

AMENDED Exhibit 108

PLAINTIFFS' OMNIBUS OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

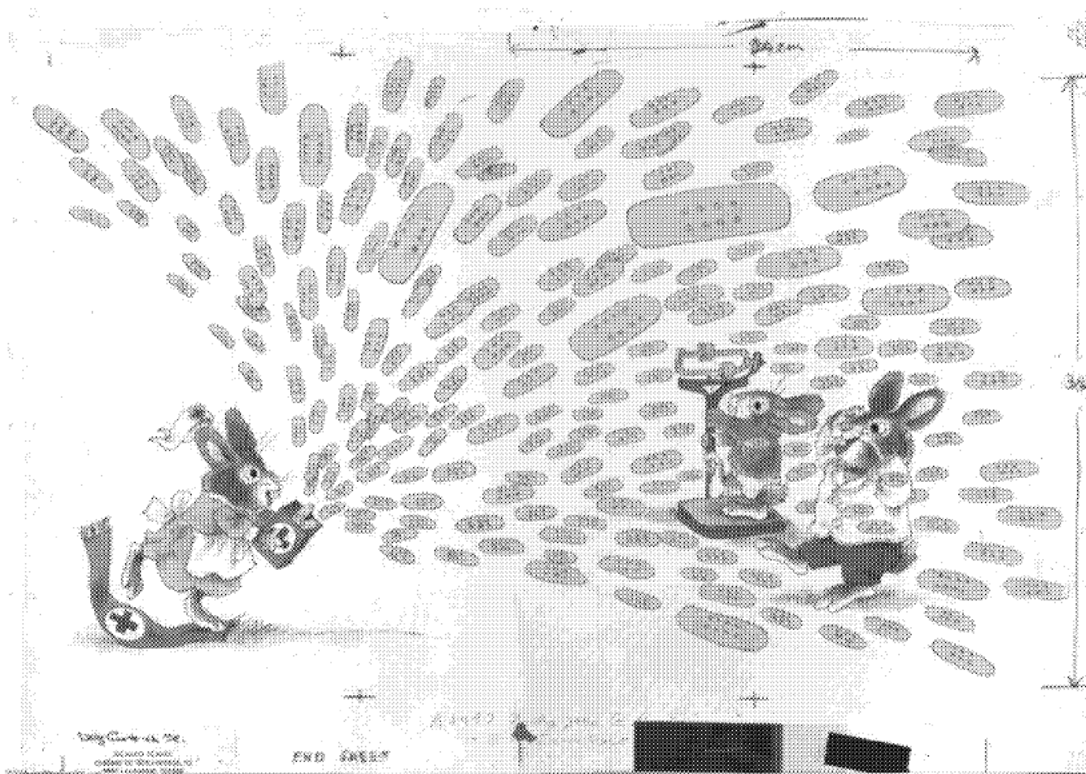
Case No.: 4:22-md-03047-YGR

MDL No. 3047

In Re: Social Media Adolescent Addiction/Personal Injury Products Liability Litigation

2019-09-20/ DRAFT:

User-Engagement vs User-Value



Many thanks for comments: [REDACTED], [REDACTED], [REDACTED], [REDACTED]

TLDR

1. We often make product decisions based on the effects on user-engagement (e.g. sessions, clicks, etc.) — and are uncertain about the

impact on user-value — i.e. what's the effect on the true utility and value that users get from using the product?

2. **There is a great deal of existing work on this, I'm going to try to summarize & organize some from the perspective of economic decision-making.**
3. **Here are three broad reasons why increasing user-engagement might not increase user-value, in this framework:**
 - a. **First, when the product *stretches* value out** — i.e., it spreads value thinly, to maximize engagement (sometimes called the Spence distortion).
 - b. **Second, when the product is misleading** — i.e., the product makes promises about value without fulfilling them.
 - c. **Third, when users' choices don't represent their own best interests** — e.g. when users have self-control problems.
4. **Q: How could we make product decisions that maximize user-value instead of user-engagement?**
 - a. **Choosing engagement metrics carefully can help:** E.g., (1) goaling on sessions, instead of time-spent, has some positives for the "stretching" problem, although it also may have negatives with respect to self-control issues; (2) capping metrics — e.g. only counting time-spent up to the first hour — so we don't incentivize growing time-spent for heavy users.
 - b. **Monitoring sentiment or WTA-value metrics can help** — however they're often too noisy to make regular product decisions with.
 - c. **A strong quality culture probably helps.**

