Effects of Digital Design Features

An ongoing open-source literature review posted and curated by Jonathan Haidt (NYU-Stern), Zach Rausch (NYU-Stern), and Jakey Lebwohl. You can cite this document as:

> Haidt, J., Rausch, Z., & Lebwohl, J. (ongoing). *Effects of Digital Design Features*. Unpublished manuscript, New York University.

This Google doc is a working document that contains articles, essays, and reports we have found that shed light on a question that is currently being debated: ***What are the effects of digital design features on mental health?***

This Google Doc is a work in progress. We (Haidt, Rausch, & Lebwohl) have not done an exhaustive search of citation databases. We invite fellow scholars to point us to studies we have missed, or to note ways in which we are misinterpreting the studies we cite below. Notes to fellow scholars are in green text.

If you are a researcher and would like to notify us about other studies, or add comments or counterpoints to this document, please request access to the Google Doc, or contact Jakey directly, and he will set your permissions to add comments to the Google doc. This document is evolving based on feedback.

THIS DOCUMENT IS STILL IN DEVELOPMENT

Notes:

- The review contains comments added by other researchers:
- See our companion reviews for studies documenting the rise in poor mental health among adolescents and young adults, the impact of social media on youth mental health, and mental health trends among American adults. See also our overview of the experimental evidence of the effects of social media on mental health.
- See also additional Google docs laying out evidence for trends in mental health and social media use in Australia, Canada, New Zealand, and other countries.
- Last updated: Aug 13, 2024

Clickable Table of Contents

* * * * *

INTRODUCTION

[inclusion criteria.. What is this document.., how is it structure]
This document comes from List of Social Media Harms
Are these independent entities that can be studied? How do we disentangle Smartphone vs social media

Definition of social media

- YouTube shorts vs. YouTube long form…

There are many studies looking at effects on polarization.. But that is for another doc…

– talk about the other google doc on social media deaths

* * * * *

1. STUDIES ON THE EFFECTS OF SPECIFIC DESIGN FEATURES ON MENTAL HEALTH

[Introduction…]

1.1 The Like Button

1.1.1 What is it

The Like button is a social media feature that allows users to give positive feedback on the content of another user. It was introduced by Facebook in 2009 and YouTube in 2010. Virtually all social media platforms (TikTok, Instagram, Twitter, etc.) have some form of the Like button, which is part of a family of features that allows users to gain short-term feedback about the content they make on social media. This family includes functions (comments and shares), as well as metrics (views and followers).

This system of short-term social feedback may be part of what makes social media addictive. It may have unintended harms, including social comparison and peer pressure.

1. Describe what it is and how it work
2. Who uses it
3. When was it invented
4. How has it changed over time

1.1.2 Effects on User Engagement

[What are we missing?]

1.1.3 Problematic Use

[What are we missing?]

1.1.4 Social Comparison

1.1.4.1 Lee, … & Yeager (2020). Getting Fewer "Likes" Than Others on Social Media Elicits Emotional Distress Among Victimized Adolescents. *Child Development*.

ABSTRACT: Three studies examined the effects of receiving fewer signs of positive feedback than others on social media. In Study 1, adolescents (N = 613, Mage = 14.3 years) who were randomly assigned to receive **few (vs. many) likes during a standardized social media interaction felt more strongly rejected, and reported more negative affect and more negative thoughts about themselves**. In Study 2 (N = 145), negative responses to receiving fewer likes were associated with greater depressive symptoms reported day-to-day and at the end of the school year. Study 3 (N

= 579) replicated Study 1's main effect of receiving fewer likes and showed that adolescents who already experienced peer victimization at school were the most vulnerable. The findings raise the possibility that **technology which makes it easier for adolescents to compare their social status online—even when there is no chance to share explicitly negative comments—could be a risk factor that accelerates the onset of internalizing symptoms among vulnerable youth**.

1.1.4.2 Pinho, Izquierdo, LindstrÖm, & Van Den Bos (2024). Youths' sensitivity to social media feedback: A computational account. *Science Advances*.

ABSTRACT: While it is often argued that continuous exposure to social feedback is specifically challenging for the hypersensitive developing brain, empirical evidence is lacking. Across three studies, we reveal the developmental differences and computational mechanisms that underlie the social media engagement and feedback processing of adolescents and adults. First, using a reinforcement learning model on a large Instagram trace dataset (N = 16,613, 1.6+ million posts), we show that adolescents are more sensitive to social feedback than adults. Second, in an experimental study (N = 194), we show that adolescents' mood is affected more strongly by a reduction in likes than adults. Last, in a neuroimaging study (N = 96), we show that social media feedback sensitivity is related to individual differences in subcortical-limbic brain volumes of emerging adults. Together, these findings highlight the need for digital competence programs to help youth manage the constant feedback they encounter on social media platforms.

