019;3(2):173-182. doi:10.1038/s41562-018-0506-1.

eating “potatoes.” In that context, they conclude that “the outsized weight given to digital screen time in the scientific and public discourse might not be merited.”<sup>537</sup>

533. There are several notable limitations of the analyses, some of which the authors themselves acknowledge:

534. First, the study is cross-sectional and accordingly cannot draw causal inferences.

535. Second, the measure of technology use is heterogenous, blunt, and dated. For example, “computer use” includes doing homework or video chatting neither of which have been implicated either theoretically or empirically in wellbeing effects. Mobile phone use (the term itself is dated) includes talking which again is not viewed as being harmful. Including variables that are unlikely to have associations, dilutes those that do or might and biases findings towards the null.

536. Third, all of these measures relied on self-report of media usage which correlates only weakly to moderately ( $r=.38$ ) with actual usage.<sup>538</sup>

537. Fourth, some of the studies collected data from as far back as 2007 which explains why the surveys included questions about “cell phones” instead of “smartphones,” and “television” which is no longer a predominate media. The media landscape has evolved considerably since then. In fact, teen usage of Facebook peaked in 2014-15 a full 7 years after initial data collection.

538. Fifth, they treat mediators as confounders in their analyses. Recall our prior exegesis on this distinction. A mediator is in the causal pathway and should not be adjusted for

---

<sup>537</sup> Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. *Nature Human Behaviour*. 2019/02/01 2019;3(2):173-182. doi:10.1038/s41562-018-0506-1.

<sup>538</sup> Parry DA, Davidson BI, Sewall CJR, Fisher JT, Mieczkowski H, Quintana DS. A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*. 2021/11/01 2021;5(11):1535-1547. doi:10.1038/s41562-021-01117-5.

but rather examined as a means of explaining a mechanism that links an exposure to an outcome (high paying job in our example linking college with subsequent wealth). Orben and Przybylski adjust for such things as negative attitudes towards school and time spent with parents (among other things). These could very plausibly be in the causal pathway between time on SM and wellbeing. In fact, Kelly et al, using one of the same databases that Orben used, the Millennial Cohort Study, looked at social media use and subsequent depression and found starkly different results. Their primary findings are presented in below extracted from their paper.<sup>539</sup>

**Figure 61: Summary of Multivariable Regressions of Depressive Symptom Scores by Social Media Use<sup>540</sup>**

Multivariable regressions, depressive symptom scores by social media use.

	Model 0 (M0)	Model 1: M0 + online harassment	Model 2: M0 + sleep	Model 3: M0 + self-esteem	Model 4: M0 + body image					
Social media use in hours/weekday										
Panel A: girls (n = 5496)										
None	0.74***	(0.62 to 0.89)	0.84*	(0.71 to 0.99)	0.87	(0.74 to 1.02)	0.77**	(0.65 to 0.91)	0.88	(0.76 to 1.01)
<1 h	0.88**	(0.80 to 0.96)	0.94	(0.86 to 1.02)	0.93	(0.86 to 1.00)	0.87***	(0.80 to 0.95)	0.96	(0.89 to 1.03)
1 to <3 h (ref)										
3 to <5 h	1.26***	(1.15 to 1.37)	1.17***	(1.08 to 1.26)	1.18***	(1.09 to 1.28)	1.20***	(1.10 to 1.30)	1.17***	(1.08 to 1.26)
>5 h	1.50***	(1.39 to 1.62)	1.30***	(1.21 to 1.40)	1.28***	(1.19 to 1.38)	1.26***	(1.17 to 1.35)	1.30***	(1.21 to 1.40)
Wald test, F(4,387)	48	P < 0.00005	21	P < 0.00005	22	P < 0.00005	31	P < 0.00005	21	P < 0.00005
Panel B: boys (n = 5408)										
None	1.01	(0.91 to 1.11)	1.11*	(1.01 to 1.23)	1.06	(0.97 to 1.16)	0.98	(0.89 to 1.08)	1.10*	(1.01 to 1.20)
<1 h	0.99	(0.92 to 1.07)	1.03	(0.95 to 1.11)	1.01	(0.94 to 1.09)	0.99	(0.92 to 1.07)	1.01	(0.94 to 1.09)
1 to <3 h (ref)										
3 to <5 h	1.21***	(1.08 to 1.35)	1.16**	(1.04 to 1.30)	1.15*	(1.03 to 1.27)	1.18**	(1.06 to 1.32)	1.17**	(1.05 to 1.31)
>5 h	1.35***	(1.23 to 1.50)	1.27***	(1.15 to 1.39)	1.21**	(1.10 to 1.34)	1.31***	(1.18 to 1.44)	1.30***	(1.19 to 1.42)
Wald test, F(4,387)	13	P < 0.00005	8	P < 0.00005	5	P = 0.007	11	P < 0.00005	10	P < 0.00005

Notes: All regressions adjust for covariates: family income and structure at age 14, internalising scores at age 11, and age and are weighted with sample weights. Confidence intervals are in parentheses. Sample sizes are unweighted. Regression coefficients have been exponentiated to aid interpretation.

\* p < 0.05.  
 \*\* p < 0.01.  
 \*\*\* p < 0.001.

539. In their analysis, using 1- 3 hours of usage at baseline as a comparator (Model 0 above), increased social media usage (3 to <5 and > 5 hours per day) was associated with a 21-

<sup>539</sup> Kelly Y, Zilanawala A, Booker C, Sacker A. Social Media Use and Adolescent Mental Health: Findings From the UK Millennium Cohort Study. *EClinicalMedicine*. Dec 2018;6:59-68. doi:10.1016/j.eclinm.2018.12.005

<sup>540</sup> Kelly Y, Zilanawala A, Booker C, Sacker A. Social Media Use and Adolescent Mental Health: Findings From the UK Millennium Cohort Study. *EClinicalMedicine*. 2019 Jan 4;6:59-68. doi: 10.1016/j.eclinm.2018.12.005. PMID: 31193561; PMCID: PMC6537508.

50% increased risk of subsequent depression (red squares). Models 1-4 go on to explore the *mechanisms* (sleep, self-esteem, and body image) that might explain this association by adding *mediators* to see if they are significant. They are; see the blue squares in the figure above. Adding “online harassment” (Model 1); “sleep” (Model 2); “self-esteem” (Model 3) and “body image” (Model 4) all attenuated the odds of increased media’s association with depression. In statistical terms, as we explored in section earlier, this means they are in the causal pathway: social media use’s association with depressive symptoms goes *through* each of them. Finally, the technology exposure variable was only measured as duration without respect to content (including SM in particular).

**K. Ferguson (2024)<sup>541</sup>**

540. This meta-analysis concluded that “*meta-analytic evidence for causal effects was statistically no different than zero.*” This meta-analysis claimed to synthesize the existing experimental data linking (or failing to link) SMS usage and mental health. Several things stood out as I read that meta-analysis. First, it had a single author. While not in and of itself dispositive, single authorship is unusual in today’s days of “team science” and especially unusual for systematic reviews and meta-analysis where subjective assessment of content is part of the adjudication process. In fact, most guidelines for how to conduct and publish systematic reviews discuss how consensus should be achieved and reported when authors disagree about relevance, findings etc. Indeed, the ROBIS criteria, widely accepted as one of the sources of best practices state: “To minimize the potential for bias and errors in these processes, titles and abstracts should be screened independently by at least two reviewers and full-text inclusion assessment should

---

<sup>541</sup> Ferguson CJ. Do social media experiments prove a link with mental health: A methodological and meta-analytic review. *Psychology of Popular Media*. 2024:No Pagination Specified-No Pagination Specified. doi:10.1037/ppm0000541

involve at least two reviewers (either independently or with one performing the assessment and the second checking the decision).”<sup>542</sup> Second, the author simply reports that 27 studies were identified as meeting criteria but does not provide a diagram of how many studies the search strategy yielded and how many studies were excluded for what reasons as is customary or systematic review. Third, there was considerable heterogeneity both in terms of approaches, outcomes, and duration in the included studies. Following best practices, heterogeneity is assessed *prior* to any summary estimate, and where it is too great, either statistically or conceptually, summary estimates are not generated. This was not done here although the author notes in the discussion that the summary estimate “*masks considerable heterogeneity between studies.*”

541. My concerns with this particular meta-analysis led me to conduct a deeper dive into it and unsurprisingly I found given its contentiousness that there was considerable conversation in the scientific “twitter sphere” especially since it ran counter to the prevailing scientific and public consensus about the role SM may play in mental health outcomes. Many scientists criticized its findings including David Stein whose substack on it begins:

Ferguson published a ‘review’ that repeatedly but falsely implied that the experiments revealed there were no beneficial [sic] impacts of SM time reductions on depression and anxiety. Ferguson misdirects the public in this manner persistently within his review, and even the title as well as the Keywords: social media, mental health, depression, anxiety information displayed at start of the review do mislead the public to think the review is about impacts on genuine MH disorders like depression and anxiety.<sup>543</sup>

542. Stein goes on to point out multiple errors of omission (missing studies) and commission (incorrect methodologies applied, wrong effect sizes with opposite signage used,

---

<sup>542</sup> Whiting P, Savović J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol.* Jan 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005

<sup>543</sup> [https://shoresofacademia.substack.com?utm\\_source=navbar&utm\\_medium=web](https://shoresofacademia.substack.com?utm_source=navbar&utm_medium=web)

inclusion of studies that do not measure mental health etc.) and concludes: “In short, Ferguson’s paper stands and falls on de facto censorship of evidence.”

543. Both Stein and Thrule go on to reanalyze the data and reach the *opposite* conclusion of Ferguson. I found a pre-print of Thrule’s analysis online and inquired if it had been peer- reviewed and published yet. Here is the reply I received to an email I sent (Nov 13, 2024).

*Hi Dimitri,*

*Thanks for your note. The re-analysis is already in press at Psychology of Popular Media, the same journal that published the Ferguson meta-analysis in the first place. I submitted corrected proofs a couple of weeks ago, so hopefully it will be online soon. I think they are waiting to get a response from Chris Ferguson to publish the re-analysis and his response together.*

*Thanks,  
Johannes*

In particular, the author of this re-analysis found that duration of abstinence from social media matters. This is not surprising. Short term outcomes during a period of abstinence might be worse as someone is going through “active withdrawal” from an addiction. In fact, Thrule et. al. report that: “Stratified analyses indicated that interventions of less than 1 week resulted in significantly worse mental health outcomes ( $d=-0.168$ ,  $SE=0.058$ ,  $p=.004$ ), while interventions of 1 week or longer resulted in significant improvements ( $d=0.169$ ,  $SE=0.065$ ,  $p=.01$ ).”<sup>544</sup> These findings are entirely consistent with what one would expect if studying abstention amongst people with an addiction. In effect then, this metanalysis, despite its conclusion, does more to *affirm* SM addiction than the *refute* it.

---

<sup>544</sup> Thrul J, Devkota J, AlJuboori D, Regan T, Alomairah S, Vidal C. Social media reduction or abstinence interventions are providing mental health benefits – reanalysis of a published meta-analysis. *Psychology of Popular Media*. In press.

544. In spite of its many and considerable limitations — enough that it should not be taken seriously — Ferguson’s study results are summarized below.

**Figure 62: Meta-Analytic Results of Social Media and Mental Health Outcomes<sup>545</sup>**

<b>Table 1</b>						
<i>Meta-Analytic Results of Social Media and Mental Health Outcomes</i>						
Random-effects model ( $k = 27$ )						
Estimate ( $d$ )	$Z$	$p$	CI lower bound	CI upper bound		
0.088	1.63	.104	-0.018	0.197		
Heterogeneity statistics						
Tau	Tau <sup>2</sup>	$I^2$	$H^2$	$df$	$Q$	$p$
0.114	0.013 ( $SE = 0.0053$ )	75.2%	4.025	26.000	91.64	<.001
<i>Note.</i> Tau <sup>2</sup> estimator: maximum-likelihood. For the purposes of analysis in jamovi, values $d$ were converted to $r$ and then converted back again. CI = confidence interval.						

His “flawed” summary estimate of the effect size .088 (small by accepted standards) and his  $p$ -value is .10 which is technically “statistically” insignificant. Recall two things we discussed earlier in this report. First, a  $p$ -value of .10 means that there is a 10% chance that the result is spurious (ie a 90% chance it is the truth). Second, even effect sizes of this magnitude can have large population effects at scale. Ferguson subsequently “amended” and corrected this meta-analysis allegedly in response to these critiques but did not redress them.

<sup>545</sup> Ferguson CJ. Do social media experiments prove a link with mental health: A methodological and meta-analytic review. *Psychology of Popular Media*. 2024:No Pagination Specified-No Pagination Specified. doi:10.1037/ppm0000541.

**L. Ferguson, Kaye, Branley-Bell, Markey (2025)**

545. Ferguson CJ, Kaye LK, Branley-Bell D, Markey P. There is no evidence that time spent on social media is correlated with adolescent mental health problems: Findings from a meta-analysis.<sup>546</sup> The considerable flaws, missteps, and overstatement in the metanalysis by Ferguson discussed in (b) above are enough to call into question his credibility to conduct them but he has done numerous of them, all of which have been heavily criticized.

546. I'll begin my critique of this paper by making a table of the 30 systematic reviews and metanalyses cited in this report. They were selected based on the criteria listed in section IV.B.

I have highlighted the title of this Ferguson one in light blue below:

Do social media experiments prove a link with mental health: A methodological and meta-analytic review. <sup>547</sup>
Alcohol use and risk of suicide: a systematic review and Meta-analysis <sup>548</sup>
Exposure and Risks of Ischemic Heart Disease and Stroke Events: Review and Meta-Analysis <sup>549</sup>
Cyberbullying Perpetration and Victimization in Youth: A Meta-Analysis of Longitudinal Studies. <sup>550</sup>
Are active and passive social media use related to mental health, wellbeing, and social support outcomes? A meta-analysis of 141 studies. <sup>551</sup>

<sup>546</sup> *Professional Psychology: Research and Practice*. 2025;56(1):73-83. doi:10.1037/pro0000589

<sup>547</sup> Ferguson CJ. Do social media experiments prove a link with mental health: A methodological and meta-analytic review. *Psychology of Popular Media*. 2024:No Pagination Specified-No Pagination Specified. doi:10.1037/ppm0000541

<sup>548</sup> Amiri S, Behnezhad S. Alcohol use and risk of suicide: a systematic review and Meta-analysis. *Journal of Addictive Diseases*. 2020/02/17 2020;38(2):200-213. doi:10.1080/10550887.2020.1736757

<sup>549</sup> Alexeeff SE, Deosaransingh K, Van Den Eeden S, Schwartz J, Liao NS, Sidney S. Association of Long-term Exposure to Particulate Air Pollution With Cardiovascular Events in California. *JAMA Network Open*. 2023;6(2):e230561-e230561. doi:10.1001/jamanetworkopen.2023.0561

<sup>550</sup> Marciano L, Schulz PJ, Camerini A-L. Cyberbullying Perpetration and Victimization in Youth: A Meta-Analysis of Longitudinal Studies. *Journal of Computer-Mediated Communication*. 2020;25(2):163-181. doi:10.1093/jcmc/zmz031

<sup>551</sup> Godard R, Holtzman S. Are active and passive social media use related to mental health, wellbeing, and social support outcomes? A meta-analysis of 141 studies. *Journal of Computer-Mediated Communication*. 2024;29(1)doi:10.1093/jcmc/zmad055

Interplay between social media use, sleep quality, and mental health in youth: A systematic review. <sup>552</sup>
Association Between Daily Alcohol Intake and Risk of All-Cause Mortality: A Systematic Review and Meta-analyses <sup>553</sup>
Prevalence of social media addiction across 32 nations: Meta-analysis with subgroup analysis of classification schemes and cultural values. <sup>554</sup>
Psychometric Properties of Screening Instruments for Social Network Use Disorder in Children and Adolescents: A Systematic Review <sup>555</sup>
Problematic Social Media Use in Adolescents and Young Adults: Systematic Review and Meta-analysis <sup>556</sup>
Media and Depression Symptoms: a Meta-Analysis. <sup>557</sup>
Fear of missing out (FoMo) and internet use: A comprehensive systematic review and meta-analysis <sup>558</sup>
Fear of missing out and social networking sites use and abuse: A meta-analysis <sup>559</sup>
Prevalence of body dysmorphic disorder: A systematic review and meta-analysis. <sup>560</sup>
The Relationship Between SNS Usage and Disordered Eating Behaviors: A Meta-Analysis. <sup>561</sup>

<sup>552</sup> Alonzo R, Hussain J, Stranges S, Anderson KK. Interplay between social media use, sleep quality, and mental health in youth: A systematic review. *Sleep Medicine Reviews*. 2021/04/01/ 2021;56:101414. doi:<https://doi.org/10.1016/j.smr.2020.101414>

<sup>553</sup> Zhao J, Stockwell T, Naimi T, Churchill S, Clay J, Sherk A. Association Between Daily Alcohol Intake and Risk of All-Cause Mortality: A Systematic Review and Meta-analyses. *JAMA Network Open*. 2023;6(3):e236185-e236185.

<sup>554</sup> Cheng C, Lau Y-c, Chan L, Luk JW. Prevalence of social media addiction across 32 nations: Meta-analysis with subgroup analysis of classification schemes and cultural values. *Addictive Behaviors*. 2021/06/01/ 2021;117:106845. doi:<https://doi.org/10.1016/j.addbeh.2021.106845>

<sup>555</sup> Schlossarek S, Schmidt H, Bischof A, et al. Psychometric Properties of Screening Instruments for Social Network Use Disorder in Children and Adolescents: A Systematic Review. *JAMA Pediatr*. Apr 1 2023;177(4):419-426. doi:10.1001/jamapediatrics.2022.5741

<sup>556</sup> Shannon H, Bush K, Villeneuve P, Hellemans K, Guimond S. Problematic Social Media Use in Adolescents and Young Adults: Systematic Review and Meta-analysis. *JMIR Ment Health*. 2022;9(4)

<sup>557</sup> Cunningham S, Hudson CC, Harkness K. Social Media and Depression Symptoms: a Meta-Analysis. *Res Child Adolesc Psychopathol*. Feb 2021;49(2):241-253. doi:10.1007/s10802-020-00715-7

<sup>558</sup> Akbari M, Seydavi M, Palmieri S, Mansueto G, Caselli G, Spada MM. Fear of missing out (FoMoFoMoFoMo) and internet use: A comprehensive systematic review and meta-analysis. *Journal of Behavioral Addictions*. 31 Dec. 2021 2021;10(4):879-900. doi:<https://doi.org/10.1556/2006.2021.00083>

<sup>559</sup> Fioravanti G, Casale S, Benucci SB, et al. Fear of missing out and social networking sites use and abuse: A meta-analysis. *Computers in Human Behavior*. 2021/09/01/ 2021;122:106839. doi:<https://doi.org/10.1016/j.chb.2021.106839>

<sup>560</sup> Rief McGrath LR, Oey L, McDonald S, Berle D, Wootton BM. Prevalence of body dysmorphic disorder: A systematic review and meta-analysis. *Body Image*. 2023/09/01/ 2023;46:202-211. doi:<https://doi.org/10.1016/j.bodyim.2023.06.008>

<sup>561</sup> Zhang J, Wang Y, Li Q, Wu C. The Relationship Between SNS Usage and Disordered Eating Behaviors: A Meta-Analysis. *Front Psychol*. 2021;12:641919. doi:10.3389/fpsyg.2021.641919

The use of social networking sites, body image dissatisfaction, and body dysmorphic disorder: A systematic review of psychological research. <sup>562</sup>
A Meta-Analysis of the Effects of Social Media Exposure to Upward Comparison Targets on Self-Evaluations and Emotions. <sup>563</sup>
The role of the media in body image concerns among women: a meta-analysis of experimental and correlational studies. <sup>564</sup>
How the exposure to beauty ideals on social networking sites influences body image: A systematic review of experimental studies. <sup>565</sup>
A meta-analytic review of the relationship between social media use and body image disturbance <sup>566</sup>
The role of media literacy in body dissatisfaction and disordered eating: A systematic review. <sup>567</sup>
A scoping review to investigate the association between social media, body image and eating disorders amongst young people. <sup>568</sup>
Problematic usage of the internet and eating disorder and related psychopathology: A multifaceted, systematic review and meta-analysis <sup>569</sup>
“Using digital media or sleeping ... that is the question”. A meta-analysis on digital media use and unhealthy sleep in adolescence <sup>570</sup>

<sup>562</sup> Ryding FC, Kuss DJ. The use of social networking sites, body image dissatisfaction, and body dysmorphic disorder: A systematic review of psychological research. *Psychology of Popular Media*. 2020;9(4):412-435. doi:10.1037/ppm0000264

<sup>563</sup> McComb C, Vanman E, Tobin S. A Meta-Analysis of the Effects of Social Media Exposure to Upward Comparison Targets on Self-Evaluations and Emotions. *Media Psychology*. 2023;26(5)

<sup>564</sup> Grabe S, Ward LM, Hyde JS. The role of the media in body image concerns among women: a meta-analysis of experimental and correlational studies. *Psychol Bull*. May 2008;134(3):460-76. doi:10.1037/0033-2909.134.3.460

<sup>565</sup> Fioravanti G, Bocci Benucci S, Ceragioli G, Casale S. How the exposure to beauty ideals on social networking sites influences body image: A systematic review of experimental studies. *Adolescent Research Review*. 2022:No Pagination Specified-No Pagination Specified. doi:10.1007/s40894-022-00179-4

<sup>566</sup> Saiphoo AN, Vahedi Z. A meta-analytic review of the relationship between social media use and body image disturbance. *Comput Hum Behav*. 2019;101:259-275.