Background

1. **There are a lot of critiques of social media that we maximize engagement at the expense of value..**
2. **We are providing value.** (1) People spend a huge share of their time on Facebook apps — something like 1% of worldwide time-spent; (2) when you ask people to put a monetary value on Facebook it's high: the average American says it would take \$50-100 before being willing to give up Facebook for a month[13]; (3) when you ask people where their meaningful social interactions took place, 1/10 were on Facebook, and another 1/10 were on WhatsApp (from a survey in Japan, India & the US, [12]).
3. **However there are reasons to be concerned.** Most importantly the relationship between usage of Facebook and psychological wellbeing is pretty murky: higher FB use is associated with worse wellbeing in the cross-section, though not in the time-series (a common finding is that *active* use of FB is positively correlated with wellbeing, but *passive* use is negative correlated). In a recent experiment, the only good one I'm aware of, they found that people who took a 1-month break from using Facebook showed a small but significant increase in self-reported psychological wellbeing.
4. **Why would engagement and value diverge?** Basic economic reasoning would imply that people engage with a product to the degree that they get value from it,

so engagement and value should be closely connected. I think this is broadly true – but there are a few theoretical reasons why value and engagement can diverge.

(1) Stretching Out Value (AKA Spence distortion)

Spreading the jam thin. Suppose you have a fixed amount of value to offer but you're trying to increase engagement, e.g. time-spent, then you can do this by spreading the value out: cut the value up into chunks, and put some filler in between.

A couple of economic papers in the 1970s by Michael Spence formalized this point: if you're making choices about the design of your product then your incentive is to maximize the *quantity* of demand – i.e., the number of people who want it, or the time-spent using your product – rather than the *quality* of demand, i.e. the total value delivered.

[formal model]

Spence applied this reasoning to television, comparing cable TV to broadcast TV. Cable TV is funded by charging a subscription, while Broadcast TV is funded by advertising. For this reason there's an incentive for broadcast TV to stretch out their value in a couple of ways: (1) they will try to make a show more broadly appealing, even at the expense of value to the core fans (they know the core fans will still watch it, and watch the ads, so it doesn't matter if they're less-happy with it); (2) they will try to increase viewing time on a show – e.g. by having recaps and cliffhangers — even if it makes the overall experience less valuable.

Stretching value could occur in Facebook products in a few ways.

1. **In notifications we limit the amount of information, because we want people to come onto the site.** Our notifications have moved towards an *uninformative* model — we now give notifications like “someone commented on your post,” without telling you who, or giving the text of the comment. We discovered with experiments that this increases engagement, though it very likely harms user-value.
2. **Lots of content on News Feed withholds information.** With titles like “*you'll never guess...*”, or “*one amazing trick...*” The titles try to get you to click through. We have a model which tries to detect these tricks and apply a demotion, though in general this catches only the extreme cases, and I think you can still see the effects of this incentive in many places: even high-quality newspapers craft their headlines to leave out some tantalizing detail.
3. **Maximizing for time-spent incentivizes bad ranking.** If we evaluated ranking by the effects on time-spent then it would *not* be optimal to rank just by expected

value – instead we would be better off by adding some randomness or filler, so that the value is stretched out. It's hard to tell how important this is for ranking in practice, because it's a subtle thing to measure and because ranking is a black-box, but I'd be surprised if it doesn't happen.

4. **More friction for high-value things.** Suppose that someone really wants to perform an activity – e.g. there's a video that they really want to watch – then any feature or experiment can increase engagement, by adding more friction, while lowering value – e.g. we can add more ads, interstitials, promotions.

[note: how does choice of ad-load fit into this framework? to what degrees are incentives aligned?]

(2) Misleading Users

A second reason for divergence between engagement & value is misleading experiences. If you give the impression to users that they're going to get some value by engaging with your product, but you don't actually give users that value, that will introduce a discrepancy between engagement & value.

This happens with Feed links all the time. Headlines will say misleading things to get you to click, and . We have a classifier and demotion for exaggerating headlines, but it only catches extreme cases.

If we blindly followed AB tests to maximize engagement, we'd lie to users all the time. The simplest way to maximize engagement is just to lie to users, e.g. by saying "click this link for guaranteed long battery life", or "accept our terms and we promise not to track you," even if it's not true.