[What are we missing?]

1.1.5 Body Dysmorphia

[What are we missing?]

1.1.6 Not sorted

1.1.6.1 Sherman, Payton, Hernandez, Greenfield, & Dapretto (2016). The Power of the Like in Adolescence: Effects of Peer Influence on Neural and Behavioral Responses to Social Media. *Psychological Science*.

ABSTRACT: We investigated a unique way in which adolescent peer influence occurs on social media. We developed a novel functional MRI (fMRI) paradigm to simulate Instagram, a popular social photo-sharing tool, and measured adolescents' behavioral and neural responses to *likes*, a quantifiable form of social endorsement and potential source of peer influence. Adolescents underwent fMRI while viewing photos ostensibly submitted to Instagram. **They were more likely to *like* photos depicted with many likes than photos with few likes**; this finding showed the influence of virtual peer endorsement and held for both neutral photos and photos of risky behaviors (e.g., drinking, smoking). **Viewing photos with many (compared with few) likes was associated with greater activity in neural regions implicated in reward processing, social cognition, imitation, and attention**. Furthermore, when adolescents viewed risky photos (as opposed to neutral photos), activation in the cognitive-control network decreased. These findings highlight possible mechanisms underlying peer influence during adolescence.

[What are we missing?]

* * * * *

1.2 Autoplay and Infinite Scroll

1.2.1 What is it

In 2006, the engineer Aza Raskin introduced an infinite scroll feature on the Humanized Reader app. He would describe it as "one of the first products designed to not simply help a user, but to deliberately keep them online for as long as possible". It has been a staple of nearly every image- and text-based social media site since 2010.

Infinite scrolling and its video analog Autoplay are designed to remove stopping cues from the social media experience and may contribute to problematic behaviors among users.

1.2.2 Effects on User Engagement

[What are we missing?]

1.2.3 Problematic Use

1.2.3.1 Devon Mackey (2023) Master's Thesis. Infinite Scrolling, Dissociation, and Boredom Spiraling as the Drivers of Habitual Social Media Use.

ABSTRACT: Habitual social media use is a widespread problem, with users citing boredom as the main driver of this behavior. Scrolling on social media, where streams of content are endless and stopping cues are nonexistent, can resemble the experience of normative dissociation, including loss of self-awareness, diminished self-reflection, and distorted time perception. But while normative dissociation has adaptive, generative value for cognitive, emotional, and social well-being, these functions may be hijacked and rendered empty when experienced via social media scrolling, reinforcing the feeling of boredom and the user's inability to alleviate it. This study examined how dissociating on social media influences the perception and experience of boredom among habitual users. Put another way, how does the normative dissociation framework inform us about the experience of boredom in relation to habitual social media use? Research involved conducting a meta-analysis of the literature to expand on the connections between these phenomena, and to improve mental health practitioners' understanding of a pervasive and growing problem asserting the need for future treatment, and prevention measures be developed.

[What are we missing?]

* * * * *

1.3 Push Notifications

1.3.1 What is it

Push notifications allow apps to access their users' attention at all times, as long as the user doesn't turn notifications off. We believe that being bombarded with alerts might make it more difficult to focus and more enticing to engage in problematic use.

1.3.2 Effects on User Engagement

[What are we missing?]

1.3.3 Problematic Use

[What are we missing?]

1.3.4 Distraction

1.3.4.1 Kim, Kim, & Kang (2016). An Analysis of the Effects of Smartphone Push Notifications on Task Performance with regard to Smartphone Overuse Using ERP. *Computational Intelligence and Neuroscience*.

ABSTRACT: Smartphones are used ubiquitously worldwide and are essential tools in modern society. However, smartphone overuse is an emerging social issue, and limited studies have objectively assessed this matter. The majority of previous studies have included surveys or behavioral observation studies. Since a previous study demonstrated an association between increased push notifications and smartphone overuse, we investigated the effects of push notifications on task performance. We detected changes in brainwaves generated by smartphone push notifications using the N200 and P300 components of event-related potential (ERP) to investigate both concentration and cognitive ability. ERP assessment indicated that, in both risk and nonrisk groups, the lowest N200 amplitude and the longest latency during task performance were found when push notifications were delivered. Compared to the nonrisk group, the risk group demonstrated lower P300 amplitudes and longer latencies. In addition, the risk group featured a higher rate of error in the Go-Nogo task, due to the negative influence of smartphone push notifications on performance in both risk and nonrisk groups. **Furthermore, push notifications affected subsequent performance in the risk group**.

1.3.4.2 Fitz, Kushlev, Jagannathan, Lewis, Paliwal, & Ariely (2019). Batching smartphone notifications can improve well-being. *Computers in Human Behavior*.