<sup>567</sup> McLean SA, Paxton SJ, Wertheim EH. The role of media literacy in body dissatisfaction and disordered eating: A systematic review. *Body Image*. Dec 2016;19:9-23. doi:10.1016/j.bodyim.2016.08.002

<sup>568</sup> Dane A, Bhatia K. The social media diet: A scoping review to investigate the association between social media, body image and eating disorders amongst young people. *PLOS Global Public Health*. 2023;3(3):e0001091. doi:10.1371/journal.pgph.0001091

<sup>569</sup> Ioannidis K, Taylor C, Holt L, et al. Problematic usage of the internet and eating disorder and related psychopathology: A multifaceted, systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*. 2021/06/01/ 2021;125:569-581. doi:https://doi.org/10.1016/j.neubiorev.2021.03.005

<sup>570</sup> Pagano M, Bacaro V, Crocetti E. “Using digital media or sleeping ... that is the question”. A meta-analysis on digital media use and unhealthy sleep in adolescence. *Computers in Human Behavior*. 2023/09/01/ 2023;146:107813. doi:https://doi.org/10.1016/j.chb.2023.107813

Interventions to control children's screen use and their effect on sleep: A systematic review and meta-analysis. <sup>571</sup>
School Start Times, Sleep, Behavioral, Health, and Academic Outcomes: A Review of the Literature <sup>572</sup>
Effects of school start time on students' sleep duration, daytime sleepiness, and attendance: a meta-analysis <sup>573</sup>
The use of wearable technology to measure and support abilities, disabilities and functional skills in autistic youth: a scoping review <sup>574</sup>
The relationship between screen time and mental health in young people: A systematic review of longitudinal studies <sup>575</sup>
Impacts of digital social media detox for mental health: A systematic review and meta-analysis <sup>576</sup>
Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: a meta-analysis of longitudinal studies <sup>577</sup>
Association between suicide reporting in the media and suicide: systematic review and meta-analysis <sup>578</sup>
Social media use and self-injurious thoughts and behaviors: A systematic review and meta-analysis <sup>579</sup>

<sup>571</sup> Martin KB, Bednarz JM, Aromataris EC. Interventions to control children's screen use and their effect on sleep: A systematic review and meta-analysis. *Journal of Sleep Research*. 2021;30(3):e13130. doi:<https://doi.org/10.1111/jsr.13130>

<sup>572</sup> Wheaton AG, Chapman DP, Croft JB. School Start Times, Sleep, Behavioral, Health, and Academic Outcomes: A Review of the Literature. *J Sch Health*. May 2016;86(5):363-81. doi:10.1111/josh.12388

<sup>573</sup> Bowers JM, Moyer A. Effects of school start time on students' sleep duration, daytime sleepiness, and attendance: a meta-analysis. *Sleep Health*. Dec 2017;3(6):423-431. doi:10.1016/j.sleh.2017.08.004

<sup>574</sup> Black MH, Milbourn B, Chen NTM, et al. The use of wearable technology to measure and support abilities, disabilities and functional skills in autistic youth: a scoping review. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*. 2020;8(1):48-69. doi:10.21307/sjcapp-2020-006

<sup>575</sup> Tang S, Werner-Seidler A, Torok M, Mackinnon AJ, Christensen H. The relationship between screen time and mental health in young people: A systematic review of longitudinal studies. *Clin Psychol Rev*. Jun 2021;86:102021. doi:10.1016/j.cpr.2021.102021

<sup>576</sup> Ramadhan RN, Rampengan DD, Yumnanisha DA, et al. Impacts of digital social media detox for mental health: A systematic review and meta-analysis. *Narra J*. Aug 2024;4(2):e786. doi:10.52225/narra.v4i2.786

<sup>577</sup> Ribeiro JD, Franklin JC, Fox KR, et al. Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: a meta-analysis of longitudinal studies. *Psychol Med*. Jan 2016;46(2):225-36. doi:10.1017/S0033291715001804

<sup>578</sup> Niederkrotenthaler T, Braun M, Pirkis J, et al. Association between suicide reporting in the media and suicide: systematic review and meta-analysis. *BMJ*. 2020;368:m575. doi:10.1136/bmj.m575

<sup>579</sup> Nesi J, Burke TA, Bettis AH, et al. Social media use and self-injurious thoughts and behaviors: A systematic review and meta-analysis. *Clinical Psychology Review*. 2021/07/01/2021;87:102038. doi:<https://doi.org/10.1016/j.cpr.2021.102038>

Prevalence estimates and nature of online child sexual exploitation and abuse: a systematic review and meta-analysis <sup>580</sup>
Longitudinal associations between digital media use and ADHD symptoms in children and adolescents: a systematic literature review <sup>581</sup>
A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. <sup>582</sup>
There is no evidence that time spent on social media is correlated with adolescent mental health problems: Findings from a meta-analysis <sup>583</sup>

547. A casual read will notice a salient difference in the titles. Some pose questions; many are entirely descriptive; only one is definitively declarative, and negatively so at that. In scientific settings, proving a negative is exceedingly difficult to do as it requires ruling out all possibilities (viz “There are no black swans”). Even the reviews above that *do* find significant associations are more circumspect in their titles. Ferguson’s title is definitive and polemical. Ironically in the introduction to his paper, he states, “Specifically concerning academic debate, the issue of social media use and mental health remains polarized.” And later in the introduction he cites multiple methodological issues with the existing literature including reliance on self-report, absence of content etc. Yet he then goes on to summarize and synthesize those same studies and concludes that there is no evidence. As such, his title is incendiary and designed to draw attention,

<sup>580</sup> Fry D, Krzeczowska A, Ren J, et al. Prevalence estimates and nature of online child sexual exploitation and abuse: a systematic review and meta-analysis. *The Lancet Child & Adolescent Health*. doi:10.1016/S2352-4642(24)00329-8

<sup>581</sup> Thorell LB, Buren J, Strom Wiman J, Sandberg D, Nutley SB. Longitudinal associations between digital media use and ADHD symptoms in children and adolescents: a systematic literature review. *Eur Child Adolesc Psychiatry*. Aug 2024;33(8):2503-2526. doi:10.1007/s00787-022-02130-3

<sup>582</sup> Parry DA, Davidson BI, Sewall CJR, Fisher JT, Mieczkowski H, Quintana DS. A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*. 2021/11/01 2021;5(11):1535-1547. doi:10.1038/s41562-021-01117-5

<sup>583</sup> Ferguson CJ, Kaye LK, Branley-Bell D, Markey P. There is no evidence that time spent on social media is correlated with adolescent mental health problems: Findings from a meta-analysis. *Professional Psychology: Research and Practice*. 2025;56(1):73-83. doi:10.1037/pro0000589

rather than accurately convey the current state of science. What is more, his own results—using “flawed” and “limited” studies—report the following:

**Figure 63: Meta-Analytic Results of Social Media and Mental Health Outcomes<sup>584</sup>**

*Meta-analytic Results of Social Media and Mental Health Outcomes*

Effect size	<i>k</i>	$\beta$	95% CI	Homogeneity test	$I^2$	$\tau$	Publication bias?
All studies	79	.061	[.047, .075]	$X^2(78) = 4404.45, p < .001$	98.8	.055	No
Biological sex							
Male	27	.044	[.025, .062]	$X^2(26) = 164.79, p < .001$	94.7	.040	No
Female	29	.075	[.050, .101]	$X^2(28) = 388.10, p < .001$	97.9	.063	No
Study type							
Correlational	48	.072	[.05, .090]	$X^2(47) = 4167.83, p < .001$	99.3	.057	No
Longitudinal	30	.044	[.023, .066]	$X^2(29) = 169.54, p < .001$	83.6	.049	No
Data set							
Bespoke	21	.044	[.012, .070]	$X^2(12) = 167.61, p < .001$	89.2	.046	No
National survey	53	.067	[.050, .084]	$X^2(52) = 99.21, p < .001$	99.2	.058	Yes
Dissertation	5	.045	[.016, .074]	$X^2(4) = 1.81, p = .770$	57.2	0	No

*Note.* *k* = number of studies;  $\beta$  = pooled effect size estimate; CI = confidence interval;  $I^2$  = heterogeneity statistic.

548. As demonstrated in the red box above, all of his results do in fact show a significant, albeit small, correlation between social media and health outcomes. Given the imprecision of the data, those findings alone could be interpreted as positive. Perhaps mindful of that, in his discussion, Ferguson states, “Overall, our findings indicate that the current research literature is unable to provide *strong* evidence for a clinically relevant link between time spend on social media and mental health issues in youth.” (*emphasis added*). Finally, it is odd that this meta review does not include some experimental and quasi-experimental studies that provide much stronger evidence of a causal linkage. These same studies were omitted from his prior meta-analysis as well.<sup>585</sup>

<sup>584</sup> Ferguson CJ. Do social media experiments prove a link with mental health: A methodological and meta-analytic review. *Psychology of Popular Media*. 2024:No Pagination Specified-No Pagination Specified. doi:10.1037/ppm0000541

<sup>585</sup> Ferguson CJ. Do social media experiments prove a link with mental health: A methodological and meta-analytic review. *Psychology of Popular Media*. 2024:No Pagination Specified-No Pagination Specified. doi:10.1037/ppm0000541; Allcott H, Braghieri L, Eichmeyer S, Gentzkow M. The Welfare Effects of Social Media. *American Economic Review*. 2020;110(3):629–76. doi:10.1257/aer.20190658; Braghieri L, Levy Re, Makarin A. Social Media and Mental Health. *American Economic Review*. 2022;112(11):3660–93. doi:10.1257/aer.20211218

**M. Hancock et al. (2022, last revised 2025)**

549. I was given this one by counsel to review.<sup>586</sup> I had not identified as one to consider including previously because the literature search in it was completed in 2018. Furthermore, it has only, to my knowledge, been “published” on a pre-print server (SSRN) and not in the peer-reviewed literature. Briefly, pre-print servers are a mechanism devised to provide rapid dissemination of findings in an open access platform that allows for commentary. Their intended purpose—and their biggest selling points—are the open and rapid dissemination of new knowledge with an opportunity to incorporate feedback and improve the science prior to peer review which remains the gold standard of quality in biomedical journals. Pre-print servers saw an explosion of submissions during COVID as rapidly proliferating scientific findings were uploaded often as fast as they were completed. Many papers uploaded to preprint servers are never published in peer reviewed journals at all but the theory is that the eventual submission will be improved by providing a period of open vetting. Experts urge caution in accepting the findings of pre-print servers for exactly this reason. This particular paper was first posted to the server in 2022 and updated in Jan of 2025. It has not, to my knowledge, been published in a journal yet. Nevertheless, I reviewed it.

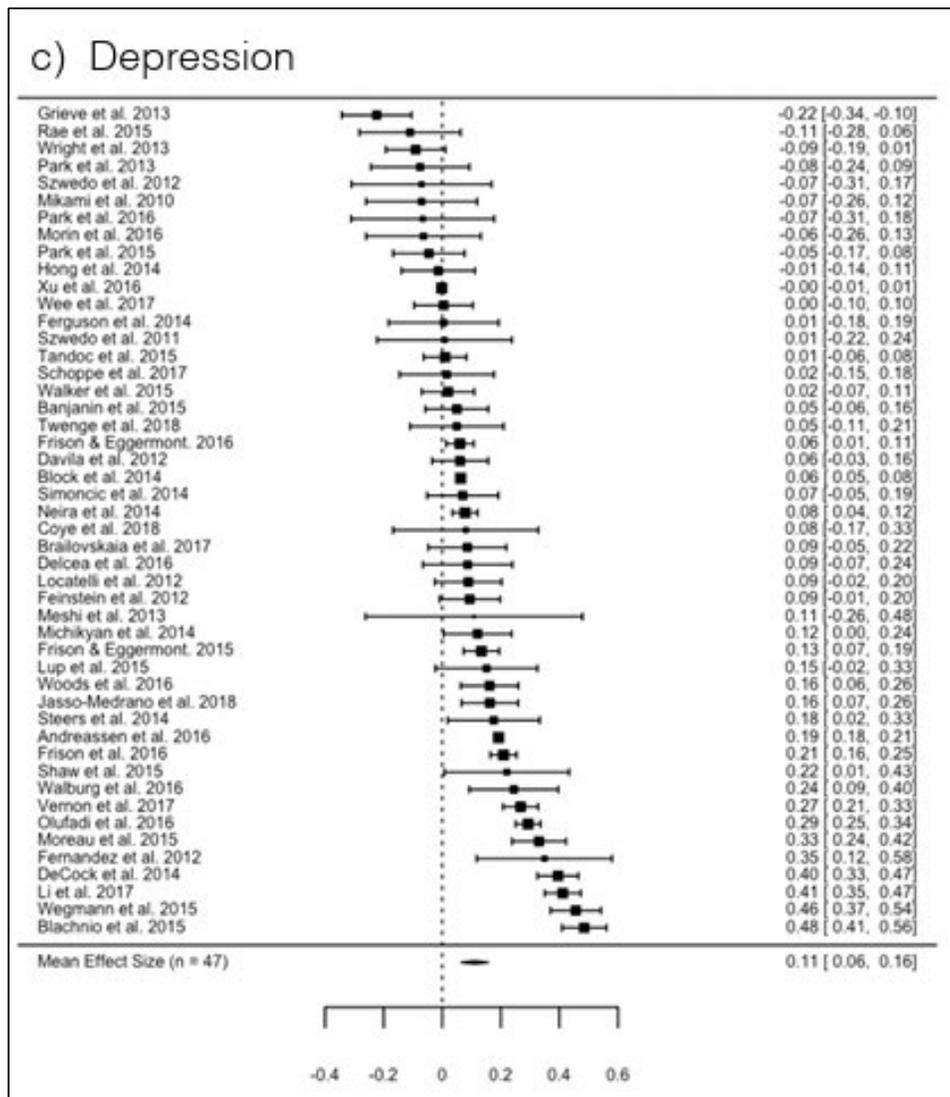
550. There was considerable heterogeneity in the included studies and so the authors appropriately deployed a random effect model as discussed earlier. In their first analysis they included all 226 studies to look at the association of social media use with overall well-being and found no effect. Next, they looked at two outcomes relevant to this report: anxiety and depression.

---

<sup>586</sup> Hancock et al. “Psychological Well-Being and Social Media Use: A Meta-Analysis of Associations between Social Media Use and Depression, Anxiety, Loneliness, Eudaimonic, Hedonic and Social Well-Being”

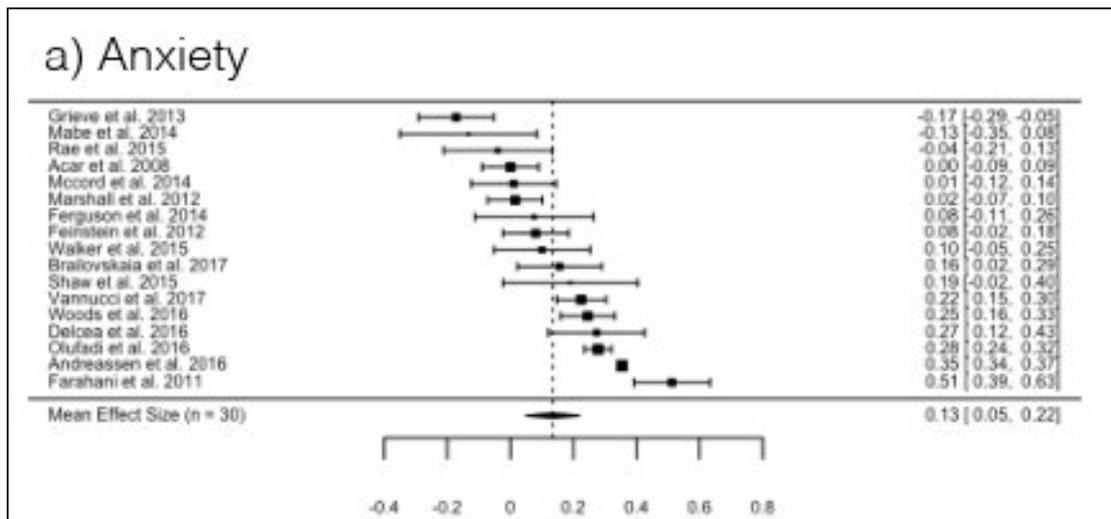
The forest plots for both are presented below. There are several notable findings. First, visually, if one scans from top to bottom, there is a notable temporal trend towards stronger associations over time and the authors note that for depression in particular, there was a statistically significant trend towards increased effect size. This is especially notable as the final included study was in 2015. There are many possible explanations for this including the evolution of SM algorithms and usage patterns over time. Second, the overall effect sizes, which the authors characterize as “small” are .13 and .12 for depression and anxiety respectively. Both effects were statistically significant and consistent with quasi-experimental and experimental studies referenced previously in this report.

Figure 64: Depression<sup>587</sup>



<sup>587</sup> Hancock et al. “Psychological Well-Being and Social Media Use: A Meta-Analysis of Associations between Social Media Use and Depression, Anxiety, Loneliness, Eudaimonic, Hedonic and Social Well-Being”

Figure 65: Anxiety<sup>588</sup>



e) Orben A. Teenagers, screens and social media: a narrative review of reviews and key studies. *Social Psychiatry and Psychiatric Epidemiology*. 2020/04/01 2020;55(4):407-414. doi:10.1007/s00127-019-01825-4

551. Counsel also provided me a narrative review of existing systematic reviews.<sup>589</sup> This short review article by a noted skeptic of the association between SM and adverse effects concludes that the field is “dominated by many cross-sectional studies” and the overall effect sizes are negative “but very small.” This review neglects that there are longitudinal studies and even some quasi-experimental studies that find significant effects that even when small, have significant public health implications at scale as well as the fact that there are subgroups for whom the effects are even larger.

<sup>588</sup> Hancock et al. “Psychological Well-Being and Social Media Use: A Meta-Analysis of Associations between Social Media Use and Depression, Anxiety, Loneliness, Eudaimonic, Hedonic and Social Well-Being”

<sup>589</sup>Orben A. Teenagers, screens and social media: a narrative review of reviews and key studies. *Social Psychiatry and Psychiatric Epidemiology*. 2020/04/01 2020;55(4):407-414. doi:10.1007/s00127-019-01825-4

## N. NASEM Report

552. The National Academy of Medicine is a highly respected independent body of scientists which, among other things, commissions and publishes reports on important medical topics. I have served as a panelist on such reports, and for 6 years was a member of their Board of Children Youth and Families. In general, I have tremendous respect for the institution and the work they do. That said, there are several notable things that negatively impact the methodology and findings of this report.

553. Lack of relevant expertise on the committee. The National Academy Handbook on committee selection states, “Committee members are chosen based on their knowledge and experience in the various aspects of the topics to be investigated.”<sup>590</sup> Of the 10 members of the panel, only 2 have actively researched social media and adolescent mental health. Others have researched related topics including AI and social media, bullying etc, and some appear to have no relevant research experience at all. There was only one member in a school of Public Health even though the report, and its implications, are squarely grounded in a public health issue.<sup>591</sup>

554. Conflict of Interest. At least two of the panelists have had their research supported by digital media companies. One received an unrestricted gift from Google and the other from Instagram. The NAM policy is that even *research* support from a relevant industry would be disqualifying from serving.

555. The majority of expert reviewers of the report, while established and credible academic scientists, were likewise, by and large, not media researchers and at least one of them

---

<sup>590</sup> National Academy of Sciences E, and Medicine. A Guide for Committee Members. Accessed April 2025

<sup>591</sup> Allem J-P. Social Media and Adolescent Health. *American Journal of Public Health*. 2024;114(10):980-982. doi:10.2105/ajph.2024.307784

espouses a very clear contrarian position about the effects of media on adolescent health. The combination of a paucity of experts on the panel and reviewers with a firmly entrenched position critiquing their work can result in undue influence and effects.

556. When discussing the “positive” effects of social media on adolescent health, panelists relied on the same types of studies as those showing harm which is to say studies that were mostly observational, often cross-sectional, relied on self-reported usage etc. However, none of those limitations were mentioned as problems in that section in the same way they were in the chapter on harms. Furthermore, the section chose to call out one of the longitudinal studies of social media use by Coyne et al.<sup>592</sup> That study did in fact not find significant negative effects of social media usage on adolescent wellbeing. However, although the demographics and sampling frame are not reported in detail in the cited paper, they are from the Flourishing Families Study. Below is a description of how that panel was assembled.