Why doesn't this process self-correct? You'd expect that consumers will learn when firms mislead or exaggerate, and adjust their expectations, giving firms an incentive to tell the truth. This doesn't work perfectly in practice for a couple of reasons: (1) the cost of disappointed expectations is shared – when users have a hard time keeping track of who's responsible for the content they're seeing they'll blame *all* publishers, and *all* app developers, for disappointed expectations. This weakens the incentive to tell the truth, because the costs are borne by other companies. (2) many features are evaluated just on *short-run* engagement – if we mislead users it'll surely cause a lack of trust down the line, but that can be slow and cumulative, and it's very hard to detect with just short-run experiments.

(3) Self-Control Problems

Finally there are sometimes discrepancies between what people choose and what gives them the most value. The classical economic point of view is that a person's choice perfectly represents their preference – they're the same thing.

However there are a number of cases where that seems overly simplistic – e.g. when people have self-control problems, or cognitive problems in processing information.

There are reasons to worry about self-control and use of our products. I think the biggest problems people have are: (1) compulsive use, to fill gaps in their time, and they end up spending more time than they'd like, instead of spending time with their families, or studying; (2) unhealthily fixating on other people – most prototypically, compulsively looking up the profile of your ex-partner. [1]

A quick rundown of evidence:

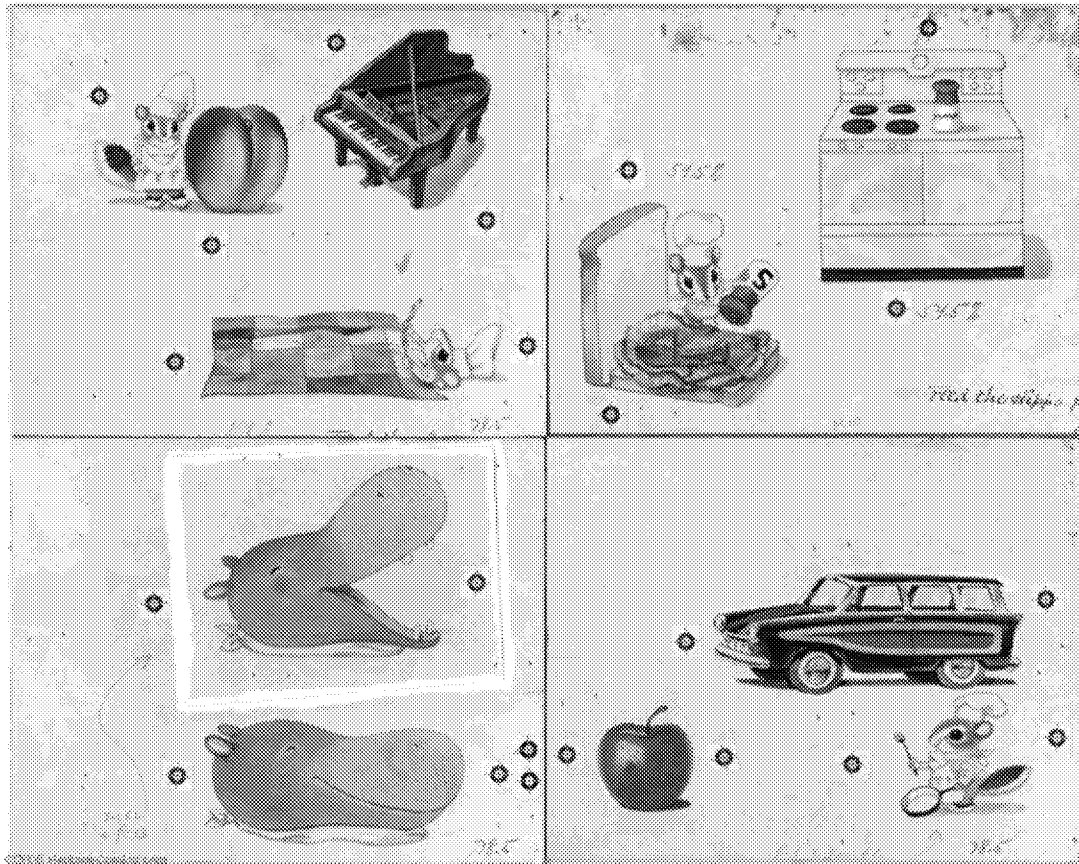
1. **Higher Facebook use is correlated with worse psychological states** (wellbeing, loneliness, etc.), although *changes* in Facebook use are uncorrelated with changes in wellbeing measures. [8,5]. It's also worth noting that higher Facebook use is typically positively associated with higher MSI, and is associated with people saying that Facebook has more value.
2. **An experiment found that a 1-month break from Facebook improved self-reported wellbeing.** [6] There is some legitimate doubt around how to interpret this experiment, unfortunately it's the only high-quality experiment that I'm aware of.
3. **A large fraction of users struggle with their Facebook/Instagram use.** Among US teens "84% had taken a break from social media, and 74% have taken a break from Instagram." [2] "More than a third of the survey respondents feel they have spent more time than they would have liked on Instagram. About one in five respondents think they have a problem controlling their time." [5]
4. **A significant minority report "problematic use".** 3.1% of Facebook users in the U.S. report "serious problems with sleep, work, or relationships that they attribute to Facebook AND concerns or preoccupations about how they use Facebook." [1]

