

Filed Under Seal

Exhibit Q

**IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA**

PEOPLE OF THE STATE OF CALIFORNIA,
et al.,
Plaintiffs,

v.

META PLATFORMS, INC, Instagram, LLC,
Meta Payments, Inc., Meta Platforms
Technologies, LLC.,
Defendants

IN RE: SOCIAL MEDIA ADOLESCENT
ADDICTION/PERSONAL INJURY
PRODUCTS LIABILITY LITIGATION

THIS DOCUMENT RELATES TO:
4:23-cv-05448

MDL No. 3047

Case No. 4:23-cv-05448-YGR

Judge: Hon. Yvonne Gonzalez Rogers

Magistrate Judge: Hon. Peter H. Kang

TRIAL REPORT OF ARVIND NARAYANAN
May 16, 2025

CONFIDENTIAL

of signals as implicit or explicit, and the same signal may be implicit or explicit depending on the user's level of knowledge of the implications of the action. However, there is a clear spectrum of explicitness: search terms, "see more/less like this" buttons, and written reviews or comments are highly explicit; likes are less explicit; and dwell time is generally considered implicit. Dwell time refers to the amount of time the user views or lingers on a post before moving on to the next one.

119. Meta uses implicit signals in its recommendation algorithms. For example, the Instagram Reels algorithm uses thirteen consumer-oriented signals, six of which are "explicit interaction based signals" and five of which are "implicit time-spent related signals" (as well as other signals not directly related to the user).⁶⁶

120. Implicit signals lead to discovery of unstated interests because they are more likely to be the result of "System 1" (impulsive) than "System 2" (deliberative) cognitive processes.⁶⁷ By the very nature of impulsive behavior, users are less likely to be able to take steps to control those inputs to the algorithm. System 1 and System 2 are a concept relevant to several disciplines – including cognitive psychology, behavioral economics, and artificial intelligence – which describes two human modes of thinking, one (System 1) instinctive and emotional, and the other (System 2) deliberate and logical.⁶⁸

B. Exploiting behavioral proclivities

121. Many users report not only that they spend a large amount of time interacting with algorithmically curated feeds that they later regret, but also that these feeds become more homogenous over time in ways they did not intend. In this section I explain why this behavior is a plausible consequence of the way these algorithms are designed.

⁶⁶ META3047MDL-014-00405874. The remaining two consumer-oriented signals are described as "learning transferred signals with models transfer-learned from non-tab data."

⁶⁷ Moehring et al. Better Feeds: Algorithms That Put People First. Knight-Georgetown Institute. Mar. 4, 2025. <https://kgi.georgetown.edu/research-and-commentary/better-feeds/>.

⁶⁸ *Id.* One former Meta employee describes these systems as "considered goals" versus "base instincts" that all humans possess and explains that humans want to "develop in ways that reflect our higher-order goals" like "staying healthy and supporting our family, [...] and not our base instincts to eat a cookie or McDonalds or whatever." Simons Dep. 97:4-17.

CONFIDENTIAL

i. Engagement optimization

122. Engagement optimization prioritizes the content that users are most likely to engage with.⁶⁹ To understand the consequences of this logic, it is important to keep in mind that the average engagement rate on every major social media platform is low. The engagement rate of a post is the number of engagements (likes, comments, and so forth) divided by the number of views.

123. According to third-party engagement statistics, Instagram and Facebook have engagement rates under 1%.⁷⁰

124. Low engagement numbers are not at odds with the fact that social media platforms are experienced by users as highly engaging and potentially addictive. It is important to note that even when a user enjoys a post, they may not engage with it in a manner that can be captured by third-party engagement statistics, such as liking or commenting. Besides, even if a post is correctly predicted to be of interest to the user, it may not interest them at a given moment.

125. Users might differ in terms of how often they engage with posts, but consider a user who is in the typical range, engaging with about 1-3% of posts in their feed. So if there is a particular type of content with which they are predicted to be 10% likely to engage, that content will score extremely highly compared to the alternatives. In other words, even if the user resists their impulse to engage with such content 90% of the time, but gives in to the impulse just 10% of the time, that will be more than enough to ensure that the algorithm ranks such content highly when creating their feed. With the rise in the importance of implicit signals, this does not even require active engagement such as a like, but simply dwelling on posts for a few extra seconds.

126. In short, the goal of engagement maximization algorithms is to maximize the time that users spend on the app (over some time horizon), and the social media platforms I examined

⁶⁹ META3047MDL-072-00032552; META3047MDL-047-00706929; META3047MDL-047-00706686; META3047MDL-072-00139365; META3047MDL-046-00328644.

⁷⁰ SocialInsider. 2025 Instagram Benchmarks. <https://www.socialinsider.io/social-media-benchmarks/instagram>. Last visited Mar. 14. 2025.; SocialInsider. 2025 Facebook Engagement Benchmarks. <https://www.socialinsider.io/social-media-benchmarks/facebook>. Last visited Mar. 14. 2025.

CONFIDENTIAL

in their feeds — which in turn influences the “inventory” of posts on the platform and consumption of those posts.⁸⁰

136. Some creators are particularly attuned to the incentives of the recommendation algorithms because Meta offers them monetary incentives by splitting advertising revenue with them.⁸¹

137. If the recommendation algorithm learns to maximize engagement by triggering addictive behavioral patterns or amplifying extreme content, creators will be incentivized to create content that satisfies these demands. Many types of creators ranging from children’s video channels⁸² to politicians⁸³ have been known to modify their posting behaviors over time in response to algorithmic incentives in ways that are harmful to users and society.

C. Problematic use as a consequence of algorithm design

138. Meta’s engagement-optimized ranking algorithms often operate in ways that functionally mirror the psychological mechanism of *intermittent variable rewards*, an idea that is foundational to persuasive technology.⁸⁴

139. Imagine a vending machine that learns your stated and unstated preferences (or weaknesses) over time. But instead of simply giving you your favorite snack every time you press a button—which would quickly become predictable and dull—it introduces an element of surprise.

⁸⁰ META3047MDL-014-00405864.

⁸¹ Scanlon. What it takes to get paid by YouTube, TikTok and other social platforms. Digiday. Feb. 28, 2025. <https://digiday.com/marketing/what-it-takes-to-get-paid-by-youtube-tiktok-and-other-social-platforms>; Instagram Help Center. About monetization for ads in Instagram profile feed. <https://help.instagram.com/5485466918184985>.

⁸² Townsend. ElsaGate: The Problem With Algorithms. United States Cybersecurity Magazine. <https://www.uscybersecurity.net/elsagate/>.

⁸³ Uberti. How Fox News dominates Facebook in the Trump era. Vice. Apr. 29, 2019. <https://www.vice.com/en/article/how-fox-news-dominates-facebook-in-the-trump-era/>; Facebook internal report. “Political Party Response to ‘18 Algorithm Change.” Apr. 1, 2019. <https://fbarchive.org/doc/odoc889111>.

⁸⁴ Fogg. Persuasive Technology: Using Computers to Change What We Think and Do. (Interactive Technologies). Morgan Kaufmann. Dec. 30, 2002; Eyal & Hoover. Hooked: How to Build Habit-Forming Products. Portfolio. Nov. 4, 2014.

CONFIDENTIAL

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 "Arvind N".

Arvind Narayanan, Ph.D.