**Figure 66: Description of Sample from Coyne et. al<sup>593</sup>**

The researchers chose a northern section of Seattle and the Provo area for their data collection. The Seattle sector includes unusual diversity within a relatively small area and also has a low crime rate, offering a safety factor for the students. The Provo area offers the possibility of comparing and contrasting Latter-day Saint families to others. To recruit families, they sent letters of invitation to every family in the area with one or two parents and at least one child in the 10- to 13-year-old age range. From those that responded, they chose a random sample, ending up with 500 families in Seattle and 200 in Provo. The Seattle group “mirrors the north Seattle census statistics,” says Day, with a wide range of education levels and financial situations; about 20 percent are families of color. The Provo area families, as expected, are less diverse.

---

<sup>592</sup> Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. 2014;111(24):8788-8790. doi:doi:10.1073/pnas.1320040111

<sup>593</sup> *Families that Flourish*, BYU (2012), located at <https://magazine.byu.edu/article/families-that-flourish/> (last accessed Apr. 16, 2025).

And here, from a separate paper using the same sample, is a description of the demographics.<sup>594</sup>

**Figure 67: Description of Sample from Padilla-Walker et. al.**

At Time 1, participants were 500 (163 single-parent and 337 two-parent) families, 96% of whom had complete data for Time 2 (N = 478, 154 single-parent and 324 two-parent families). Regarding ethnicity, 86% of fathers, 75% of mothers, and 69% of children were European American, 6% of fathers, 14% of mothers, and 13% of children were African American, and 8% of fathers, 11% of mothers, and 18% of children were from other ethnic groups or were multi-ethnic. Seventy percent of fathers and 59% of mothers reported having a bachelor's degree or higher. Average monthly income for fathers was US\$6,572 (SD = 5,316.46) and for mothers was US\$3,538 (SD = 8,231.68). Ninety-five percent of fathers and 66% of mothers reported being currently married (never divorced); 10% of mothers were single parents, never been married, 4% were separated, 15% were divorced, 3% were cohabiting, and 2% were widowed.

557. As would be expected given the sampling frame, it is not remotely representative of the US population. Why they singled out that longitudinal study when there are several other larger and more representative ones as well as several experimental deprivation ones is unclear.<sup>595</sup>

---

<sup>594</sup> Padilla-Walker LM, Christensen KJ. Empathy and Self-Regulation as Mediators Between Parenting and Adolescents' Prosocial Behavior Toward Strangers, Friends, and Family. *Journal of Research on Adolescence*. 2011;21(3):545-551. doi: <https://doi.org/10.1111/j.1532-7795.2010.00695.x>

<sup>595</sup> Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. 2014;111(24):8788-8790. doi:10.1073/pnas.1320040111

Allcott H, Braghieri L, Eichmeyer S, Gentzkow M. The Welfare Effects of Social Media. *American Economic Review*. 2020;110(3):629-76. doi:10.1257/aer.20190658

Hunt MG, Marx R, Lipson C, Young J. No More FoMo: Limiting Social Media Decreases Loneliness and Depression. *Journal of Social and Clinical Psychology*. 2018;37(10):751-768. doi:10.1521/jsoc.2018.37.10.751; Braghieri L, Levy Re, Makarin A. Social Media and Mental

558. Further, the report goes on to say, “Social media can be valuable to adolescents who otherwise may feel excluded or lack offline support, including patients with rare diseases or disabilities, and those who struggle with obesity or mental illness, or come from marginalized groups such as LGBTQ+ young people.” For this quote, they happen to cite an AAP guideline that I am an author on.<sup>596</sup> That particular statement was based on our opinion and not on any original science (which they fail to cite). Finally, the report states that “At its most extreme end, isolation and related mental health problems can manifest in suicidal thoughts and self-harm. Some evidence indicates that supportive online communities can decrease risks of suicidal ideation and improve wellbeing.<sup>597</sup> That study (authored by a member of the committee) did indeed find benefits... On Reddit. There are multiple salient differences between reddit and social media sites including the demography of users, the site’s features (generally anonymous, organized by topics rather than individuals, no algorithmic display of content, pictures etc) and of course Reddit is not a defendant in this suit.

### **XIII. Prevention Measures**

559. Below, I will walk through specific prevention steps that could have been taken by Defendants to better protect child safety, recognizing the specific vulnerabilities of children. This includes evidence that age verification tools were inadequate, and parental controls were absent,

---

Health. *American Economic Review*. 2022;112(11):3660–93. doi:10.1257/aer.20211218; Bridge JA, Greenhouse JB, Ruch D, et al. Association Between the Release of Netflix’s 13 Reasons Why and Suicide Rates in the United States: An Interrupted Time Series Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2020/02/01/ 2020;59(2):236-243. doi:<https://doi.org/10.1016/j.jaac.2019.04.020>; Niederkrotenthaler T, Braun M, Pirkis J, et al. Association between suicide reporting in the media and suicide: systematic review and meta-analysis. *BMJ*. 2020;368:m575. doi:10.1136/bmj.m575.

<sup>596</sup> Reid Chassiakos YL, Radesky J, Christakis D, Moreno MA, Cross C. Children and Adolescents and Digital Media. *Pediatrics*. Nov 2016;138(5)doi:10.1542/peds.2016-2593.

<sup>597</sup> De Choudhury M, Kıcıman E. The Language of Social Support in Social Media and its Effect on Suicidal Ideation Risk. *Proc Int AAAI Conf Weblogs Soc Media*. May 2017;2017:32-41.

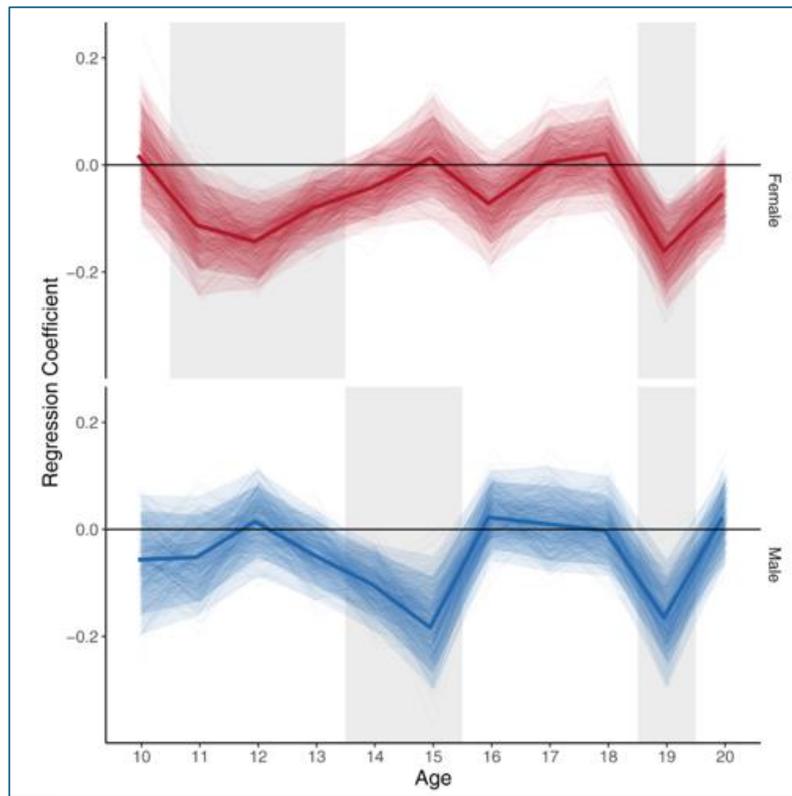
limited, or ineffectively implemented. From a public health standpoint, allowing young users on social media platforms as they existed up to present puts these vulnerable children at risk of potentially severe harms. Because children are particularly vulnerable to social media harms, protections for them should have been anything but lackluster.

560. The variable developmental sensitivity to social media sites was illustrated in a study by Orben, Przybylski et al. They used the “Understanding Society” cohort study of 17,409 10-21 year olds to look at how self-reported social media use predicted life satisfaction one year later.<sup>598</sup> Some of their findings are summarized below.

---

<sup>598</sup> Orben A, Przybylski AK, Blakemore S-J, Kievit RA. Windows of developmental sensitivity to social media. *Nature Communications*. 2022/03/28 2022;13(1):1649. doi:10.1038/s41467-022-29296-3

**Figure 68: How Social Media Use Predicts Life Satisfaction In Longitudinal Data<sup>599</sup>**



The grey rectangles illustrate ages where the effects were significantly different from the null. For both females and males, younger ages are associated with increased likelihood of negative effects on life satisfaction. Females are more vulnerable between the ages of 10-13 and males between the ages 13 ½ to 15 ½.

561. To put the vulnerabilities of younger children into perspective, consider that according to a survey of over 4,500 U.S. parents who chose to delude their children about the

<sup>599</sup> Orben A, Przybylski AK, Blakemore S-J, Kievit RA. Windows of developmental sensitivity to social media. *Nature Communications*. 2022/03/28 2022;13(1):1649. doi:10.1038/s41467-022-29296-3

existence of Santa Claus, the average age at which children stop believing in him is 8.4 years.<sup>600</sup> And a subsequent study found that as many as 20% of 10-year-olds still believe in him.<sup>601</sup>

562. Parental involvement is essential to mitigating risks to children. Research shows that open communication and social support are critical in buffering teens from the harms of online abuse. Social support from trusted adults and peers can significantly blunt the psychological impact of exposure to things like cyberbullying or hate speech. A systematic review of effective strategies to combat and mitigate cyberbullying identified families as being a key component.<sup>602</sup> Trained parents can help their children prevent, identify, and cope with cyberbullying but doing so requires that they be alerted to warning signs either by the victim or the platform. Likewise, with respect to sexting, although schools emerge as critical intervention sites, parents and family involvement is essential.<sup>603</sup>

563. Finally, although research into sextortion is limited and in early stages, notification of parents or family members is an important and effective strategy to mitigate harms.<sup>604</sup> However, fewer than 50% of teens do so highlighting the essential role that the sites must play in prevention. Defendants could provide the ability to report inappropriate interactions or CEI/CSAM to a safety

---

<sup>600</sup> Helen Brown, *When Do Children Stop Believing in Santa?*, MADE FOR MUMS (Dec. 5, 2023), located at [https://www.madeformums.com/news/when-do-children-stop-believing-in-santa/?utm\\_source=chatgpt.com](https://www.madeformums.com/news/when-do-children-stop-believing-in-santa/?utm_source=chatgpt.com)

<sup>601</sup> Elisabeth Beauchamp, Lora Novak, *At What Age Did Americans Stop Believing in Santa?*, TODAY'S HOMEOWNER (Nov. 13, 2024), located at <https://todayshomeowner.com/blog/guides/not-believing-in-santa-by-state/>

<sup>602</sup> Tozzo P, Cuman O, Moratto E, Caenazzo L. Family and Educational Strategies for Cyberbullying Prevention: A Systematic Review. *International Journal of Environmental Research and Public Health*. 2022;19(16):10452.

<sup>603</sup> Ojeda M, Del Rey R. Lines of Action for Sexting Prevention and Intervention: A Systematic Review. *Archives of Sexual Behavior*. 2022/04/01 2022;51(3):1659-1687. doi:10.1007/s10508-021-02089-3

<sup>604</sup> Ray A, Henry N. Sextortion: A Scoping Review. *Trauma, Violence, & Abuse*. 2025;26(1):138-155. doi:10.1177/15248380241277271

support team that provides immediate feedback to the child. If a child reports such inappropriate interactions or illegal content, the Defendant could also notify the parent. I have not seen evidence that this happens effectively and have seen evidence to the contrary. For example, Jayakumar acknowledged at her deposition that, as of March 30, 2020, Instagram did not have a specific way for people to report CSAM on its platform.<sup>605</sup> Further, despite publicly claiming it was increasing staff to review reports of CSAM, Meta did the opposite and slashed the total workforce assigned to that job.<sup>606</sup>

564. In order for parents and guardians to provide the support that children require, they need adequate ability and access to control social media usage for their children. Parents also need to receive full and accurate information regarding the mental health harms that can be caused by social media so that they can make informed decisions for their children. As discussed below, Defendants could have provided adequate parental controls and information to parents but failed to do so. Relatedly, they failed to verify the ages of the users, a meaningful predicate to ensuring that any parental controls are effective.

565. As part of my review of materials, I reviewed interrogatory answers provided by the Defendants. Some of these answers contained relevant information regarding various safety features, their adoption rates, and whether they were default/not default. I have summarized some of the more relevant information below.

---

<sup>605</sup> Vaishnavi Jayakumar Dep. Tr. at 274:6-11; *see also* Jayakumar Dep. Ex. 31 (“We don’t have an in-app reporting option for CEI that I’m aware of.”); Vaishnavi Jayakumar Dep. Tr. at 275:23-276:2 (“[R]eporting CEI is sort of baked into most platforms. It was a surprise to me, I was only two months into the company, it was a surprise to me that you couldn’t report CEI within the app.”).

<sup>606</sup> Vaishnavi Jayakumar Dep. Tr. at 352:9-356:15.

566. Tiktok has a series of screentime management tools: daily screen time limit, screen time breaks, and sleep reminders. TikTok defined these as follows:

567. Daily Screen Time Limit: Feature launched in February 2019 and then made default in March 2023 that allows users to set a personalized, in-app screen time limit that are locked behind a user-set four digit passcode. Starting in 2019, parents were able to set a cap on their teen's screen time through the platform's Family Center feature. In March 2023, TikTok introduced a 60-minute default daily screen time limit for under-18 users (which can be overridden with a user-set four-digit passcode).<sup>607</sup> This feature was made default for U18 in March 2023 (the adoption rate below, however, is presumably less than 100% because minors could turn off limits).<sup>608</sup>

568. Screen Time Breaks: Feature that launched in June 2022 that allows users to set custom reminders to take breaks from the Platform after certain allotted periods of time.<sup>609</sup> The notification prompts the user to either "(1) tap OK to dismiss the notification; (2) tap Snooze to restart the timer and remind you again in 10 mins; or (3) tap Edit Reminders to change the screen time break schedule."<sup>610</sup> This feature was not on by default and had to be enabled and configured by the user.<sup>611</sup>

569. Sleep Reminders: Feature that launch in in July 2023 that allows users to set a specific time of night that they would like to stop using the app. Once the time is reached, the user is sent a "reminder notification" that prompts the users to "wind down and log off the platform."

---

<sup>607</sup> TikTok's Supplemental Response to Interrogatory No. 8 at 31-32.

<sup>608</sup> TikTok's Supplemental Response to Interrogatory No. 8 at 46.

<sup>609</sup> TikTok's Supplemental Response to Interrogatory No. 8 at 32-33.

<sup>610</sup> TikTok's Supplemental Response to Interrogatory No. 8 at 32-33.

<sup>611</sup> TikTok's Supplemental Response to Interrogatory No. 8 at 33.

It also mutes notifications for “seven to eight hours depending on the user’s age.”<sup>612</sup> Sleep Reminders was available as an “opt-in” feature and was not default.<sup>613</sup>

570. The table below reflects the adoption rates for these “Screentime Management Tools between January 2024 to January 2025:<sup>614</sup>

**Table 6: Adoption Rates of Daily Screen Time, Screen Time Breaks, and Sleep Reminders on TikTok’s Platform between January 2024 and January 2025**

DATE	Daily Screen Time	Screen Time Breaks	Sleep Reminders
January 2024	77.5%	1.4%	0.7%
February 2024	77.3%	1.5%	0.7%
March 2024	86.0%	1.5%	0.8%
April 2024	91.2%	1.5%	0.8%
May 2024	90.6%	1.5%	0.8%
June 2024	89.8%	1.5%	0.8%
July 2024	88.4%	1.5%	0.8%
August 2024	86.8%	1.5%	0.9%
September 2024	85.9%	1.5%	0.9%
October 2024	85.7%	1.5%	0.9%
November 2024	85.5%	1.5%	1.8%
December 2024	85.0%	1.5%	0.9%
January 2025	83.8%	1.4%	0.9%

<sup>612</sup> TikTok’s Supplemental Response to Interrogatory No. 8 at 34.

<sup>613</sup> TikTok’s Supplemental Response to Interrogatory No. 8 at 34.

<sup>614</sup> TikTok’s Supplemental Response to Interrogatory No. 9 at 36-37.

571. **Restricted Mode:** A feature that limits a user’s exposure to certain content that “may not be comfortable for everyone, such as content that contains mature or complex themes.” It also limits access to other platform features including: (1) the “following feed;” (2) “going LIVE; and (3) “Gifting on Live.”<sup>615</sup> This feature was not on by default and had to be enabled/configured by the user.<sup>616</sup> The table below reflects the adoption rate between April 2024 to December 2024.<sup>617</sup>

**Table 7: Adoption Rate of the TikTok Platform’s “Restricted Mode from April to December 2024**

DATE	13-15 Users	16-17 Users
April 2024	.61%	.18%
May 2024	.63%	.18%
June 2024	.63%	.19%
July 2024	.65%	.19%
August 2024	.63%	.18%
September 2024	.67%	.19%
October 2024	.70%	.20%
November 2024	.74%	.22%
December 2024	.76%	.23%

572. Similar responses were provided from Meta. Those features and relevant launch and adoption rates are summarized below:

<sup>615</sup> TikTok’s Supplemental Response to Interrogatory No. 8 at 30-31.

<sup>616</sup> TikTok’s Supplemental Response to Interrogatory No. 8 at 30-31.

<sup>617</sup> TikTok’s Supplemental Response to Interrogatory No. 9 at 37.

**Table 8: Summary of Adoption Rates and Default Setting Status of Various Safety Features on Instagram**

<b>Feature</b>	<b>Description</b>	<b>Launch Date</b>	<b>Adoption Rate</b>	<b>Default at Launch (U18)?</b>	<b>Default Today (U18)?</b>
Hide Likes & Views	Allows users to hide like counts on all posts that appear in their feed, posts they create, and/or for other users' posts	5/26/2021	0.0348 (1/8/2023-3/23/2025)	No	No
Take a Break	Optional setting on Facebook and Instagram that sends a notification suggesting that the user "take a break" after they've spent an allotted amount of time on the platform.	12/7/2021	0.0045 (1/8/2023-9/24/2024)	No	N/A
Quiet Mode	Optional setting that disables all notifications and sends an automatic reply when the user receives a direct message.	1/19/2023	0.1352 (7/9/2023-4/1/2024)	No	No
Nighttime Nudge	A prompt that encourages teen users to enable Quiet Mode after using Meta's platforms during "nighttime hours."	1/19/2023	0.5111 (7/9/2023-4/1/2024)	Yes	Yes
"New" Nighttime Nudge for Teens	Expanded version of the Nighttime Nudge feature that prompts teens to activate Quiet Mode after spending 10 minutes on Reels or Direct Messages during nighttime hours	1/18/2024	0.5123 (1/18/2024-4/1/2024)	Yes	Yes
Not Interested	A user control that allows users to "hide content that they do not want to see" and alter Instagram's algorithms to "avoid recommending such content in the user's Explore, Reels, and Search pages."	1/19/2023	0.029 (11/28/2023-4/1/2024)	N/A	N/A

<b>Feature</b>	<b>Description</b>	<b>Launch Date</b>	<b>Adoption Rate</b>	<b>Default at Launch (U18)?</b>	<b>Default Today (U18)?</b>
Limit Unwanted Interactions	Optional tool that allows Instagram users to limit comments and direct message requests from other users who do not follow them or only recently followed them on Instagram	8/10/2021	0.0008 (2/24/2025-3/23/2025)	No	Yes – only U16
Daily Limit	Tool that allows users to receive an alert after spending a predetermined amount of time on Instagram.	8/1/2018	0.161 (1/8/2023-3/23/2025)	No	Yes
Your Activity	A feature that allows users to view and manage their activity on Instagram. Currently, this feature allows users to (1) view time spent on Instagram per day; (2) archive multiple shared posts at once; (3) delete multiple shared photos and reels at once; (4) edit and delete highlights; (5) review and delete multiple likes and comments at once; (6) view and manage posts in which they have been tagged; (7) view and manage their account history; (8) view and clear their recent search history; (9) view their recent link history; (10) view and manage recently deleted and archived content; (11) view advertisements they have recently interacted with; (12) view and manage reviews they have left for purchases on Instagram; and (13) download a copy of all information they have shared on Instagram.	2/8/2022	0.1572 (9/7/2024-3/24/2025)	N/A	N/A
Single-Tap Teen Safety & Privacy	A feature that prompts teen users to update their privacy settings on Instagram with “a single tap.”	1/9/2024	0.0407	N/A	N/A

Feature	Description	Launch Date	Adoption Rate	Default at Launch (U18)?	Default Today (U18)?
Settings Prompts			(12/26/2024-3/24/2025)		
Family Center Supervision	A set of parental controls that grants parents and guardians access to “supervision tools and resources from leading experts.” These tools include (1) Teen Reporting Notifications; (2) Teen Follows; (3) Teen Followers; (4) Teen Time Spent; (5) Time Limits and Set Specific Times of Use; and (6) Education Hub.” <sup>618</sup>	3/22/2022	0.0038 (3/23/2025-3/24/2025)	No	Only U16
Teen Accounts	A set of built-in protections that defaults teen users into the strictest privacy, messaging, sensitive content, interaction, and time management settings. It also requires teens under 16-years-old to enroll in parental supervision and obtain parental permission before changing their more restrictive default settings.	9/17/2024	Unknown <sup>619</sup>	Yes	Yes

**O. At What Age Should Adolescents Use Social Media?**

573. Because of the increased risk of harm to children and adolescents, in my opinion, defendants’ social media platforms should restrict children under the age of 16 from using their platform or at a minimum, require informed parental consent and have extensive, effective parental

<sup>618</sup> Meta did not provide adoption rates for these tools individually in its responses to Interrogatory No. 12.