What do Do?

Despite lots of work I think it's still an open problem, of how to make sure that maximizing user-engagement doesn't harm user-value. I'm planning to spend some more time thinking about this, in the meantime here are a few observations:

1. **Choosing engagement metrics carefully can help.** –

- a. **Meaningful engagement vs engagement.** Since early 2018 many FB-app products have been goaling on “meaningful” metrics.
 - b. **Sessions vs time-spent.** Over the last couple of years News Feed slowly switched from maximizing time-spent towards maximizing sessions. This should be good from the point of view of the Spence distortion: the frequency of sessions should be roughly proportional to the total expected value, so there’s less incentive to *stretch* value out to maximize time-spent. However this could be bad from the perspective of self-control problems: we actually find that the frequency of sessions is a strong predictor of problematic use, and our own internal researchers recently recommended we should “help people consolidate their use of Facebook into fewer sessions.” [1]
 - c. **Per-item survey metrics.** (WYT, pStar45,)
 - d. **Capped metrics.** From discussion with [[HYPERLINK "https://quip.com/MFXAEAzD1cT"](https://quip.com/MFXAEAzD1cT)], having a “capped” seems a very good idea: that we .
2. **Monitoring sentiment metrics could help.** – e.g. CAU (“Does Facebook cares about users?”), GFW (“Is Facebook good for the world?”). My impression is that they’re often too noisy to give guidance to everyday product decisions, even for half-length holdouts.
 3. **Economic Valuation metrics could help.** — We could monitor changes in consumers WTA (“what is your minimum willingness to accept for not using Facebook?”). This would directly address the Spence distortion, though less clear that it would help with problems #2 and #3.
 - a. (Note: could address #3 with intertemporal choices — e.g., choose again after finishing the experiment, or state your *future* WTA for using FB.)
 4. **A strong quality culture could help.** Most Facebook product decisions, large and small, are based on experimental effects on on engagement metrics. It does not have to be this way — other high-quality apps have been built with far less experimentation — e.g. WhatsApp. [a little more about likely effects]



Imaged by Heritage Auctions, HA.com

References

[1] Cheng, Burke, Davis, "Understanding Perceptions of Problematic Facebook Use", [HYPERLINK "https://www.dropbox.com/s/bph75wxf1u27hnr/problematic_use_2019-01-07_final.pdf?dl=0" \h]

- See also summary at [HYPERLINK "<https://our.internmc.facebook.com/intern/research/p/6gcvu18gbp/>" \h]
- 3.1% of Facebook users in the U.S. experience "Serious problems with sleep, work, or relationships that they attribute to Facebook AND concerns or preoccupations about how they use Facebook (e.g., a fear of missing out (FOMO) or lack of control)"
- "people who reported problematic use spent significantly more time on Facebook than people who did not ... they spent 21.6% more time on the site ... had 13.5% more distinct sessions."

- We should “Help people consolidate their use of Facebook into fewer sessions, with more control over notifications.”

[2] ██████████, 2018, “Taking a Break from Tech” – [HYPERLINK "https://fb.workplace.com/groups/ResearchInsights/permalink/1824949684220305/" \h]

- “84% [of US teens] had taken a break from digital technology in the last year”
- “84% had taken a break from social media, and 74% have taken a break from Instagram.”