<sup>619</sup> Meta did not provide adoption rates for Teen Accounts in its responses to Interrogatory No. 12.

controls. For teenagers age 16 and 17, I would recommend parental consent be required and effective parental controls be implemented for use of social media.

574. The legal “13-year-old minimum” age for creating social media accounts in the United States emanates from the Children’s Online Privacy Protection Act (COPPA). Enacted in 1998 and implemented by the Federal Trade Commission (FTC) in 2000, COPPA was designed to give parents control over what data websites and online services can collect from children under 13 years of age. (It is not, in fact, a “legal limit,” but rather just sets certain requirements for usage of platforms by minors under the age of 13.) COPPA was enacted prior to the creation of the social media companies I discuss in my report. Further, COPPA is not a scientific standard for what is appropriate, healthy, or beneficial for adolescents. Rather, COPPA is a legal, statutory minimum regarding data collection and restrictions on monetization of internet collected data. As Dr. Alison Lee, a senior UX researcher at Meta, testified: “There’s also likely a lot of harm that may happen as a result of lack of support to those young people, especially in those digital spaces that were not designed for young people in the first place.”<sup>620</sup>

575. As discussed above, the white matter of the brain has not reached full maturity at age 13 and is decidedly less mature at age 10. Internally, Meta documents recognized that “[t]he teenage brain is usually about 80% mature... At this time teens are highly dependent on their temporal lobe where emotions, memory and learning, and the reward system reign supreme.”<sup>621</sup>

576. Speaking as a pediatrician, an epidemiologist and a parent, age milestones must be viewed from a developmental perspective. I am not aware of any scientific support that 13 is the age at which children can safely engage with social media without supervision. My opinion is that

---

<sup>620</sup> Alison Lee Dep. Tr. at 21:19-23

<sup>621</sup> Mark Zuckerberg Dep. Exhibit 30 at -5452

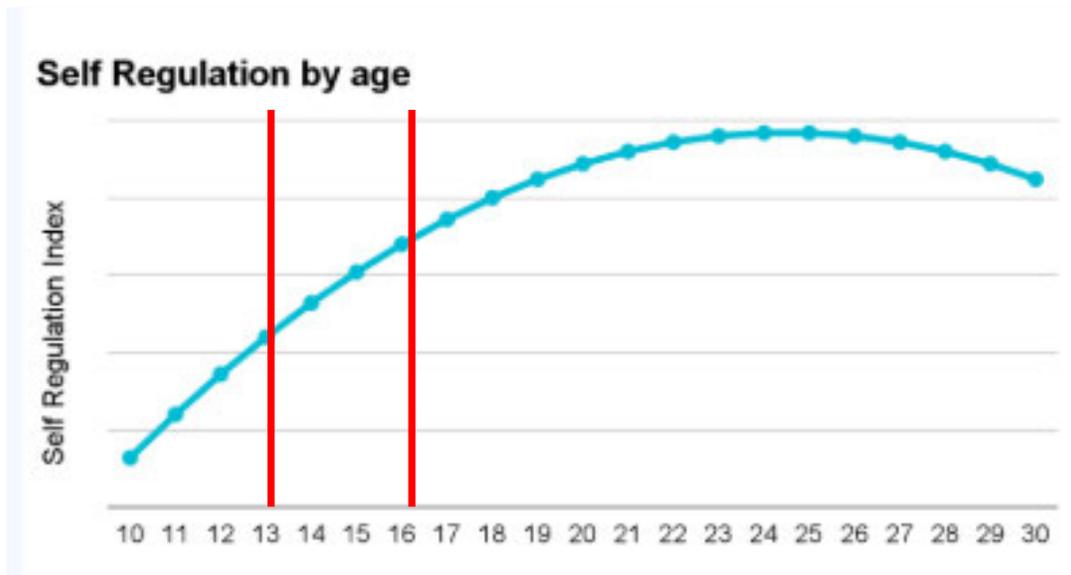
much of the science actually demonstrates that social media causes mental health and other harms, especially to children. Meta's documents quote David Kleeman, Senior Vice President at Dubit, as saying, "While most platforms have set their minimum age for participation at 13, there's no on/off switch that makes someone ready to be a fully media-literate participant on that birthday."<sup>622</sup> And when asked in her deposition if there was any longitudinal data to support the 13 year and older age restriction. Dr. Moira Burke, a Meta user experience (UX) research scientist, responds, "I am not aware of longitudinal research on anything related to the ages that people start."<sup>623</sup>

577. Restricting social media use and providing effective parental controls for pre-teens and teens seems entirely justifiable if one looks at *Figure 14* and *Figure 16*. By age 21, both brain development and executive function are by and large fully complete. There is still rapid growth of executive function at age 13 (see super imposed dotted red lines on figures). Internal documents from YouTube appear to recognize this:

---

<sup>622</sup> Alison Lee Dep. Exhibit 29 at 2-3

<sup>623</sup> Moira Burke Dep. Tr. at 50:24-51:1



*Document 191: GOOG-3047MDL-01719787 at Slide 17 (emphasis added)*

578. Indeed, the Handbook of Children and Screens recommends a minimum age for “social media use as 16 given the consistent links between social media use and depression of evidence of causality.”<sup>624</sup> As part of these recommendations, we also note a need for effective age verification in order to consistently enforce age minimums.<sup>625</sup>

579. Some believe that denying children access to social media during early adolescence deprives them of the ability to connect with others and share insights, information, entertainment etc. Further, some believe that especially for vulnerable and potentially isolated populations (e.g. LGBTQ+ youth in rural areas) social media sites might enable them to connect with others with whom they identify in a safe space that might be affirming and even lifesaving. I am cognizant and sensitive to that position and it is discussed at length in the Handbook of Children and Screens. But scientific research indicates the data on the benefits versus harms for LGBTQ+ youth in

<sup>624</sup> Handbook of Children and Screens at 139

<sup>625</sup> Handbook of Children and Screens at 139

particular are mixed<sup>626</sup> and the Thorn report (discussed above) revealed that LGBTQ youth are more likely to be bullied than their heterosexual peers on line. Moreover, there are other, safer ways to enable community-building without relying on extant social media sites.

**P. Inadequate Age Verification**

**i) Meta**

580. Despite COPPA requiring restrictions for users under 13-year-old, children under 13 are still able to (and do) access social media platforms and accounts, including those operated by Meta. Indeed, in 2018, Meta’s own analysis reported that there were “4 million people under 13 in 2015 on IG. This represents around 30% of all 10–12-year-olds in the US.”<sup>627</sup>

- Looking at people we predict to be 13 and 14 today, we can estimate that there were 4M people under 13 in 2015 on IG. This represents around 30% of all 10-12 years old in the US. Finally, assuming similar numbers today, 75% of US teens are MAP on IG.

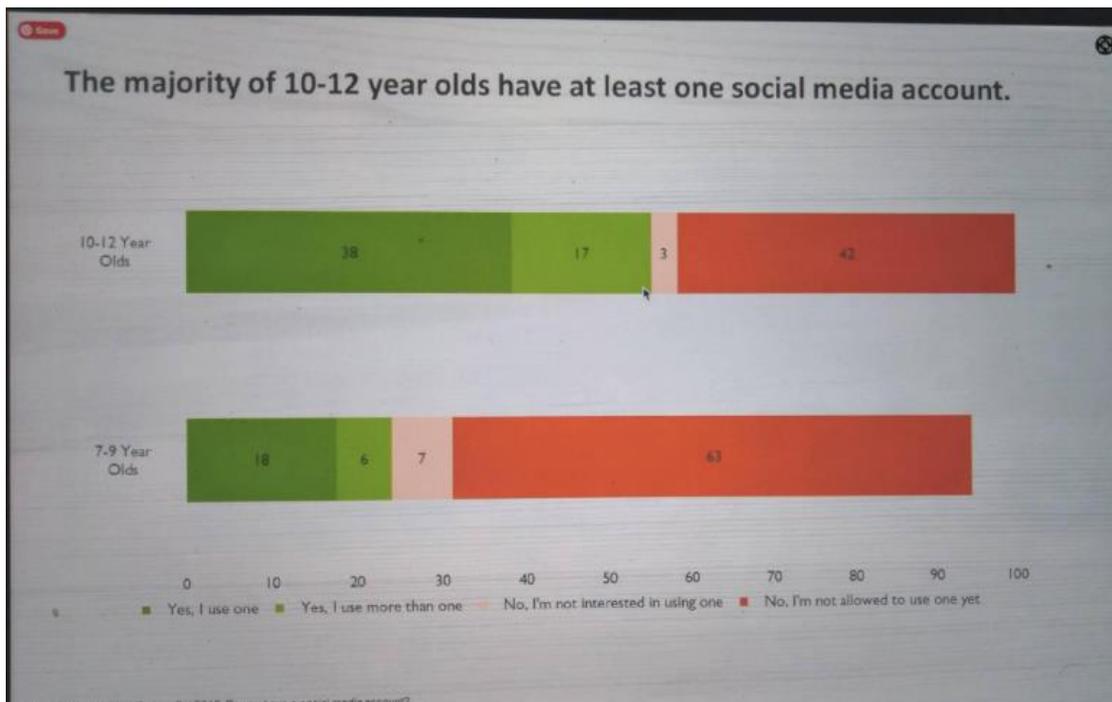
*Document 192: META3047MDL-014-00133717, -3721*

581. Likewise, a 2017 study commissioned by one of Facebook’s research managers, Anja Dinhopf, found that the majority of 10–12-year-olds and 24% of 7–9-year-olds have at least one social media account in spite of age “gating”:

---

<sup>626</sup> Escobar-Viera CG, Whitfield DL, Wessel CB, et al. For Better or for Worse? A Systematic Review of the Evidence on Social Media Use and Depression Among Lesbian, Gay, and Bisexual Minorities. *JMIR Ment Health*. 2018/07/23 2018;5(3):e10496. doi:10.2196/10496

<sup>627</sup> META3047MDL-014-00133717, -3721



*Document 193: Haugen\_00023849, -3866*

This “mixed methods” study is notable given it included focus groups of 10–12-year-olds, child parent interviews, and a survey of 1450 7–12-year-old children about their social media habits.<sup>628</sup>

Arturo Bajur, in his testimony, stated that if Meta’s executives say that they do not allow kids under 13 on Instagram they would be lying.<sup>629</sup>

582. In 2020, Thorn (a not for profit dedicated to child online safety) issued a report that was circulated within Meta that discussed TikTok, Snap, and YouTube user ages. Among other findings, it reported:

<sup>628</sup> Haugen\_00023849. Interestingly, the study cites the “sensitive nature of the subject” in explaining that “research findings, materials and raw data about tweens’ social media usage are only available upon request.” Haugen\_00023849, -3855.

<sup>629</sup> Arturo Bejar Dep. Tr. at 190:11-17

<b>Platform usage habits: At least once a day</b>			
<b>Company</b>	<b>All minors</b>	<b>Ages 9-12</b>	<b>Ages 13-17</b>
Discord	19%	12%	25%
Facebook	36%	45%	28%
Google Hangouts/Meet	33%	36%	30%
Instagram	50%	40%	59%
Messenger	32%	38%	27%
Pinterest	16%	13%	17%
Reddit	10%	8%	12%
Snapchat	47%	40%	52%
Telegram	6%	10%	3%
TikTok	45%	41%	49%
Tumblr	13%	18%	9%
Twitter	28%	30%	27%
WhatsApp	26%	39%	16%
YouTube	80%	78%	81%

*Document 194: META3047MDL-031-00245499, -5503*

This report is entirely consistent with Meta’s own data from 2017, which found that 55% of 10–12-year-olds have at least one social media site and the very limited steps Meta took to mitigate it since then. In response, Pavni Diwanji (VP of UX Research) responded to the group via email, “This is a big WHOA, if these numbers are to be believed.... If we feel that these numbers are in the right ballpark, it’s hard to justify bringing more kids onto our platform before we make it better, afford better protections. And if the numbers are not in the right ballpark, can we make a case or refute these?”<sup>630</sup>

<sup>630</sup> META3047MDL-031-00245499, -5501

583. Notably, Instagram did not even start *asking* users for their date of birth until December 2019, nine years after it was launched and it did not *require* it until March of 2021.<sup>631</sup> Merely “asking” for a birthdate is a minimal and easily circumventable safeguard which many of the defendants readily acknowledge in their internal communications. Children as young as 9 or 10 (possibly younger) were (and likely still are) on the platform.

584. Internal documents reflect that Meta executives recognized that its age limits were not really working. In 2019, Nick Clegg, Meta President of Global Affairs, texted his team, “The fact that we have age limits that are unenforced (unenforceable?) and that there are, as I understand it, important differences in the stringency of our policies on IG vs Blue App [Facebook] makes it difficult to claim we are doing all we can.”<sup>632</sup> Similarly, in a 2019 email chain, Monika Bickert (VP of Public Policy), informed others she was getting an error message when trying to report an underage account brought to her attention by someone at a child’s school. She adds: “The reporting flow was pretty bad. I wondered if we should look into it. *It was obviously structured to deter any reports.*”<sup>633</sup>

585. Troublingly, the “final word” in this email chain came from ██████████ (Product Manager) who stated, “Improving this is not currently in the plans for FRX and looking at these numbers seem quite small (*only* 15K completions per week)”.<sup>634</sup> But 15,000 children is the equivalent of approximately 20 entire average sized US high schools per week. Moreover, Mr. ██████████ reply overlooks that the “bad reporting flow” may, in part, explain the “low” numbers.

---

<sup>631</sup> Diego Castaneda Dep. Ex. 4; *See also* Diego Castaneda Dep. Tr. at 79:12-80:2.

<sup>632</sup> META3047MDL-003-00175144, -5153

<sup>633</sup> META3047MDL-014-00166515, -6517 (emphasis added)

<sup>634</sup> META3047MDL-014-00166515, -6515 (emphasis added)

Facilitating reporting to alleviate or prevent harm is an essential component of public health safety strategies. It is foundational to the institution of mandatory reporting.

586. The situation does not seem to have improved by 2021. In an internal chat that year, Meta employees indicated they did not even know the age of ~30% of Instagram users. Despite claiming externally that “we age gate,” these employees privately acknowledged that “age-gating” was never designed nor used for Integrity purposes, but rather was implemented only by the Instagram advertising team:

[REDACTED] (10/11/2021 10:04:43 PDT):  
>Specifically, in the past, the only requirement for age-gating on IG was by advertisers, around brand safety.

[REDACTED] (10/11/2021 10:05:06 PDT):  
>The relevant infra and code are owned by the IG Ads team ("IG4B") as opposed to Well-being. We never used that infra for Integrity. I don't even know how reliable it is.

[REDACTED] (10/11/2021 10:05:31 PDT):  
>So whoever wants to start using it for Integrity should chat with IG4B, see what gaps there are, etc.

[REDACTED] (10/11/2021 10:06:18 PDT):  
>Not opposed to enhancing it, just sharing that it was never designed or used for Integrity purposes in the past.

[REDACTED] (10/11/2021 10:11:25 PDT):  
>It sounds like some folks just assumed age-gating would work for Integrity, when it was never planned to.

[REDACTED] (10/11/2021 10:12:14 PDT):  
>It sounds like "we had some wishful thinking, found out it's wrong, so now we're filing a SEV to force someone else to fix it fast" :) None of this is a criticism, I'm probably missing something. Mostly just curious.

[REDACTED] (10/11/2021 10:12:17 PDT):  
>I think that's a fair point. The more important point is that this gap doesn't seem to be well known.

[REDACTED] (10/11/2021 10:12:42 PDT):  
>we can give the heads up and then decide if it should be a sev?

[REDACTED] (10/11/2021 10:12:54 PDT):  
>Sure. Can someone write it including what depends on it?

[REDACTED] (10/11/2021 10:13:05 PDT):  
>It wasn't a "gap" until now because nothing depended on it or tried to use it (outside the well-known Ads use-cases).

[REDACTED] (10/11/2021 10:13:18 PDT):  
>Someone can you frame the email for Miki and Yoav to take a look

[REDACTED] (10/11/2021 10:13:20 PDT):  
>I imagine some new functionality depends on this now?

[REDACTED] (10/11/2021 10:44:21 PDT):  
>we may be violating some policies without it

[REDACTED] (10/11/2021 10:47:30 PDT):  
>Can we separate the infra from the policy compliance problem? I don't know if the original intent was to use this infra specifically or what, but it's clear that something was overlooked here years ago, possibly before all of our time

[REDACTED] (10/11/2021 10:47:45 PDT):  
>[Sorry for maybe dup message, messages is flaky for me right no]  
>  
>The gap is external community standards pages mentioning we age gate but not implemented in IG product. Found this via action->UX audit. <https://transparency.fb.com/policies/community-standards/violent-graphic-content/>

[REDACTED] (10/11/2021 10:48:09 PDT):  
>Sure.

[REDACTED] (10/11/2021 10:48:17 PDT):  
>Yup, understood, not contested.

[REDACTED] (10/11/2021 10:48:34 PDT):  
>Y'all know we don't even have age on IG for a big chunk (last figure I saw was ~30%) of IG users, right?

*Document 195: META3047MDL-014-00355780, -5780-81*

587. Jayakumar, child safety lead at Instagram, states in her deposition that “Insta didn’t do enough to identify under 13’s.”<sup>635</sup> As of September of 2020, Instagram’s own documents reveal that there “was a backlog of 450,000 reports or noted incidences of potential users under the age of 13 that need to be reviewed and addressed.”<sup>636</sup> But later in the same text exchange, [REDACTED] (Facebook child safety) states that there are over 2.5 million.<sup>637</sup> That considerably larger estimate is plausible given this exchange between Jayakumar and [REDACTED], an Insta programmer:

[REDACTED] (11/11/2021 16:18:54 PST):  
>It's actually the opposite for minors who state they are adults. Here's some data from August:  
<https://fb.workplace.com/groups/1411637962558841/permalink/1675656312823670/>

[REDACTED] (11/11/2021 16:19:00 PST):  
>Comparing stated vs predicted age

Vaishnavi Jayakumar (11/11/2021 16:20:14 PST):  
>Oh this is great! Just to make sure I am understanding it right - around half of teens lie about their age, but most people who state that they are teens are actually teens

*Document 196: META3047MDL-020-00298458, -8458*

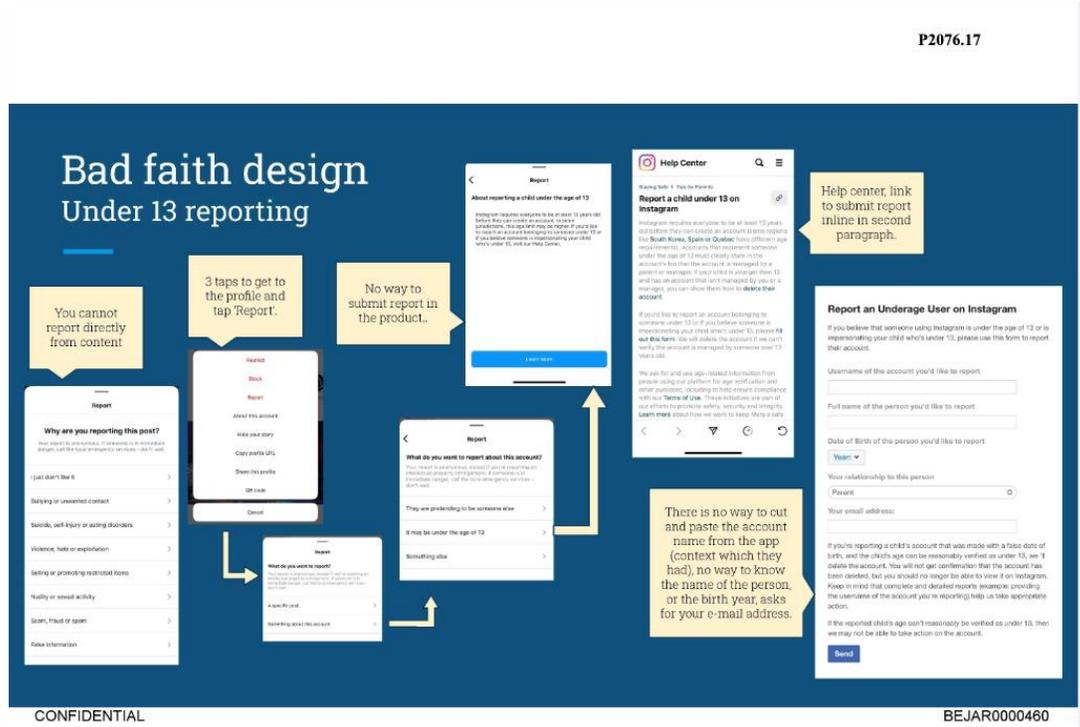
To call it out for emphasis, Insta was aware that “**around ½ of teens lie about their age.**” It should be a surprise to no one that given the ready access to adult content teens will lie about their age especially if there are no consequences to doing so.

<sup>635</sup> Vaishnavi Jayakumar Dep. Tr. at 82:19-21

<sup>636</sup> Vaishnavi Jayakumar Dep. Tr. at 89:14-18; *See also* Vaishnavi Jayakumar Dep. Exhibit 8.

<sup>637</sup> Vaishnavi Jayakumar Dep. Tr. at 96:21

588. Meta’s failure to identify and protect children under the age of 13 – who by law and Meta’s own stated policies should not have been on the SM – is evinced by the design features themselves:



Document 197: BEJAR0000444 at -0460

589. Internal documents recognize top-level discussions regarding targeting children under the age of 13:

[REDACTED] (5/04/2017 10:20:50 PDT):  
>on good, we re going after <13 year olds now?

[REDACTED] (5/04/2017 10:21:12 PDT):  
>zuck has been talking about that for a while

[REDACTED] (5/04/2017 10:21:37 PDT):  
>yeah it was gross the last time he mentioned it

*Document 198: META3047MDL-072-00304285 at -4289*

Later, in the same exchange:

[REDACTED] (5/04/2017 10:23:02 PDT):  
>im ok with hiring people, but targetting 11 year olds feels like tobacco companies a couple decades ago (and today)

[REDACTED] (5/04/2017 10:23:16 PDT):  
>like we're seriously saying "we have to hook them young" here

[REDACTED] (5/04/2017 10:23:58 PDT):  
>[https://fb.facebook.com/groups/e/permalink/1466622773393901/?comment\\_id=1468740833182095](https://fb.facebook.com/groups/e/permalink/1466622773393901/?comment_id=1468740833182095) I'm sure it was a harmless comment, just thought it was funny  
shared: <http://www.facebook.com/449949138693181>

[REDACTED] (5/04/2017 10:23:59 PDT):  
>well, to be fair, a lot of other services trying to compete with facebook explicitly target under-13's and create plausible deniability for knowing their age

*Document 199: META3047MDL-072-00304285 at -4290*

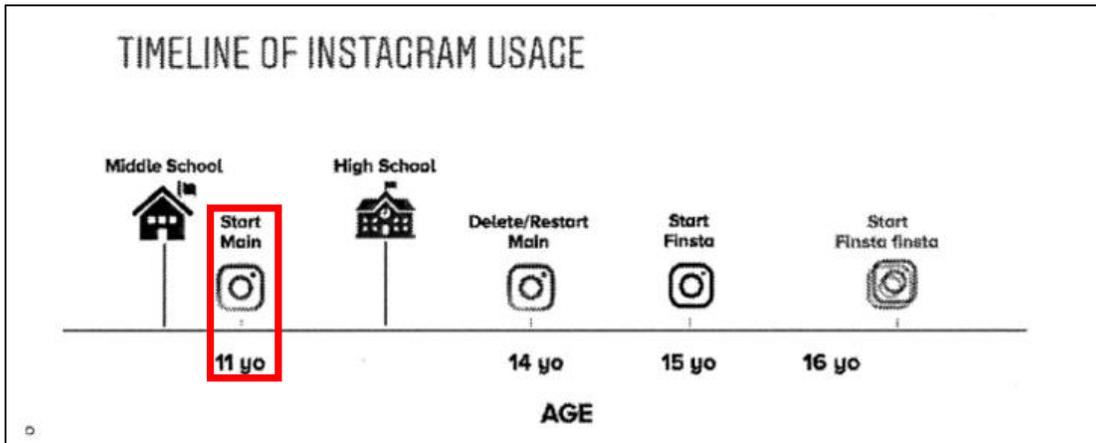
590. In fact, concern was expressed that the Federal Trade Commission might determine that Meta could determine the age of users and would then push for Meta to do so:

**From:** Antigone Davis [REDACTED]  
**Date:** Tuesday, August 1, 2017 at 9:53 AM  
**To:** [REDACTED]  
**Cc:** [REDACTED]  
**Subject:** Re: Data for FTC Meeting and External Communications

Hi, from the policy perspective I'd recommend 6 or 7-12, definitely not younger than 6 as 7-11 is the usual bucket. In terms of using our data, Dubit data or public data, we should use Dubit data or public data. If the FTC sees our data they will try to extrapolate from it that we can tell the exact age of people on our platform and could start pushing for us to use something other than stated age to ensure we remove age-liars. If we can't use Dubit's data, I believe Rally found public data and in early decks the product team cited some public data as well. Happy to lend a hand given the time crunch.

*Document 200: META3047-014-00123522 at -3523*

591. Meta's failure to understand the age of the children on its platform (or facilitate and act on reports of users under the age of 13 makes sense when paired with documents indicating that the company was actively looking to grow the number of young people who use its platforms. Indeed, while several Meta employees testified that under 13-year-old children were not a growth strategy, and when they were found using the site they were expediently removed, those claims are undermined by the following graphics, which were shown within the company during a presentation regarding sustaining and promoting growth of the app (in part through "Finstas," fake Instagram accounts).



Document 201: META3047MDL-031-00086272, -6273

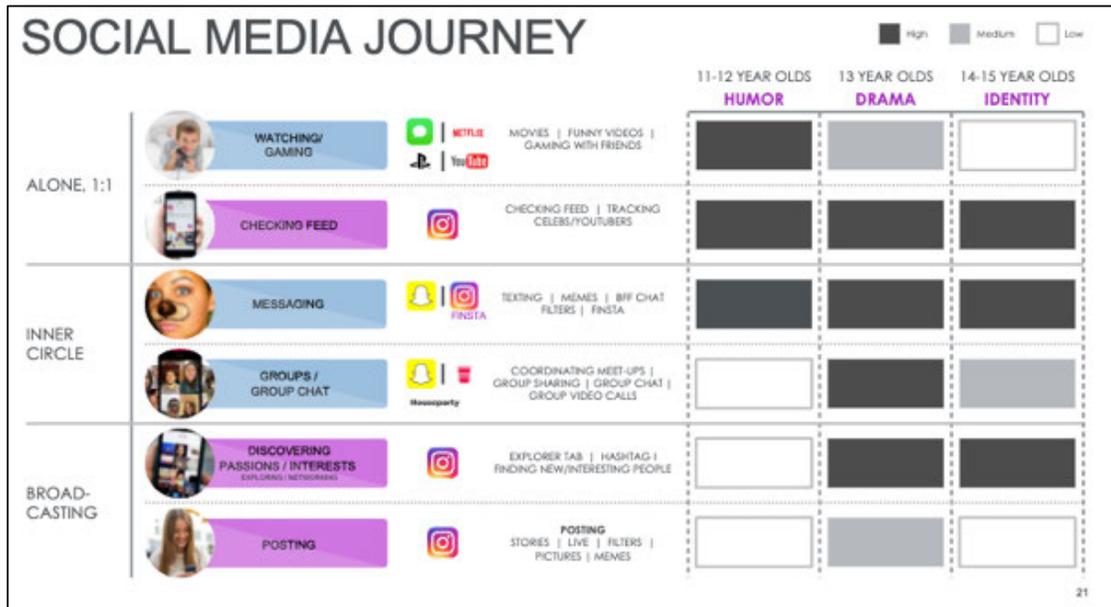
MAIN	FINSTA
<ul style="list-style-type: none"> <li>Start: First account between age 11-12</li> <li>Called "Main"</li> <li># of Followers: 500</li> <li>Videos: Less videos</li> <li>Tone: Serious or Official</li> </ul>	<ul style="list-style-type: none"> <li>Start: Second account, between 14-15</li> <li>Called "Finsta", "Private", or "Spam"</li> <li># of Followers: 50</li> <li>Videos: More videos than Main</li> <li>Tone: Raw emotional connection - sad, angry, happy</li> </ul>

Document 202: META3047MDL-031-00086272, -6273

Both slides show that 11 years is the “age” at which, in Meta’s timeline, the typical child starts using Instagram (red boxes added) and, for what it’s worth, the age at which they typically create a “Finsta” account. “Finstas” are further addressed below in the context of parental controls.

592. Consider, too, an international qualitative study of 220 children that Meta conducted in 2017 (before Meta was collecting age data through Instagram). This study, entitled “Early Teen Illumination Research,” used a sampling frame of children aged 11-15 and the reported objective was to “understand early teens and create illumination foundation. Foundation

that will inspire high impact marketing messages and campaigns that drive Instagram platform production and engagement and bring in new users.”<sup>638</sup>



Document 203: META3047MDL-019-00059532 at Slide 21

<sup>638</sup> META3047MDL-019-00059532 at Slide 7

# JOINING INSTAGRAM



**INSTAGRAM SELLS ITSELF:**  
In 6<sup>th</sup> grade, age 11-12, teens get first phone and are "invited" in one of two ways...



**BFF**  
"My friend just got it and told me to download it."  
– Lilly, 12, LA



**TOP ACCOUNTS**  
"I want to see what my favorite YouTubers and soccer players are doing in real-life (off the screen... off the field)!"  
– Anthony friend group, 14, S. Brazil

*Document 204: META3047MDL-019-00059532 at Slide 26*

These slides demonstrate that Meta noted the age of initiation preceded 13 and in fact, "Instagram sells itself" to 6<sup>th</sup> grades ages 11-12.

593. Finally, in 2021, Meta conducted a study to ascertain "barriers" to using Instagram among 10–12-year-olds (see below).

**We have signal from tweens (10-12 yr olds) that they have preconceived barriers to using Instagram that rest mainly on the social pressure the platform presents.**

<p><b>Requires perfection</b></p> <p>● -----</p> <p>"I have Instagram, I rarely go on it...I don't really like it. <b>When you're trying to post a picture it always has to be perfect and stuff and I'm just not good at that...</b>"</p> <p>- Tween</p>	<p><b>Just for showing off</b></p> <p>● -----</p> <p>"...they could go to an expensive store and take a picture of expensive shoes, but in real life they don't really buy those shoes, <b>they just take a picture and post it on Instagram...to get people to like them and stuff.</b>"</p> <p>- Tween</p>	<p><b>Don't have anything "shareworthy"</b></p> <p>● -----</p> <p>"...we see our friends everyday at school and I <b>don't have anything I particularly need to share to anybody...</b>"</p> <p>- Tween</p>
---	--	---

How Tweens and Teens Make and Maintain Close Friends, Jimenez, Weber, Ng, An, April 2021

*Document 205: META3047MDL-020-00349969, -0038*

594. I have seen documents indicating that Meta leaders discussed strategies to get younger teens on their platforms “safely.” For instance, I have seen preparatory documentation motivating the launch of Facebook’s “Project Kid” which was a site designed with additional controls specifically for children under the age of 13.

## High Level Overview

- **Allow <13s to register on Facebook**, with varying levels of **parental approval** required depending on the kid's country
- Parents have **control over certain features** that allow their kids to form connections or have private conversations, and they have the option to disable their kid's account.
- Parents are **subscribed to updates** regarding their kids' friend requests/connections and actions that their kids take that they are allowed to see
- Parents can **access their child's account** and are considered joint administrators.
- More **restrictive privacy defaults/maximums** for kids – primarily a "friends only" experience
- Additional **product restrictions** for kids (e.g. age gating of objects, limited ads, etc.) to protect them from inappropriate content
- **Safety collateral** for parents and kids

a/c privileged

*Document 206: META3047MDL-034-00385870 at Slide 2*

Notably, although these controls were developed and deployable, when Meta scrapped the plans to roll out Facebook Kids because of public backlash, they neglected to offer some or all of those same safeguards to the parents of 13 year and older children.

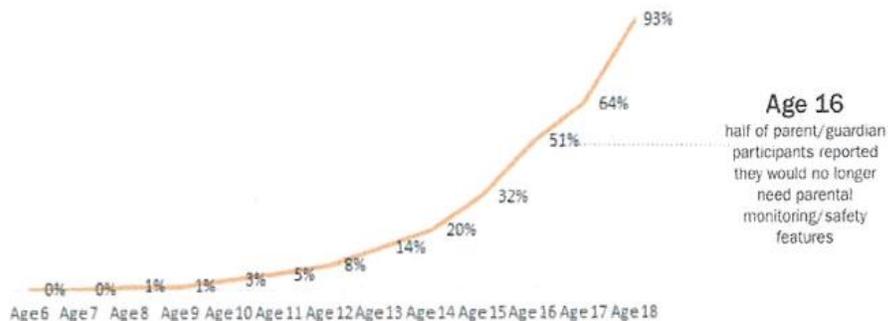
595. That is particularly concerning given Meta's awareness that over 70% of parents reported concerns about the content their tweens might be exposed to and the people they might meet on Meta's platforms.<sup>639</sup> Those parental concerns were further validated by internal Meta research led by Kramer in 2020, which surveyed over 3500 parents/guardians:

---

<sup>639</sup> Haugen\_00023849, -3870

## Parent/guardian participants wanted to monitor child app usage well into the teen years.\*

Age at Which Parent/Guardian Participants Would Feel Comfortable with Child Using Apps and Websites without Monitoring or Supervision Features  
(Sabino, June 2021)



\* Internal-led research with US sample, n=3,503 parents/guardians

Document 207: META3047MDL-020-00350316, -0360

596. In addition to demonstrating that parents “want[] to monitor child app usage well into the teen years,” this study is notable because the lower limit Meta used to query parental interest in monitoring was 6 years. Again, that is considerably below the current 13-year-old COPPA limit, suggesting Meta was exploring parents’ comfort levels for children well below the current COPPA age limit. In fact, Meta had plans to “build compelling products for users as young as 5-10 years of age”<sup>640</sup>

**4: Pre-teens are under-served and fast-growing.** We can build compelling social products for users as young as 5-10 and grow usage via their parents. Building for kids allows us to serve the social needs of future teens, help them establish their first social graph, and grow with them over time.

Document 208: META3047MDL-046-00241542, -1556

<sup>640</sup> META3047MDL-046-00241542, -1556

597. Meta’s failure to age gate (and/or its perfunctory approach to doing so) are particularly concerning in light of Meta’s understanding of the significant mental health harms that children on the platform can experience. In 2020, Thorn (a not for profit dedicated to child online safety) issued a report that was circulated within Meta. Among other findings, the report stated that “[c]hildren report having online sexual interactions at high rates...25 percent of kids 9-17 reported having had a sexually explicit interaction with someone they thought was 18 or older.”<sup>641</sup> Those numbers are consistent with Meta’s own. Meta reports that 50% of Instagram direct messages (DMs) to children come from adults, which is concerning to say the least.<sup>642</sup> Indeed, in June of 2020, an internal Meta communication reported the results of an internal study showing that **“500,000 IG underage accounts receive IIC on a weekly basis.** The IG prevalence is 3x Messenger’s.”<sup>643</sup> “IIC” is Meta’s acronym for “inappropriate interactions with children.”

598. As part of my review of internal exhibits, I watched exhibits from the Behar deposition in which a series of young underage girls sing in front of the camera and state their name, their age, and color.<sup>644</sup> It appears that they were all following a “trend” that Instagram provided to produce these videos. It is unclear to me why Instagram would permit so many underage children to remain on their site when they are clearly stating that they are underage – and some are very young. In other video exhibit to Bejar’s deposition, the videos show young girls either dancing suggestively or asking if they are cute - both videos elicits comments that are obscene and offensive.<sup>645</sup>

---

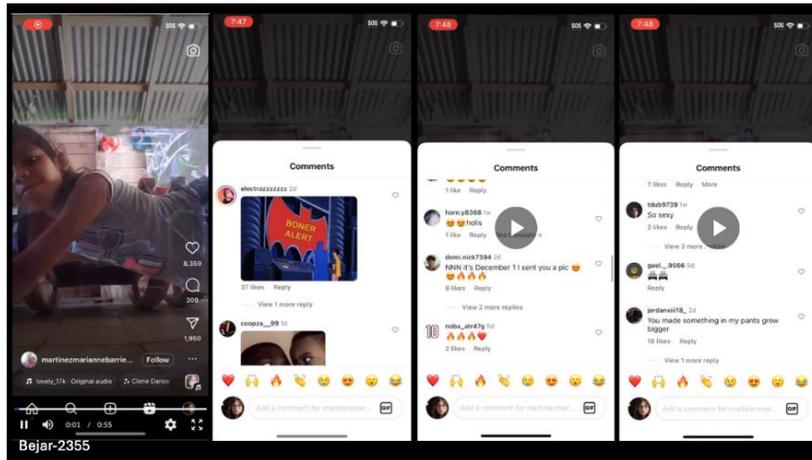
<sup>641</sup> META3047MDL-031-00245499, -5502

<sup>642</sup> Diego Castaneda Dep. Tr. at 280:9-12

<sup>643</sup> META3047MDL-014-00350154, -0159 (bold in original)

<sup>644</sup> See Arturo Bejar Dep. Exhibit 13 (BEJAR0002542)

<sup>645</sup> See Arturo Bejar Dep. Exhibits 13 (BEJAR0002542), 32 (BEJAR002355) 34 (BEJAR0002543)



*Document 209: Arturo Bejar Dep. Exhibit 32 (BEJAR002355)*

599. Furthermore, according to Meta’s TRIPS data, only 1% of people with a negative experience on Insta report it, and only 2% of those reports lead to action. (P0063.13) Given that “action” could entail as little as informing the parents of children in some format—including via parental controls—that their child had lodged a report, the failure to fulfil a minimal duty to warn is striking.

### Under 13 users and rewarding exploitative content

It is trivial to find large cohorts of teenagers under 13 using Instagram. Basic testing with Reels can create a feed made up almost exclusively of under 13 accounts with large reach, primarily posting content that sexualizes minors but is not CSAM.

Posts of minors that are sexually exploitative posts have tens or hundreds of thousands of views vs. more innocent content that has a few hundred views.

Anonymized public comments from posts by minor presenting accounts:

- GIFs that include “Boner Alert” and images implying masturbation.
- “Sexy.”
- “You should go get a chair and a rope.”

*Document 210: Arturo Bejar Dep. Exhibit 50 at Slide 16*

600. To put this in human terms, consider this example from a 2019 mixed-methods study commissioned by Meta of child users of Instagram in the Los Angeles area all of whom were over 13 when the research was conducted by Answer Lab, an independent contractor. One 14-year-old participant reported:

*Fake robot accounts DM you and say 'watch my free sex videos' and the person sends it to 50 different kinds of people...This was kind of like 'hey get a \$500 Visa gift card', I did that one time, it was the first time I ever saw it, I was really young, I thought, 'I am going to get some money!' It asked me for my address, plain old little me, I was kind of ditzy, and what it said to me after that was 'We have your address now, we will stalk you.' I was eleven, I freaked out...That's probably one of the reasons I got a private account. - P12, 14*

*Document 211: Diego Castaneda Dep. Exhibit 9 at -6904*

601. Notwithstanding that their own data validated Thorn's findings, in their public (and internal) response, Meta "thanked" Thorn, then pointed out that Thorn excluded Apple's iMessage from their study even though it is "bigger than Messenger and IG Direct Combined." While it might be true that iMessage is used more frequently than DMing on Meta, it is a texting platform that is not tied to any social media site. There is no built-in algorithm in iMessage that makes children's contact information available to potential predators. In this context, iMessage is closer to a proverbial "landline" which was not a commonly deployed pedophilic strategy. In any event, Meta's effort to divert attention to another platform is not persuasive, again given its own awareness of the problem.

**ii) TikTok**

602. On June 12, 2019, Rey Allie, Trust and Safety Strategy Expert at Tik Tok, asked Patrick Nommensen, Head of Global Public Policy for Bytedance, "if we are doing anything to verify that those who say there are over 13 actually are." Nommensen responded, "We do not

verify age beyond user input.”<sup>646</sup> Blake Chadlee, President for Global Business solution, provided further details in his response:

Please note the kids mode is enabled only in US.

And regarding your questions:

1. No, currently we only use age gate information to identify user's age. We do not have any third party service to verify the authenticity of how old a user really is.
2. Currently, no.
3. This should be divided into two parts:

##US

- For new user signup, yes, user can kill the app to re-start the age gate signup process.
- For existing users, no, as long as they selected an age and tap "Next", it will be an irreversible move for user. If they selected a wrong age, then they have to submit an appeal in order to change the birthdate.

##Non-US

- In TikTok-m, The age gate is disabled for China, Russia, Egypt, South Africa, Brazil. For other countries and regions, age gate is enabled for new user signup: user who selected 13- will be blocked during signup process. In UK, Germany, France, if the user selected 13- in the first time, then in the next 24 hours, no matter what birthdate user enters, TikTok will continue blocking users from signing up an account.