[3] ██████████ et al., “The uses and abuses of Facebook: A review of Facebook addiction” - [HYPERLINK "https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4189307/" \h]

- Review finds 9 studies of FB addiction, but methodologically inconsistent.

[5] ██████████ (2018) “User Perception of Time-Spent on Instagram”

- [HYPERLINK "https://fb.workplace.com/groups/432156973909421/permalink/482477452210706/" \h]
- “survey in May 2018 with mid to high time spent users on Instagram,”
- “More than a third of the survey respondents feel they have spent more time than they would have liked on Instagram. About one in five respondents think they have a problem controlling their time.”
- 42% of the survey respondents have taken a break from Instagram. The top reason for taking a break is “I had something at work or school that I needed to focus on.” More high TS users report that the break has made them feel worse than before than low TS users.
- People who spend longer time think their time spent is more meaningful than those who spend shorter time. Those who spend longer than 3 hrs/day have lower positive affect, lower life satisfaction, and higher negative affect than people who spend shorter time on Instagram.

[6] Allcott et al. “The Welfare Effects of Social Media”

- [HYPERLINK "https://web.stanford.edu/~gentzkow/research/facebook.pdf" \h]
- “Deactivation caused small but significant improvements in well-being, and in particular in self-reported happiness, life satisfaction, depression, and anxiety ... Our overall index of subjective well-being improved by 0.09 standard deviations. As a point of comparison, this is about 25-40 percent of the effect of psychological interventions including self-help therapy, group training, and individual therapy, as reported in a meta-analysis by Bolier et al. (2013).”
- (after deactivation, persistently lower use)

[7] Burke & ██████████ (2016) “The Relationship Between Facebook Use and Well- Being Depends on Communication Type and Tie Strength”

- [HYPERLINK "https://onlinelibrary.wiley.com/doi/full/10.1111/jcc4.12162" \h]
- only looks at *changes* in wellbeing over a month, doesn’t look at

[8] ██████████ et al. (2018) “A Longitudinal Analysis of Facebook Use and Well-Being”

- “people who use Facebook more have lower well-being compared to people who use Facebook less”
 - No relationship between time-spent and *change* in well-being
- [9] ██████████, “More time spent per session is associated with a lower risk of problematic use”
- Findings:
 - a. $\text{is_problematic} \sim \text{sess_per_day} (+.01) + \text{controls}$
 - b. $\text{is_problematic} \sim \text{session_per_day} (+.007) + \text{ts_per_session} (-.003) + \text{controls}$
 - that, when you control for both, session-frequency positively associated with problematic use, and session-length negatively associated.
 - [HYPERLINK "https://fb.workplace.com/groups/558331381190793/permalink/903051943385400/" \h]
 - notebook @ [HYPERLINK "https://our.internmc.facebook.com/intern/anp/view/?id=132931" \h]
- [10] ██████████ (2019) “The Good, the Bad, and the Murky: Initial Findings on Session Quality”
- [HYPERLINK "https://fb.workplace.com/notes/██████████the-good-the-bad-and-the-murky-initial-findings-on-session-quality/484659979043924/" \h]
- [11] ██████████ (2019) “Facebook’s Superpower”
- [HYPERLINK "https://fb.workplace.com/notes/██████████facebook-superpower/780275142321216/" \h]
- [12] ██████████ Moira, ██████████ (2018) “We’re Meaningful”
- Survey in US, India & Japan, not branded. Finds that 1/10 meaningful interactions were on FB, and another 1/10 on WhatsApp.
 - [HYPERLINK "https://fb.workplace.com/notes/██████████were-meaningful/753700794978651/" \h]
- [13] ██████████ (2019) “Notes on the Economic Impact of Facebook”
- [HYPERLINK "https://fburl.com/cmk66hwl" \h]
- [14] ██████████ et al. (2018-oct), [HYPERLINK "https://quip.com/4JJQArloTuTg" \h]
- Recommendations: (1) cap sessions metrics at 8-15 sessions; (2)
- [15] Haidt & Twenge, “Social Media & Mental Health”
- [HYPERLINK "https://docs.google.com/document/d/1w-HOfseF2wF9YlpXwUUtP65-olnkPyWcgF5BiAtBEy0/edit" \h]