*Document 212: TIKTOK3047MDL-079-LARK-02017133, -7136*

603. There appears to be considerable variability in how (or if) age gating is present and enforced but it is readily gameable in all contexts. Presumably recognizing that the problematic nature of misrepresenting age gating is universal, Nommensen adds:

**Text:**

In the past we've been very cautious about not making reps that the age gate is global. This is generally not a detail need/should disclose. Focus more on our policy and strict adherence to local regulations and give an example, eg "for example we enforce this in the US and Europe with an age gate"

*Document 213: TIKTOK3047MDL-079-LARK-02017133, -7138*

---

<sup>646</sup> TIKTOK3047MDL-079-LARK-02017133, -7133

604. The ineffectiveness of Tik Tok’s age gating approach was acknowledged by Kristelle Collins, of the Youth Safety and Wellbeing team, in a text exchange on Aug 11, 2022. In her words, “I know I sound like a broken record on this, but I think we need to continually highlight that age gate is predicated on entering a birth date—this is a feeble safety precaution and needs to be understood for what it is.”<sup>647</sup> In a chat exchange on August 26, 2021, Amy Classen, Global Issue Owner at TikTok, stated that “10% of users are underage” and “[she’s] about to ban all of them.”<sup>648</sup> But it appears that even when the age is known at TikTok, some users—at least very popular ones, were not immediately banned. In his deposition, Han was asked about a particular 11-year-old creator with 4.1 million followers.<sup>649</sup> The child, whose age was confirmed as accurate by TikTok, was blacklisted but still able to post for at least five months.<sup>650</sup>

605. The problem of underage users may be especially pronounced at TikTok which distinguishes itself from its competitors in its marketing materials as having a younger (and more female) demographic. Almost 59% of their users are in the “14 and Under” category.

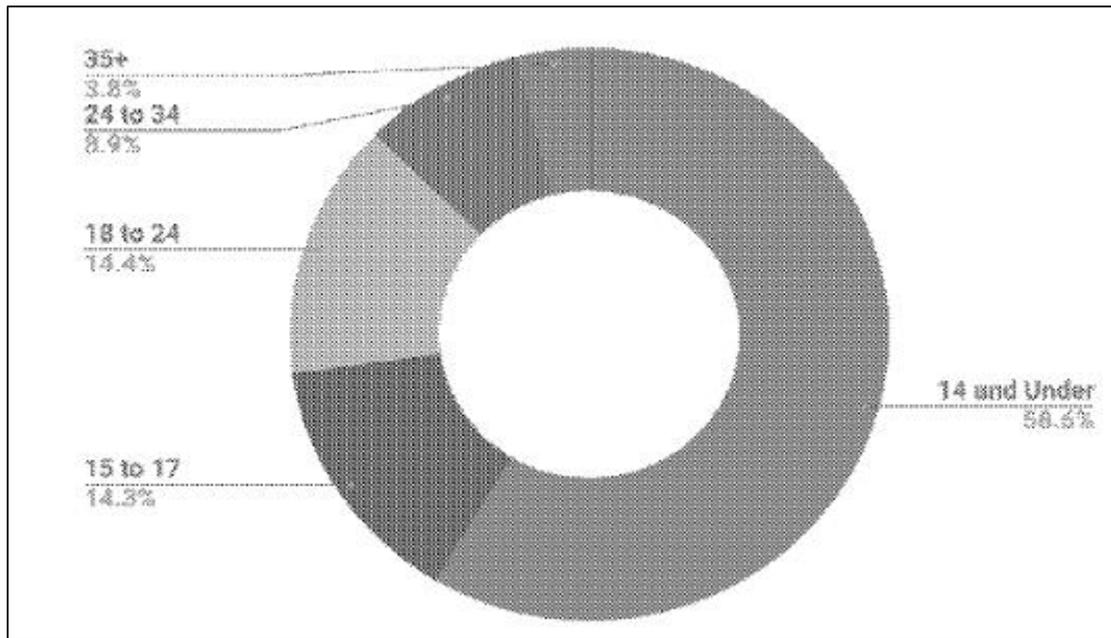
---

<sup>647</sup> TIKTOK3047MDL-021-LARK-00005510, -5516

<sup>648</sup> TIKTOK3047MDL-038-LARK-00192063, -2064

<sup>649</sup> Eric Han Dep. Tr. at 365:9-12

<sup>650</sup> Eric Han Dep. Tr. at 369:7-14



*Document 214: TIKTOK3047MDL-004-00290586, -0586*

As an aside, it is notable that the “Under 14” category does not have a lower bound. Every other age band is bounded at both ends except 35+ which understandably doesn’t have one since the upper age of human life expectancy is unknown. Given that being over 13 is “required” to have an account, one might expect it would have been labeled “13-14.” Nevertheless, the “14 and Under” category consumes more videos per day (225 on average) than any other group. Younger users are also more likely to share, which also drives their usage.<sup>651</sup>

---

<sup>651</sup> TIKTOK3047MDL-004-00290586, -0587

**iii) YouTube**

606. YouTube has determined that users under the age of 13 (u13) should not be permitted to access YouTube Main unless enabled by a parent or legal guardian (the “u13 Policy”).<sup>652</sup>

607. YouTube’s primary age-verification mechanism for users, including u13 users, is declared age.<sup>653</sup> But YouTube knows that using declared age is an unreliable metric and internal documents note that (“those who say they're under 18 are likely to be u18, but only a small fraction of those who are actually u18 are declaring accurately,”<sup>654</sup> and that “[d]eclared age isn’t a reliable signal to build teen specific models.”<sup>655</sup> Yet another YouTube document concluded that “[m]ost actual YT Teens users did not declare themselves between 13-17. This partially explains why only around 3-4% of the YT DAU are Teens users based on the declared age.”<sup>656</sup>

608. YouTube’s age verification is not only easily circumventable by entering a false birthday, but also prompts the user to “update” their birthday if its incorrect or use the birthdate of the “business owner” who manages the account.<sup>657</sup> The screen that appears when a user attempts to sign-up with an under-13 birthdate is below:

---

<sup>652</sup> YouTube, *Terms of Service*, Age Requirements, <https://www.youtube.com/t/terms#eb887a967c>.

<sup>653</sup> GOOG-3047MDL-04585554 (November 2023 document noting that “[i]n the United States and most of the ROW [rest of the world], YT currently relies on declared age.)

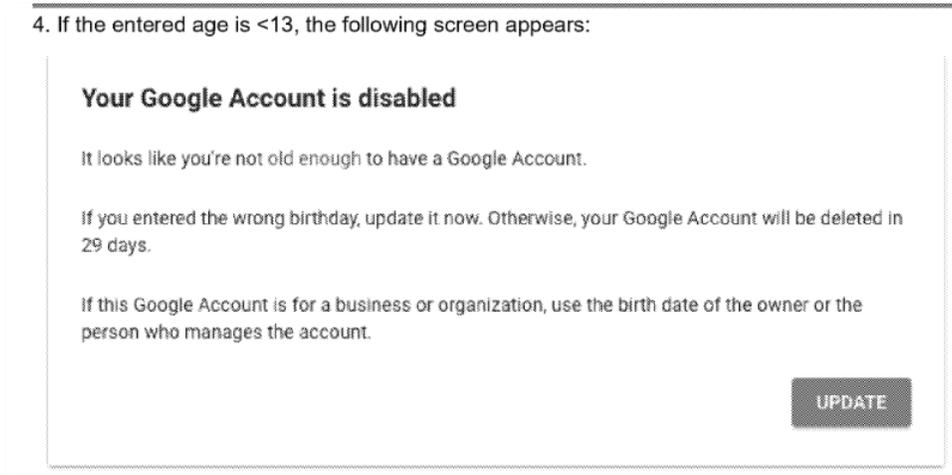
<sup>654</sup> Kim Ex. 18 (GOOG-3047MDL-04703742) at -742 (Oct. 21, 2022)

<sup>655</sup> GOOG-3047MDL-04683365 at -366 (July 24, 2023); *see also* GOOG-3047MDL-01339056 at -071 (Mar. 20, 2023) (“most actual YT Teens users did not declare themselves between 13-17”); James Beser Dep. Tr. at 515:15-17 (“[W]e do know that not every under 13 or under 18 is truthful.”).

<sup>656</sup> GOOG-3047MDL-01435767 at Slide 16

<sup>657</sup> GOOG-3047MDL-02027137 at -7139, -7140

4. If the entered age is <13, the following screen appears:



*Document 215: GOOG-3407MDL-02027137 at -140*

609. YouTube has some limited age verification. According to YouTube, “[i]f a user provides a birthdate that indicates they are above the age of consent [13 years old] at account creation, a combination of YT classifier and human review may later indicate that they may not meet our age requirements. In that case users can either provide a government-issued ID or credit card demonstrating that they are over the age of consent or add supervision to their account. If they fail to do so within 14-days, their account is disabled. YT disables hundreds of thousands of accounts each month as part of this process.”<sup>658</sup>

610. But the internal documents indicate that there are significant limitations to this review. YouTube documents indicate that suspected u13 accounts are flagged through the following processes: (1) human trust and safety reviews; (2) Cyclops Classifier for livestreams;

---

<sup>658</sup> GOOG-3047MDL-04585554 at -5555.

(3) Athena Classifier; and (4) external reports of u13s (anyone can flag a video or user and report to YouTube).<sup>659</sup>

611. Yet these tools appear to be designed in a manner that would limit their effectiveness. For example, YouTube documents state that the Athena Classifier is the most prominent avenue by which accounts are flagged for u13 review.<sup>660</sup> Yet, Athena Classifier only looks at user uploaded videos, not other forms of activity such as viewing or commenting on videos.<sup>661</sup>

612. As recently as February 2021, using its “Athena v3 classifier,” “YouTube flag[ged] 300K accounts per week as being suspicious for underage.” By this estimate YouTube scans only 0.3% of accounts daily.<sup>662</sup> As of 2021, Athena has allegedly “terminated over 35 million channels.”<sup>663</sup>

613. The review of livestream videos was similarly resource-constrained. In 2019, there was a two-month backlog of reviews that were queued through the Livestream Classifier:

---

<sup>659</sup> GOOG-3047MDL-01342809 (describing Athena Classifier and subsequent review process) (2024); Ostergaard Dep. Ex. 16 (diagram of suspected u13 age verification process); Adi Jain 30(b)(6) Deposition Ex. 1 (deposition notes on history of u13 detection at YouTube).

<sup>660</sup> GOOG-3047MDL-01342809 at -2809.

<sup>661</sup> GOOG-3047MDL-01342809 at -2809.

<sup>662</sup> GOOG-3047MDL-01342809, -2810

<sup>663</sup> GOOG-3047MDL-01342809, -2809

## Underage accounts - sustainable?

- We have a 60k backlog based on 2 months of enqueues from Livestream
- We can continue to terminate these accounts (85% actionability rate in Underage Account review) but is this what we want?
  - The kids will just go to other platforms where child safety is not top priority nor do these platforms have sophisticated abuse detection mechanisms
  - Long term, these kids are our future user base so we should not isolate them
- What experience do we want these users to have on Google?

Confidential & Proprietary



*Document 216: GOOG-3047MDL-00246776 at Slide 17*

It appears that YouTube was less concerned with getting through the backlog of accounts to review and more concerned that a review had an 85% actionability rate – in other words, 85% of accounts were deleted for being under the age of 13.

614. In recognizing that 85% of accounts flagged were actually u13, YouTube lists considerations and asks “is this what we want?”. While “[t]erminat[ion]” is one proposed option (bullet 2 above), that same bullet goes on to question, “is this what we want?” rationalizing that “kids will go on to other less safe platforms” and acknowledging they are YouTube’s “future base”—a powerful motivation to minimize or ignore the problem.

615. A subsequent slide proposed coming up with a way for kids as young as nine to post videos on the grounds that “[k]eeping kids safe on YT should not mean chasing them away.”<sup>664</sup>

---

<sup>664</sup> GOOG-3047MDL-00246776 at Slide 21

This was the tension at the heart of YouTube’s age verification efforts: even as YouTube recognized its legal obligation to keep u13 children off its platform, from a business perspective it was desperate to keep them as users. This is a powerful motivation to minimize or ignore the problem.

616. YouTube, too, recognizes the problem of “grooming” on its platform with 8% of minors claiming to have had a sexual interaction on it and 17K adult CI seeking comments removed daily (see below).

**External organizations** are raising alarm bells over the rising prevalence of online grooming & the enticement of minors - **1 in 7** minors are asked for nudes by strangers daily or weekly.

**8% of minors claim to have had a sexual interaction on YT (10-14% on other social media platforms).** Experts agree that while DM'ing is the biggest risk vector, predators can begin extorting minors through relationship building on YT before moving the conversation off platform.

- YT removes 17K adult CI-seeking cmnts daily
- **77%** of CSAM on Shorts is "self-gen" & we don't know if it is innocent/solicited
- YT is expanding the ways in which creators can engage leading to increased risk (e.g., @mentions, Posts, Shorts Comment Stickers)

Today we detect & remove comment-level grooming violations, but we have known detection gaps in account & convo level detection that can be exploited by bad actors.

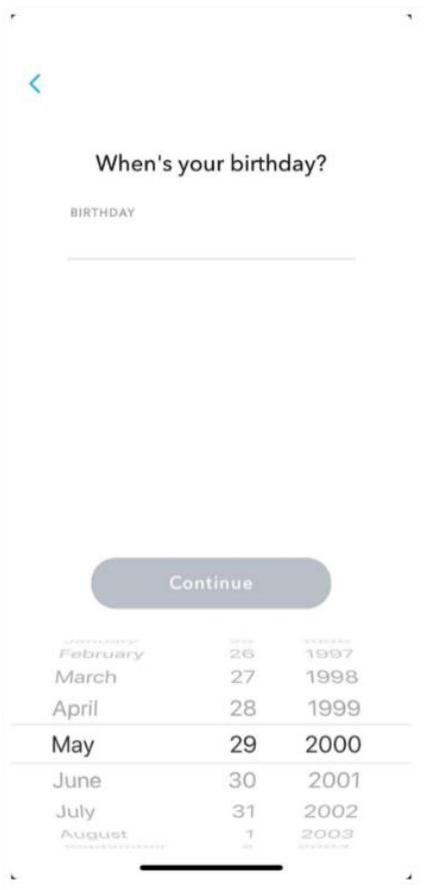
Lastly, if we don't prioritize this, regulations like the EU CSAM bill may force us to.

*Document 217: GOOG-3047MDL-00864164 at Slide 26*

The final sentence of this excerpt, which emphasizes that inappropriate interactions should be a high priority, goes on to say that failing to do so might result in increased regulation that might force them to.

**iv) Snap**

617. Snapchat’s age gate is similarly lax. Until 2016, Snap did not collect birthday information from its users at all. Since then, when signing up for the app, Snapchat users have been required to enter a birthday, and if they enter a birthday under the age of 13, they are not allowed to create an account. However, since 2017, the birthday entry screen has defaulted to eighteen years before that day’s date—essentially providing children with a pre-loaded fake birthday. An example of this from May 29, 2018 is included below.<sup>665</sup>



*Document 218: SNAP2367438, -67440*

<sup>665</sup> SNAP2367438, -67440

A prospective user can simply press “Continue” and be allowed to create a Snap account without ever having to affirmatively enter their birthday. Indeed, the eighteen-year-old default makes it easier to create an account and thereby increases Snap’s engagement metrics, including time spent.<sup>666</sup> Snap trialed truly neutral age gates, where there is no pre-selected birthday, for other markets, but it has never adopted this design in the United States.<sup>667</sup>

618. Josh Siegel, a former manager at Snap, described this default as a “minor design change” but internal Snap correspondence reveals that it was anything but.<sup>668</sup> So many children took advantage of Snap’s ready-made fake birthday that it created the appearance of a “dramatic drop” in the number of 13-17 year old users that was of “great concern” to Snap’s leadership.<sup>669</sup>

619. Snap’s ineffective age verification systems also mean that Snap is unable to prevent children under the age of 13 from using Snap.<sup>670</sup> At one point Snap tested a form of cookie that would lock new users out of the signup process if they entered an age below 13, but perversely, scrapped it because it was having too much of an effect on the number of user registrations and undercutting Snap’s growth metrics.<sup>671</sup> And once children under 13 are on Snap, Snap will only remove them if they affirmatively identify themselves to Snap as being under 13 or if someone else reports them. In 2021, Jennifer Stout, Snap’s Vice President of Global Public Policy, wrote

---

<sup>666</sup> SNAP3129214, -9214; *See also* Josh Siegel Dep. Tr. at 302:13-303:13:6-15

<sup>667</sup> SNAP6399042, -9042

<sup>668</sup> Josh Siegel Dep. Tr. at 288:20-22

<sup>669</sup> SNAP2367515, -7515 (“Nima found that our default birth year when a user signs up is 2000, which may explain a significant amount of what we’re seeing.”); SNAP2367438, -7438 (“I believe part of the problem is that default year in the registration flow is year 2000 and that’s why top birthday year for us is 2000 by far.”)

<sup>670</sup> SNAP2294924

<sup>671</sup> SNAP4833189, -3189 (“In the past, we’ve tested a session cookie that would lock out users from new sign ups if they entered an age below 13. However, the impact this had on new user registrations was so significant that the product team scrapped it.”)

that “we’re often asked what we do proactively [to delete underage accounts] and our answer truthfully is nothing.”<sup>672</sup>

**Q. Inadequate Parental Controls**

620. Parental controls are an important tool through which parents can monitor their kids’ usage of social media platforms and try to keep their kids safe. Importantly, parental controls, while important, should be complemented by increased industry safety standards. The limits of parental controls as an effective tool for safety are explained well by former Meta employee Ms. Jayakumar, who testified in her deposition: “I think parental controls are a complementary tool for child safety online. They cannot be the foundational tool. In addition to the vast impracticality of the suggestion, there is no evidence that every child has an engaged, knowledgeable parent with plenty of time to spare to monitor their child’s activity. It also kind of ignores the widespread prevalence of online experiences. Most of us are online for much more of the day than we are offline, and having a parent monitor every single minute of that would essentially be more than a full-time job.”<sup>673</sup>

621. Parental controls implemented by Defendants have been late in time, cumbersome, and minimally effective in their implementation.

**i) Meta**

622. Meta’s top executives were put on notice about the importance of parental controls as early as February 2009. That month, Mr. Zuckerberg was sent an email from one of Meta’s founding engineers, Jeff Rothschild, entitled “Let parents be parents on Facebook.”<sup>674</sup> Mr. Rothschild wrote: “I suspect that this feature may be somewhat controversial, so to limit the

---

<sup>672</sup> SNAP4833189, -3189

<sup>673</sup> Vaishnavi Jayakumar Dep. Tr. 434:11-24

<sup>674</sup> Mark Zuckerberg Dep. Ex. 91.

distraction, I'll bounce this off the three of you first to get some feedback on whether this is worth exploring further. I would like to see us add an opt-in feature, which would allow a Facebook user (child) to designate another user (the parent) to have certain auditing rights and limited controls over the child account.”<sup>675</sup> He went on: “The dynamic that this creates is to give parents an opportunity to act as parents on Facebook as they would in other dimensions of their children’s lives, shifting the primary responsibility for protecting and supervising children from Facebook to parents, which I believe is the only scalable and effective way to address the issues of minors on Facebook.”<sup>676</sup>

623. Despite this early warning, “parental controls, including tools developed for parental supervision of Teens, were first made available to users on Facebook in 2023”—fifteen years later.<sup>677</sup> And, despite acquiring Instagram in 2012, Meta did not make parental controls available on *that* platform until 2022, fully a decade later.<sup>678</sup>

624. While delay seems inexplicable, it makes sense when paired with documents indicating that Meta affirmatively sought to thwart parents’ supervision of their children’s use of its platforms—in order to ensure increased usage by young people. For example, Mr. Zuckerberg sent an email to other company executives in February 2016 discussing Facebook Live, a feature that allowed users to livestream video. After declaring his optimism for this product, he stated: “I’m worried that even if Live is a new raw format that young people enjoy, they may quickly migrate to a standalone product with a clean graph if we don’t fix our graph issues for this audience. That is, they may like Live, **but still not want to live stream to their parents** and all their FB

---

<sup>675</sup> Mark Zuckerberg Dep. Ex. 91.

<sup>676</sup> Mark Zuckerberg Dep. Ex. 91.

<sup>677</sup> Meta’s Amended Responses to Request for Admission No. 4.

<sup>678</sup> Meta’s Amended Responses to Request for Admission No. 3.

friends.”<sup>679</sup> He then went on: “If we tell teens’ parents and teachers about their live videos, that will probably ruin the product from the start.”<sup>680</sup> Finally: “My guess is we’ll need to be very good about **not notifying** parents / teachers.”<sup>681</sup>

625. Parental controls also ran contrary to the objectives of Instagram’s “Finsta growth team”:

- For people, especially teens, their Instagram profile represents their image online, and they are concerned about how they appear on their profile grid. One workaround is creating Finstas, the ultimate Identity & Audience Control.
  - WHAT WE’RE DOING: Favorites, Finsta Growth

*Document 219: META3047MDL-031-00086272, -6272*

As the name “Finsta Growth” implies, Meta actively promoted usage of finstas by teenagers:

- Finsta Growth is an effort on the Growth team to encourage teens to create their first Finsta account and to teach them to use the multi-account switcher.
- Future Opportunities: The team hasn’t explored teaching about Multiple Account Switching (MAS) in more depth, an opportunity to consider in the future.

*Document 220: META3047MDL-031-00086272, -6274*

626. Meta’s promotion of Finstas ran counter to parental supervision, as a January 2020 memo made clear: most parents “did not become aware of teen finstas/spam accounts, until long after the teen had created it.”<sup>682</sup> It is my opinion as a pediatrician and public health expert, that Meta’s promotion of fake Instagram accounts to teenagers is akin to a liquor store getting into the fake ID business.

---

<sup>679</sup> META3047MDL-014-00378779, -8779 (emphasis added)

<sup>680</sup> META3047MDL-014-00378779, -8780

<sup>681</sup> META3047MDL-014-00378779, -8780

<sup>682</sup> META3047MDL-034-00078516, -8516

627. I have reviewed a recent interview by Mr. Zuckerberg on the Joe Rogan podcast, in which he makes the claim that, “from a values perspective, where we should be is just trying to, like, be an ally of parents.”<sup>683</sup> During this interview, Mr. Zuckerberg touted a suite of new parental supervision features called “Instagram Teen Accounts,” which the company began rolling out in late 2024. However, despite these good intentions, Meta’s own internal data shows that only “0.38% of Youth users predicted to reside in the U.S. [sic] were enrolled in Supervision through Family Center on Instagram” as recently as March 24, 2025.<sup>684</sup> Whether Meta’s position as of 2025 is to be an ally of parents, it seems clear to me Meta did not take this approach earlier in its history, when I understand many of the children pursuing this litigation allege they were harmed.

**ii) TikTok**

628. TikTok’s efforts to eliminate or mitigate the impact of harms on underage users were either insufficient or not implemented due to competing growth concerns. Despite creating tools like parental controls, the company’s leadership acknowledged “we have awareness issues for multiple minor safety features, including restricted mode, parental controls, etc.”<sup>685</sup> An internal study concluded that while users could often find information about these tools, “their paths to the correct information were often indirect [sic] which suggests that participants often expected to find information in alternative locations.” The study further noted that “several items were relatively difficult for participants to find, including family paring, control who messages you and screen time management.”<sup>686</sup>

---

<sup>683</sup> Joe Rogan Experience #2255 – Mark Zuckerberg, <https://youtu.be/7k1ehaE0bdU> (1:49:16 - 1:50:35)

<sup>684</sup> Meta’s Sixth Supplemental and Amended Response to Interrogatory No. 12 at p. 39.

<sup>685</sup> TIKTOK3047MDL-004-00147649, -7658; *See also* TIKTOK3047MDL-115-04366552, -6557-58 (Reporting that “teens lack awareness of our screen time management offering” with a mere 0.6% usage rate.).

<sup>686</sup> TIKTOK3047MDL-004-00311638, -1640.

629. The need for more restrictive safety features was echoed by the app’s own users who reported that they “want[ed] restrictive solutions since permissive, ignorable, and unrestrictive tools are useless” to reduce harms they experienced on the platform.<sup>687</sup>

630. These issues were exacerbated by TikTok’s failure to prioritize these safety efforts compared to other growth-related initiatives.<sup>688</sup> Its internal documents reported that while it would “need to address incentives and make tradeoffs in order to avoid common points of failure” to address user wellbeing, its efforts were being impaired by “1) the lack of cross-functional cohesion caused by no shared definition of wellbeing and unclear decision-making processes, roles and priorities, and 2) a lack of resources and visibility” that made wellbeing work “mostly one-off, reactive, and [an inconsistent] priority across teams.”<sup>689</sup> These problems clearly persisted since as late as 2024, TikTok was relegating user wellbeing at the expense of engagement when it decided to launch “Streaks” despite “prior research on this feature on other platforms [finding] an association with anxiety, problematic overuse, and FoMo” and concluding that it “does not adhere to our current practice of promoting healthy digital habits for U18s.”<sup>690</sup>

---

<sup>687</sup> TIKTOK3047MDL-120-LARK-06208410, -8415-16.

<sup>688</sup> See e.g., TIKTOK3047MDL-021-LARK-00006955, -6959 (Permitting videos instructing children on how to circumvent Family Pairing because “[t]he fact is that the content doesn’t directly cause any direct harm to teens (physical, psychological, exploitation, or developmental).”); TIKTOK3047MDL-081-02351723, -1723 (Describing a product development cycle in which the Trust and Safety team’s safety recommendations are ignored due to time constraints.).

<sup>689</sup> TIKTOK3047MDL-002-00077113, -7136.

<sup>690</sup> TIKTOK3047MDL-150-LARK-07285061, -5064; See also Ivan Mehta, *TikTok is Testing Snapchat-Like Streaks*, TECHDIRT (June 6, 2024), [https://techcrunch.com/2024/06/06/tiktok-is-testing-snapchat-like-streaks/?guccounter=1&guce\\_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce\\_referrer\\_sig=AQAAAKEmu-yvF\\_SbcoCS5ipjNuvFSTOkk01TZX-v\\_C3mpK1rm4\\_B22dbW-UXTRFjEc0-77kDxjuR1baXCv4LDOxpNgI5C\\_E6Qq-NcJ5THxckI3o8tH\\_338rKDUzyrwcqZWJKM-XWFaqDA-b3IWx2PguJ19juUkNXv19IKtWWWhs0TZfF](https://techcrunch.com/2024/06/06/tiktok-is-testing-snapchat-like-streaks/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAKEmu-yvF_SbcoCS5ipjNuvFSTOkk01TZX-v_C3mpK1rm4_B22dbW-UXTRFjEc0-77kDxjuR1baXCv4LDOxpNgI5C_E6Qq-NcJ5THxckI3o8tH_338rKDUzyrwcqZWJKM-XWFaqDA-b3IWx2PguJ19juUkNXv19IKtWWWhs0TZfF).

### iii) YouTube

631. The history of parental controls on YouTube is confusing to say the least, and varies greatly depending on the level of parental control (device or app), the particular YouTube product (Main or Kids), and the operating system (Android or iOS). This constellation of measures would be very challenging for any parent to keep track of let alone deploy effectively.

632. Until 2017, there were no parental controls made available by Google or YouTube for YouTube Main. That year, Google introduced Family Link, which provided device level parental controls for children under 13 using Android devices.<sup>691</sup> Parental controls via Family Link did not become available for Chromebook, and could not be used for children above the age of 13, until 2018.<sup>692</sup> Family Link did not allow for app-specific time limits until 2019.<sup>693</sup> Parental controls, via Family Link, were not available for iOS until 2018 (and the iOS version continues to suffer from reduced capabilities).<sup>694</sup> The ability for parents to restrict the use of particular applications, including YouTube, during school hours was not introduced to Family Link until 2024, despite internal discussions regarding app-specific time-of-day restrictions since at least 2018.<sup>695</sup>

---

<sup>691</sup> GOOG-3047MDL-01621942, 1942-43; James Beser 30(b)(6) Dep. Ex. 1.

<sup>692</sup> Helping more families set digital ground rules with Family Link, September 18, 2018, <https://blog.google/technology/families/helping-more-families-set-digital-ground-rules-family-link/> (last visited April 11, 2025).

<sup>693</sup> The evolution of Family Link parental controls, May 7, 2019, <https://blog.google/technology/families/evolution-of-family-link-parental-controls/> (last visited April 11, 2025).

<sup>694</sup> James Beser 30(b)(6) Dep. Ex. 1; GOOG-3047MDL-05630293.ECM at -0294.ECM; FAQ, <https://families.google/familylink/faq/> (last visited April 13, 2025) (“Can children or teens be supervised by Family Link on iOS devices and web browsers?” “Children or teens signed into iOS, web browsers, or other unsupervised devices can only be partially supervised.”)

<sup>695</sup> James Beser 30(b)(1) Dep. Exs. 40, 41, 42, 43, 44, 45.

633. YouTube did not introduce its *own* (app level) parental controls for YouTube Main until 2021 with the launch of a “supervised experience” for “tweens” under the age of 13, internally referred to as “SupeX.”<sup>696</sup> Parental controls in SupeX are limited to: content settings, channel blocking, disabling autoplay, and the ability to disable or delete the child’s search and watch history.<sup>697</sup> Notably, the introduction of SupeX created a loophole whereby Android and ChromeOS could bypass the SupeX onboarding process and access YouTube Main using their EDU accounts.<sup>698</sup> Despite YouTube’s February 2021 announcement that it was introducing supervised experiences for “tweens *and* teens,” YouTube did not introduce parental controls for teens until 2024.<sup>699</sup> Parental controls for teens are limited to reviewing channel activity and information for teen users who post content.<sup>700</sup> This so-called “Supervised Experience for Teens” does not allow a parent to view a teen’s time watched statistics, which is a chart of the daily break down of time spent watching videos over the past seven days with an average calculation of time watched per day at the top.<sup>701</sup>

634. YouTube knew that parental controls in Family Link and YouTube didn’t provide parents an easy way to monitor and control their children’s screen time on YouTube.<sup>702</sup> For instance, Family Link Screen Time controls only applied to Android and Chromebook devices, despite YouTube’s knowledge that approximately 65% of U.S. users were on iOS.<sup>703</sup> Similarly, Family Link accounts and enforcement of YouTube app settings was limited to signed-in users on

---

<sup>696</sup> James Beser 30(b)(6) Dep. Ex. 1; GOOG-3047MDL-00000280.

<sup>697</sup> James Beser 30(b)(6) Dep. Ex. 1; GOOG-3047MDL-00000280.

<sup>698</sup> GOOG-3047MDL-01693424; *see also* James Beser 30(b)(6) Dep. Ex. 2.

<sup>699</sup> *Compare* GOOG-3047MDL-00000280 *with* James Beser 30(b)(6) Dep. Exs. 1, 4.

<sup>700</sup> James Beser 30(b)(6) Dep. Ex. 1.

<sup>701</sup> Raj Iyengar Dep. Tr. 82:21-86:9.

<sup>702</sup> GOOG-3047MDL-05214601, -4601.

<sup>703</sup> GOOG-3047MDL-05214601, -4601.

Android and ChromeOS devices.<sup>704</sup> Therefore, children accessing YouTube in a logged-out state (i.e. not signed into an account) on iOS could “circumvent policy restrictions and parental controls.”<sup>705</sup> And, of course, YouTube knew that child users could also input a false age, which would similarly allow them to view anything an adult accessing YouTube in this way could view.

706

635. It wasn’t until 2022 that concerns about regulation and competition caused YouTube to reevaluate its “inadequate” offerings, acknowledging that parents could not “access their child’s screen time controls” in YouTube.<sup>707</sup> This included SupeX.<sup>708</sup>

636. Parental controls on YouTube Kids have taken a different, though equally circuitous and incomplete path. YouTube Kids launched on February 23, 2015, but parents were not given the ability to block channels on YouTube Kids until the following year.<sup>709</sup> Two years later, in September 2018, parental controls were “added to the YouTube Kids app – allowing parents to handpick videos and channels in the app.”<sup>710</sup> However, at this point, members of the YouTube

---

<sup>704</sup> James Beser 30(b)(6) Dep. Ex. 1.

<sup>705</sup> GOOG-3047MDL-05630293.ECM at -0294.ECM.

<sup>706</sup> Woojin Kim Dep. Tr. at 229:22-230:32 (Kim testified that he was not aware of an age being assigned to a YouTube user accessing YouTube Main in a logged-out state); Matt Halprin Dep. Tr. at 169:5-25 (Matt Halprin was aware from discussions at meetings at YouTube that people were not honest about their ages when creating an account); James Beser Dep. Tr. at 491:23-492:1 (“So YouTube signed out is very easy to access, and it’s very likely that some of those users are under 13.”).

<sup>707</sup> GOOG-3047MDL-05214601, -4604.

<sup>708</sup> GOOG-3047MDL-05214601, -4604; GOOG-3047MDL-01195859.

<sup>709</sup> History of YouTube’s Responsibility Efforts. [www.youtube.com/howyoutubeworks/progress-impact/timelines](https://www.youtube.com/howyoutubeworks/progress-impact/timelines) (Retrieved April 26, 2024); GOOG-3047MDL-00000048; *see also* GOOG-3047MDL-01625570, -5570 (There was concern, at launch, that “[p]arental controls security in the app was very weak” and something kids could “easily work around”; however, the suggested response to this concern was simply to state that YouTube would listen to user feedback and continue to improve the experience).

<sup>710</sup> GOOG-3047MDL-00000922, -0928

Parent Panel had been “encouraging the YouTube Kids team to put more control in the hands of parents” for nearly the entire time YouTube Kids had been in existence.<sup>711</sup>

637. In April 2019, Google received a report it had commissioned from Fluent titled *Digital Wellbeing of Families*, based on a study designed to “examine the role digital technology plays in the wellbeing of families.”<sup>712</sup> The report described parents’ overall frustration with and lack of faith in parental control apps and tools.<sup>713</sup> It contained statements from parents that described two separate concerns: one, that YouTube was able to distract and keep their children quiet, but also the fear that the same YouTube videos might be addictive.<sup>714</sup>

638. Even so, YouTube delayed implementation of parental control of the Autoplay feature in YouTube Kids until 2021 when it was defaulted to “off” and parents given the option to deploy it.<sup>715</sup> Prior iterations of YouTube Kids did not include an Autoplay toggle option.<sup>716</sup> The rationale for not putting this decision in parents’ hands for over six years was that if a parent wanted a child to “watch 15 minutes of YouTube, but then every two minutes...had to go and play the next movie for them, then that would not be the experience that parents wanted.”<sup>717</sup> However, this meant

---

<sup>711</sup> GOOG-3047MDL-00080597, -0599.

<sup>712</sup> GOOG-3047MDL-05705191 at 5192.

<sup>713</sup> GOOG-3047MDL-05705191 at 5296, 5300.

<sup>714</sup> See e.g., GOOG-3047MDL-05705191 at Slide 46 (“My 15-year-old—I heard some noise. Walked into his room at two in the morning and he’s got a tablet in there watching YouTube videos. I’m like ‘Dude. You took this out of my room without permission.’ Stuff like that. It’s almost like an addiction.”); GOOG-3047MDL-05705191 at Slide 94 (“I think sometimes we’ve all used our devices to kind of distract our kids. We’re kind of taking away from parenting them because we just got to get something done and we know if we hand them a YouTube video to watch, they’ll watch that YouTube [video].”).

<sup>715</sup> GOOG-3047MDL-04922012, p. 14.

<sup>716</sup> GOOG-3047MDL-04922012, p. 14.

<sup>717</sup> Shimrit Ben Yair Dep. Tr. at 91:5-12.

that, if a child’s guardian did not set a timer on YouTube Kids, videos would continue to play forever without intervention.<sup>718</sup>

639. Years before YouTube granted parents the ability to control the Autoplay feature on YouTube Kids, internal documents acknowledged a study concluding the “most common reason parents provide kids with mobile devices is to distract them” and depicting the resulting cycle in which kids relied on this distraction as a coping mechanism and melted down when not given the device, ultimately eroding “parents authority [and] ability to set and enforce limits[.]”<sup>719</sup>



Document 221: GOOG-3047MDL-00408442 at Slide 15

<sup>718</sup> Shimrit Ben Yair Dep. Tr. at 93:17-21 (Videos would continue to play “[f]or as long as the parent allowed their kids to use the product.”)

<sup>719</sup> GOOG-3047MDL-00408442 at Slide 15.

640. Other internal documents evidence findings that “young kids are often the ones holding the tablet” and a preference by parents and children for larger screens.<sup>720</sup> In contrast, Shimrit Ben Yair, Product Manager Lead for YouTube Kids at its inception, testified that YouTube Kids was slated to begin on tablet devices because they “heard from parents that... iPads and tablets is typically where they watch content as a family.”<sup>721</sup> When asked whether this research regarding families gathered around tablets to watch video were provided to her in written form, Ben Yair cited her “own experience as a parent” as the guiding force for making YouTube Kids available on tablets first, ahead of other devices.<sup>722</sup> Eighteen months after its launch, YouTube Kids remained a smartphone and tablet only product, touted in advertising-related materials as “optimized for tablets.”<sup>723</sup>

**iv) Snap**

641. Snap’s only parental controls are a feature called “Family Center.” Family Center was only implemented in August, 2022, making Snap among the last of Defendants to implement parental controls.<sup>724</sup> Furthermore, the controls that Snap implemented through Family Center were extremely limited and barely used. The initial version of Family Center only permitted parents to view kids’ friends and recent conversations, without the ability to limit use or control account settings. Other designs considered would have allowed greater visibility into account settings and allowed parents to actually change kids’ settings, but that capability was removed at the direction of Snap’s CEO, Evan Spiegel.

---

<sup>720</sup> Shimrit Ben Yair Dep. Ex. 12 at Slide 43; *see also* Ben Yair Dep. Ex. 12 at Slide 6 (“TV is still THE ‘device’ most often used by kids [5-7 years old]” and Slide 19 (“Some parents prefer TV because: - Kids have better posture and sit further from the device – It’s easier to monitor”).

<sup>721</sup> Shimrit Ben Yair Dep. Tr. at 76:16-24.

<sup>722</sup> Shimrit Ben Yair Dep. Tr. at 106:3-21.

<sup>723</sup> Shimrit Ben Yair Dep. Tr. at 131:14-132:13; *see also* Ben Yair Dep. Ex. 14 at Slide 24.

<sup>724</sup> SNAP0017949



*Document 222: SNAP2619258, -9264*

While additional features were added to Family Center in 2023, allowing parents to see more parts of users' profiles, Family Center still does not allow parents to actually place guardrails on their children's use of Snapchat.

642. In 2022, when family center was deployed, internal documents reflected that Snap recognized that "there are users in the system who falsify their age to be older than 18, when in fact they are younger."<sup>725</sup> At the time, the idea of notifying parents if their child's age in-app was recorded as 18+ was floated.<sup>726</sup>

643. Later, Snap did formative qualitative research with parents to determine what their concerns about their teens' use of SM (presumably to remediate them). In July 2023, an internal PowerPoint's dominant theme was "tech loopholes," which Snap defined as "even when thoroughly supervised, their teens would be able to find ways around monitoring."<sup>727</sup> And later in the same report, the following slide appears:

---

<sup>725</sup> SNAP0010984 at -0010985

<sup>726</sup> SNAP0010984 at -0010985

<sup>727</sup> SNAP0019103 at SNAP0019110.

- Parental control tools and an enhanced reporting system to manage concerns or monitor suspicious activities were among the more popular potential actions Snap could take to give parents increased supervisory agency in their teens' activities.
- More rigorous age verification and retaining messages on Snapchat's servers were popular potential actions that most parents wished Snapchat would have already implemented, though many acknowledged operational loopholes and difficulties associated with putting them in place.

*Document 223: SNAP0019103 at -9118*

644. I find no evidence that Snap alerted parents about this “loophole” in their parental monitoring system. To the contrary, the age defaults when signing up were set to make everyone above age.<sup>728</sup> In fact, SNAP was worried about the “growth impact of implementing some sort of block on retrying registration.”<sup>729</sup> But in thinking of ways to make age verification more effective and secure, Snap identified “risks” to that strategy, specifically that “regulators might demand that [they] expand [their] usage of age estimate tools to other, more consequential parts of Snapchat.”<sup>730</sup>

645. Additionally, the number of families actually using Family Center is miniscule. Ten months after it launched, only 0.33% of kids actually used Family Center.<sup>731</sup> Perhaps driving this very low adoption of Family Center is the fact that it's only available when a child's reported age

---

<sup>728</sup> SNAP2382505.

<sup>729</sup> SNAP3664412

<sup>730</sup> SNAP3664412 at -4414

<sup>731</sup> SNAP0017949, -7952

is under 18.<sup>732</sup> But Snap is not only aware that a significant number of children lie about their age and say that they are older in order to get on Snap, it actively facilitates that lying by making the default age at sign-up 18—a default that makes users ineligible for Family Center.<sup>733</sup> Snap doubles down on this policy by not notifying parents if the reason their Family Center request has been rejected is that their child’s reported age is over 18. As a Snap employee pointed out in 2022, this creates the impression that Snap is “prioritizing a user’s right to essentially falsify their age (something we ostensibly don’t support) over a parent’s ability to utilize the Family Center functionality.”<sup>734</sup>

646. Overall, as Abby Tran, the Snap product manager responsible for launching Family Center, wrote in 2022, “if we are asking the question of ‘does this feature give parents everything they need to improve safety for their teen’ - obviously not.”<sup>735</sup> In internal communications, Tran was forthright about the fact that this limited utility was by design. The point of Family Center was to provide just enough features to create a public talking point, without actually doing anything that might affect kids’ use of Snap. Snap’s image, not children’s safety, was paramount.

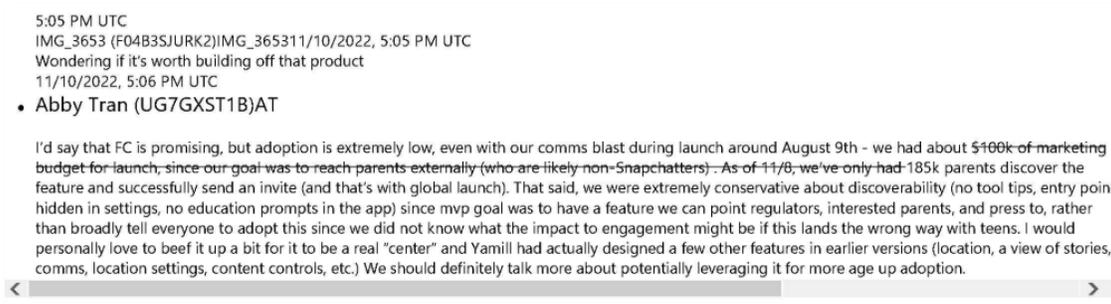
---

<sup>732</sup> SNAP0010984, -0985, *Why Can’t I Access Family Center*, SNAP, located at <https://help.snapchat.com/hc/en-us/articles/8132746171796-Why-can-t-I-access-Family-Center#:~:text=And%20as%20a%20reminder%2C%20it's,to%20accept%20Family%20Center%20invites> (last accessed Apr. 14, 2025).

<sup>733</sup> SNAP0010984, 0984, *Why Can’t I Access Family Center*, SNAP, located at <https://help.snapchat.com/hc/en-us/articles/8132746171796-Why-can-t-I-access-Family-Center#:~:text=And%20as%20a%20reminder%2C%20it's,to%20accept%20Family%20Center%20invites> (last accessed Apr. 14, 2025).

<sup>734</sup> SNAP0010984, -0984

<sup>735</sup> SNAP1837692, -0984



*Document 224: SNAP1186209, -6211*

**v) Exogenous Parental Controls Efforts**

647. Teen's desire for autonomy coupled with their emerging curiosity about "adult" content makes restricting their access to digital spaces challenging. As discussed above, extant parental control features of social media platforms demonstrate minimal uptake in large part because of design features that make them cumbersome or hard to understand. Given the shortcomings in the parental controls made available by Defendants themselves, and in light of the fact that the vast majority of US parents have concerns about their pre-teen and teenagers' screen use,<sup>736</sup> experts (including myself) generally recommend providing additional, effective parental controls to mitigate the risks to children and teenagers.

648. I define "exogenous" filters as those that are not app specific and commercially available as third-party, stand-alone solutions for parents to deploy whereas "endogenous" filters are ones that apps deploy or make available for parents. Endogenous filters and their limitations are discussed in the immediately preceding sections. As for exogenous ones, a SmithMicro random digit dial survey of 2000 US parents of children 5-18 years of age revealed that 90% of parents use

---

<sup>736</sup> META3047MDL-020-00349969, -0013

digital parenting technology to manage their children’s internet activity and 86% of parents report they have regular talks with their children about online safety.<sup>737</sup>

649. In spite of parents’ efforts, teens have proven adept at using a variety of techniques to bypass common controls, including VPNs to redirect web traffic, incognito or private browsing windows, altering device or app time settings, or creating “fake” accounts altogether to evade detection or monitoring (e.g. finstas discussed in Section XI.C.(i) above). Almost ½ (45%) of parents reported that their child had attempted to disable or bypass parental controls.<sup>738</sup> That percentage is especially worrisome given the ease with which children can disable many existing SM controls. A 2015 study of over 15,000 parents found filters to be essentially of no utility in screening content.<sup>739</sup> Further, a systematic review of 40 studies examining the effectiveness of existing digital technologies to moderate children’s screen use showed minimal to no effect.<sup>740</sup>

650. One major limitation of exogenous filters is their imprecision in identifying inappropriate content within apps themselves. These tools are primarily designed to block access to entire websites (e.g., pornographic sites) or prevent the installation of or limit access to specific apps. However, they offer minimal—if any—visibility into what teens are actually exposed to inside those platforms, such as sexually explicit direct messages, grooming attempts, inappropriate

---

<sup>737</sup> Software S. *Understanding Digital Parenting*. 2021. [https://info.smithmicro.com/hubfs/Surveys/eBook\\_Smith-Micro-Digital-Parenting-Survey.pdf?hsLang=en-us](https://info.smithmicro.com/hubfs/Surveys/eBook_Smith-Micro-Digital-Parenting-Survey.pdf?hsLang=en-us)

<sup>738</sup> Przybylski AK, Nash V. Internet Filtering and Adolescent Exposure to Online Sexual Material. *Cyberpsychology, Behavior, and Social Networking*. 2018/07/01 2018;21(7):405-410. doi:10.1089/cyber.2017.0466

<sup>739</sup> Przybylski AK, Nash V. Internet Filtering and Adolescent Exposure to Online Sexual Material. *Cyberpsychology, Behavior, and Social Networking*. 2018/07/01 2018;21(7):405-410. doi:10.1089/cyber.2017.0466

<sup>740</sup> Stoilova M, Monica B, and Livingstone S. Do parental control tools fulfil family expectations for child protection? A rapid evidence review of the contexts and outcomes of use. *Journal of Children and Media*. 2024/01/02 2024;18(1):29-49. doi:10.1080/17482798.2023.2265512

images, or sextortion schemes. This is why platform-level controls, starting with reliable and enforceable age verification mechanisms, are so critical to teen safety.

651. In conclusion, while parents play a vital role in monitoring and guiding their children's use of digital technology, tech companies have an essential—and in many cases, indispensable—responsibility to support these efforts. This is especially true for parents who lack the technical know-how or financial means to access and manage third-party software solutions effectively.

#### **R. Other safety features**

652. A review of Defendants' documents demonstrates multiple missed opportunities to deploy effective safety features. For example, self-limiting tools, such as Meta's "Take a Break" feature, were tested to ensure they did not reduce engagement too much. Likewise, failure to provide parents and teenagers the ability to select default limits on the length and frequency of sessions, or to block time of usage during the day, such as nighttime or school facilitated over usage. As discussed above, distracted learning and disrupted sleep are particularly harmful for developing adolescents. In my opinion, based upon my clinical experience, medical training, and the academic literature, tools that help reduce time on the apps will decrease the risk of harm. This is particularly true for any tool that decreases usage of social media platforms during the school day or at nighttime.

##### **i) Lack of default limits on the length and frequency of sessions**

653. The absence of default time limitations on these social media platforms creates an environment where pre-teens and teens, whose prefrontal cortex and self-regulatory capacities are still developing, can engage in prolonged and frequent usage patterns that significantly increase their vulnerability to addiction-like behaviors and associated mental health harms. Without built-

in constraints, these platforms effectively rely on the external regulation that children and teens have not fully developed. It's analogous to expecting the proverbial kid set loose in a candy shop to volitionally limit what they eat. It is an unreasonable and ineffective expectation.

**ii) Defective opt-in restrictions to the length and frequency of sessions, FTW**

654. To the extent they exist, current opt-in restriction models for social media usage represent a fundamentally flawed approach to protecting vulnerable pre-teens and teens. By defaulting to unlimited engagement and requiring active self-limitation, platforms effectively place the burden of protection on the very individuals whose developmental stage makes them least equipped to exercise such judgment. Unlike other products with known risks to developing minds, these platforms provide minimal transparent communication about potential psychological harms, leaving adolescents and their caregivers inadequately informed about documented risks. The defective nature of opt-in time restrictions becomes evident when examined through a developmental lens. These mechanisms incorrectly assume teens and pre-teens possess the same risk assessment capabilities and impulse control as fully developed adults. This misalignment between platform safety design and neurobiological reality disproportionately impacts teens and pre-teens. Platforms deliberately selected opt in (vs opt out) because they knew uptake would be less and the overall impact on the number of daily average users and time online would be minimally impacted.

**iii) Defective self-limiting tools**

655. Inadequate self-limiting tools on social media platforms represents a critical failure point in protecting adolescent mental health. When these tools are difficult to access, unintuitive to operate, or inconsistently implemented across features, they fail to provide the protection necessary for developing minds—contributing to increased anxiety, depression, and diminished

psychological well-being. From a clinical perspective, the defective nature of existing self-limiting mechanisms directly undermines pre-teens and teens developing capacity for healthy self-regulation. These poorly designed tools create a false sense of protection while simultaneously exposing teens and pre-teens to algorithmic engagement strategies engineered to override impulse control, thereby exacerbating vulnerability to mental health harms, including compulsive usage patterns. The current implementation of self-limiting features on major social media platforms demonstrates a concerning disregard for developmental science. By designing ostensible protective measures that are easily circumvented, frequently reset, or buried within complex settings menus, platforms effectively nullify their utility for the population most in need of protection—contributing to documented increases in social comparison, sleep disruption, and attention difficulties among adolescent users.

**iv) No blocks to usage during certain times of day**

656. The absence of default time-of-day restrictions on social media platforms creates significant vulnerability during critical developmental periods. Without automated evening limitations, adolescents—who already experience biologically-driven delayed sleep onset—frequently engage with stimulating content during pre-sleep hours, potentially disrupting circadian rhythms and reducing both sleep quality and quantity, which research has consistently linked to compromised emotional regulation, cognitive performance, and mood stability. From an educational perspective, the unrestricted availability of social media during school hours and designated study periods represents a substantial barrier to academic engagement and cognitive development. The constant accessibility of highly stimulating, dopamine-rewarding content creates an attention competition that developing brains are neurobiologically disadvantaged to

resist, potentially contributing to documented decreases in sustained attention, comprehension, and academic performance.

657. Without time-specific usage limitations, platforms effectively undermine parental and educational boundary-setting efforts, creating digital environments that can disrupt essential activities including family interactions, academic engagement, and the consolidated sleep necessary for optimal psychological functioning.

**v) Defective barriers to deactivation/deletion of accounts**

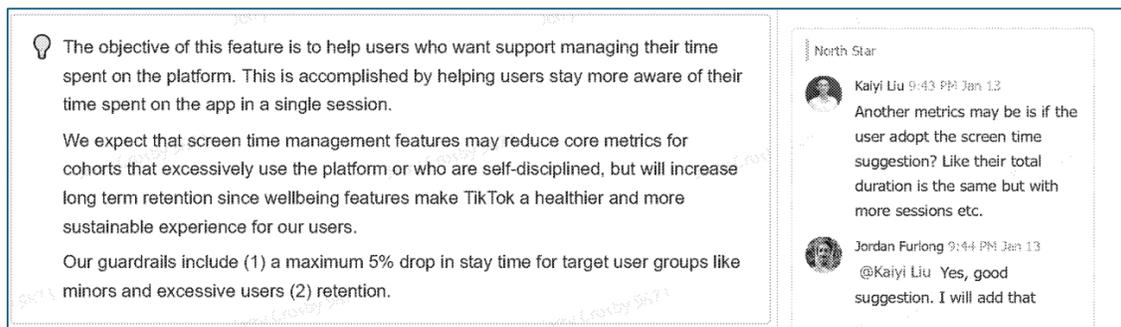
658. The implementation of complex, multi-step account deactivation and deletion processes creates significant obstacles for adolescents attempting to disengage from potentially harmful social media use. These convoluted exit pathways exploit developing executive function capabilities, potentially prolonging exposure to platforms that clinical evidence suggests may be contributing to psychological distress for vulnerable pre-teens and teens. From a developmental perspective, the deployment of emotional tactics during account deletion attempts—including messages about friends who will 'miss' the user—exploits adolescents' heightened sensitivity to social evaluation.

659. By designing systems that capitalize on pre-teens and teens' powerful biological desire to "fit in," platforms effectively undermine genuine attempts at self-regulation, potentially extending exposure to psychological harms including social comparison, anxiety, and addiction-like usage patterns. The implementation of easy entry paired with high-friction exit disproportionately impacts developing users. This imbalance effectively traps adolescents in digital environments increasingly associated with negative mental health outcomes while simultaneously undermining their developing sense of digital autonomy and self-efficacy.

660. In that regard, limits on length and frequency of social media sessions, blocks on usage during critical hours of the day, and making it easier for children and parents to delete and deactivate accounts would help reduce harm to children. It is incumbent upon the social media companies to make these safety features be the default; requiring children or their parents to identify and maneuver a complicated “opt-in” process will decrease the number of adolescents who use these safety features. Former Meta employee Volichenko recognized this phenomenon at his deposition when he testified that opt-in features are implemented by fewer users than opt-out features.<sup>741</sup> Mr. Zuckerberg said the same thing.<sup>742</sup>

661. There are some examples of Defendants offering such tools, albeit in a fashion that prioritized continued engagement over real reduction in harm. For example, Meta noted in the “Teen Mental Health Deep Dive” that teens found time spent tools “easy to ignore.”<sup>743</sup>

662. TikTok was no different. TikTok also introduced screen time management tools, as noted below:



*Document 225: TIKTOK3047MDL-001-00060515, -0520*

<sup>741</sup> Volichenko Dep. 106:6-107:1237.

<sup>742</sup> Zuckerberg Dep. 237:21-23 (“Stats show us that most people just use whatever the default setting or filter is”).

<sup>743</sup> Gross Dep. Ex. 12 at -1773.

663. However, even as they introduced what was deemed a “safety” feature, the communication here (and elsewhere) references “guard rails” related to a potential reduction in the amount of time spent and the overall retention of users. The term “guard rails” in this context is a bit ironic. One normally associates them with safety features on highways that are erected to prevent drivers from going off a cliff or crossing a lane into oncoming traffic: money spent to reduce risk to lives. At TikTok they are intended to preserve revenue: lives put at risk to make money.

664. As it turns, out, uptake of the initial screen time management system was exceedingly poor (0.29%) although the percentage of TikTok’s users who said they would use them was high (20-43%) leading TikTok to explore building out their offerings to make them more appealing.<sup>744</sup> True to form, those changes were subjected to A/B testing prior to being introduced with the following “guardrails” set:

<b>Guardrail</b>	Observe potential trade-offs and impacts	StayDuration/U	Fluctuate
		Last 30-day ActiveDays/U	Fluctuate
		Publish/U	No effect
		Average session duration from 10pm-6am	Small decrease
		App uninstall rate	Small decrease

*Document 226: TIKTOK3047MDL-001-00060515, -0520*

<sup>744</sup> *TIKTOK3047MDL-001-00060515, -0518*

**Estimated Stay Duration Impact**

**1.5% of global stay duration, to be verified in A/B testing**

- 1.5% is an global expected stay duration decrease given minors represent 10% of the TikTok user base
- Teens are ~8% of DAU in US and EU; disproportionately high stay duration (~10% of total); disproportionately low revenue (~6% of total)
- Estimates are based on extrapolating behavior from users who opt-in to screen time limits to all teens, so we expect efficacy to be slightly lower for new users as they have not decided to enable this feature
- Impact may decrease based on teens' likelihood to disable feature or exceed limits
- Impact may increase if age assurance efforts reassign current 18+ users as teens

**Commented [10]:** Since these users are ads- touching users, st loss could also lead to revenue loss, so we may need some alignment on this topic between tiktok core ads and monetization side.

**Commented [11]:** We are at the same time considering significant restrictions on personalized ads for teens to comply with upcoming European and US laws (and address public pressure), so there is some overlap of effect

**Commented [12]:** Agree, could we estimate these 2 things' total impact to monetization?  
 Josh Stickler : [OK] 2023-02-14 11:06:01  
 Laura : [FISTBUMP] 2023-02-14 11:07:18

Document 227: TIKTOK3047MDL-010-00329585, -9596 (emphasis added)

Comment 10 above (yellow highlighting added) immediately cites the need for alignment with the “monetization” side since these are “ad-touching users.”

665. The results of the A/B testing regression analysis indicated that the revamped Screen Time Management tools are estimated to reduce the daily average stay for minors by about 10 minutes on weekdays and 15 minutes on weekends which immediately begs the question “if that effect size is “acceptable.”<sup>745</sup> It is honestly unclear in this context if acceptable means big enough or too big. Even after the initial testing was completed, the plan was to run a “holdout test to measure the effect that Screen Time Management features have on long term retention.”<sup>746</sup>

666. Tik Tok’s internal assessment of the uptake of all of their control features reveals just how ineffective they all are.

<sup>745</sup> TIKTOK3047MDL-004-00151111, -1111

<sup>746</sup> TIKTOK3047MDL-001-00060515, -0526

Feature	DAU / Penetration	Problems to Address
Screen Time Management	1.8m / .29%	<ol style="list-style-type: none"> <li>1. Revamp change pin code/forget pin code process (Jordan doing in 7-8 bimonth)</li> <li>2. There is little transparency into how much time users spend on the platform and therefore how they should self-regulate usage</li> <li>3. Restrictions are not enforced on all platforms. Need to sync on uid level instead of uid+did (Jordan doing in 7-8 bimonth)</li> <li>4. People don't know about this feature</li> </ol>
Restricted Mode	600k / .11%	<ol style="list-style-type: none"> <li>1. Restricted Mode controls for inappropriate content based on Risk labels, Warning Tags, and three rounds of moderation, but <u>not</u> labeling for specific age appropriateness. (Lufan working on this)</li> <li>2. Content filtering only applies to FYP. Users can still search for and follow links that are shared to content that would be filtered out of the FYP by Restricted Mode</li> <li>3. Content filtering is not enforced on all platforms, so users of Restricted Mode can continue to see unrestricted content</li> <li>4. People don't know about this feature</li> </ol>
Family Pairing	180k (parent) 278k (teen) / .08% (total)	<ol style="list-style-type: none"> <li>1. Families do not use Family Pairing. Based on user research, it appears that a lack of awareness may be suppressing usage and perception.</li> <li>2. Family Pairing doesn't address parents' top concerns like inappropriate content, offensive interactions, and lack of privacy</li> <li>3. Teen users can disable Family Pairing without PIN, so it's less restrictive than activating either Restricted Mode or Screen Time Management individually.</li> </ol>

Document 228: TIKTOK3047MDL-006-00325873, -5885

667. Daily active user penetration ranged from 0.08% to 0.29%. From a developmental perspective, relying on teens as young as 13 to self-regulate their usage, or “opt in” to more controls runs counter to what every pediatrician, psychologist, neuroscientist or even parent knows. They lack the foresight, the discipline, the cognitive capacity, to exert self-control like adults do (recall the brain development reviewed above). As Lee’s presentation to Facebook leaders from 2020 states, “Teens don’t think deeply about safety risks until something bad happens.”<sup>747</sup>

<sup>747</sup> Alison Lee Deposition Exhibit 4 at Slide 12

668. In summary, SM platforms simultaneously created and deployed sophisticated engagement mechanisms and rudimentary, often tokenistic self-limiting features. This unbalanced digital environment for developing minds is exceedingly and unnecessarily hazardous especially for some children.

#### **XIV. Conclusion**

669. As a leading expert on the effects of digital media on children—with more than 25 years of experience as a pediatrician, researcher, public health scientist, chief science officer, and journal editor—it is my considered judgment that social media platforms are both contributing to and intensifying harm in millions of children in the United States and globally. Specifically, they are partially responsible for the alarming rise in depression, anxiety, sleep disturbances, body dysmorphia, eating disorders, suicide and self-harm, and school-related difficulties. These effects are widespread, though not evenly distributed. Certain subgroups—those already vulnerable—are disproportionately affected. Tragically, social media algorithms have amplified this inequity, often identifying and targeting children based on their susceptibility. As a result, children who engage with harmful content—knowingly or not—are frequently shown more of it, due to platform design elements operating without their or their parents' knowledge or consent.

670. It is no surprise—at least not to me—that platforms engineered by some of the brightest computer scientists and behavioral experts in the world, under business models focused on maximizing user engagement, have proven especially addictive to young users. While some external experts dispute the conclusiveness of the research, in my view—and in that of many scientists in the field—the available evidence, combined with well-established psychological theory, supports a causal link between social media use and adverse outcomes in youth. Critics often highlight that many studies are cross-sectional and that effect sizes are modest. Both points

are true. Yet there is also a growing body of longitudinal and experimental research demonstrating harm. Moreover, the principle of differential susceptibility tells us that population averages obscure significant impacts on the most vulnerable—and even small effect sizes, when applied at scale, translate into harm for millions of children.

671. Despite this, the pace of scientific discovery has been slowed—deliberately so—by social media companies’ refusal to cooperate with independent researchers. Their own internal documents and analyses, many now public, acknowledge the harms their platforms pose to children. Yet time and again, they have failed to act in the best interest of their youngest users. Instead, they have rolled out superficial safety features and minor algorithmic tweaks, often admitted to be more about public relations positionings than meaningful protection. Even the most basic tools, such as parental controls, have been poorly designed and implemented, resulting in minimal uptake and no attempts to rectify them. Engaging children and adolescents was not a side effect—it was a growth strategy, pursued aggressively and competitively across the industry.

672. In sum, these platforms were consciously and systematically engineered to maximize engagement and growth—at the direct expense of children’s well-being.