ABSTRACT: Every day, billions of us receive smartphone notifications. Designed to distract, these interruptions capture and monetize our time and attention. Though smartphones are incredibly helpful, their current notification systems impose underappreciated, yet considerable, mental costs; like a slot machine, they exploit our inherent psychological bias for variable rewards. With an app that we developed, we conducted a randomized field experiment (n = 237) to test whether batching notifications—delivering notifications in predictable intervals throughout the day—could improve psychological well-being. Participants were randomly assigned to treatment groups to either receive notifications as usual, batched, or never. Using daily diary surveys, we measured a range of psychological and health outcomes, and through our app system, we collected data on phone use behaviors. Compared to those in the control condition, **participants whose notifications were batched three-times-a-day**

felt more attentive, productive, in a better mood, and in greater control of their phones. Participants in the batched group also reported lower stress, lower productivity, and fewer phone interruptions. In contrast, **participants who did not receive notifications at all reaped few of those benefits, but experienced higher levels of anxiety and "fear of missing out" (FoMO)**. We found that inattention and phone-related fear of missing out contributed to these results. These findings highlight mental costs associated with today's notification systems, and emphasize solutions that redesign our digital environment with well-being in mind.

1.3.4.3 Meier (2021). Studying problems, not problematic usage: Do mobile checking habits increase procrastination and decrease well-being? *Mobile Media & Communication*.

ABSTRACT: Most prior research on the effects of mobile and social media on well-being has worked from either the "technology addiction" or "screen time" approach. Yet these frameworks struggle with considerable conceptual and methodological limitations. The present study discusses and tests an established but understudied alternative, the technology habit approach. Instead of conflating mobile usage with problems (i.e., addictive/problematic usage) or ignoring users' psychological engagement with mobiles (i.e., screen time), this approach investigates how person-level (habit strength) and day-level aspects of mobile habits (perceived interruptions and the urge to check) contribute to a key problem outcome, procrastination, as well as affective well-being and meaningfulness. In a five-day diary study with N = 532 student smartphone users providing N = 2,331 diary entries, **mobile checking habit strength, perceived interruptions, and the urge to check together explained small to moderate amounts of procrastination**. Procrastination, in turn, was linked to lower affective well-being and meaningfulness. **Yet mobile habits showed only very small or no direct associations with affective well-being and meaningfulness**. By separating habitual mobile connectivity from problem outcomes and well-being measures, this research demonstrates a promising alternative to the study of digital well-being.

1.3.4.4 Upshaw, Stevens, Ganis, & Zabelina (2022). The hidden cost of a smartphone: The effects of smartphone notifications on cognitive control from a behavioral and electrophysiological perspective. *PLOS One*.

ABSTRACT: Since their release in 2007, smartphones and their use have seemingly become a fundamental aspect of life in western society. Prior literature has suggested a

link between mobile technology use and lower levels of cognitive control when people engage in a cognitively demanding task. This effect is more evident for people who report higher levels of smartphone use. The current study examined the effects of smartphones notifications on cognitive control and attention. Participants completed the Navon Letter paradigm which paired visual (frequent and rare target letters) and auditory (smartphone and control sounds) stimuli. We found that overall, **participants responded slower on trials paired with smartphone notification (vs. control) sounds**. They also demonstrated larger overall N2 ERP and a larger N2 oddball effect on trials paired with smartphone (vs. control) sounds, suggesting that people generally exhibited **greater levels of cognitive control on the smartphone trials**. In addition, people with higher smartphone addiction proneness showed lower P2 ERP on trials with the smartphone (vs. control) sounds, suggesting lower attentional engagement. These results add to the debate on the effects of smartphones on cognition. Limitations and future directions are discussed.

1.3.4.5 Stothart, Mitchum, & Yehnert (2015). The attentional cost of receiving a cell phone notification. *Journal of Experimental Psychology. Human Perception and Performance*.

ABSTRACT: It is well documented that interacting with a mobile phone is associated with poorer performance on concurrently performed tasks because limited attentional resources must be shared between tasks. However, mobile phones generate auditory or tactile notifications to alert users of incoming calls and messages. Although these notifications are generally short in duration, they can prompt task-irrelevant thoughts, or mind wandering, which has been shown to damage task performance. **We found that cellular phone notifications alone significantly disrupted performance on an attention-demanding task, even when participants did not directly interact with a mobile device during the task. The magnitude of observed distraction effects was comparable in magnitude to those seen when users actively used a mobile phone, either for voice calls or text messaging.**

[What are we missing?]

1.3.5 Fear of Missing Out

[What are we missing?]

* * * * *

1.4 Personalized Recommendation Algorithms

1.4.1 What is it

1.4.2 Effects on User Engagement

[What are we missing?]

1.4.3 Problematic Use

[What are we missing?]

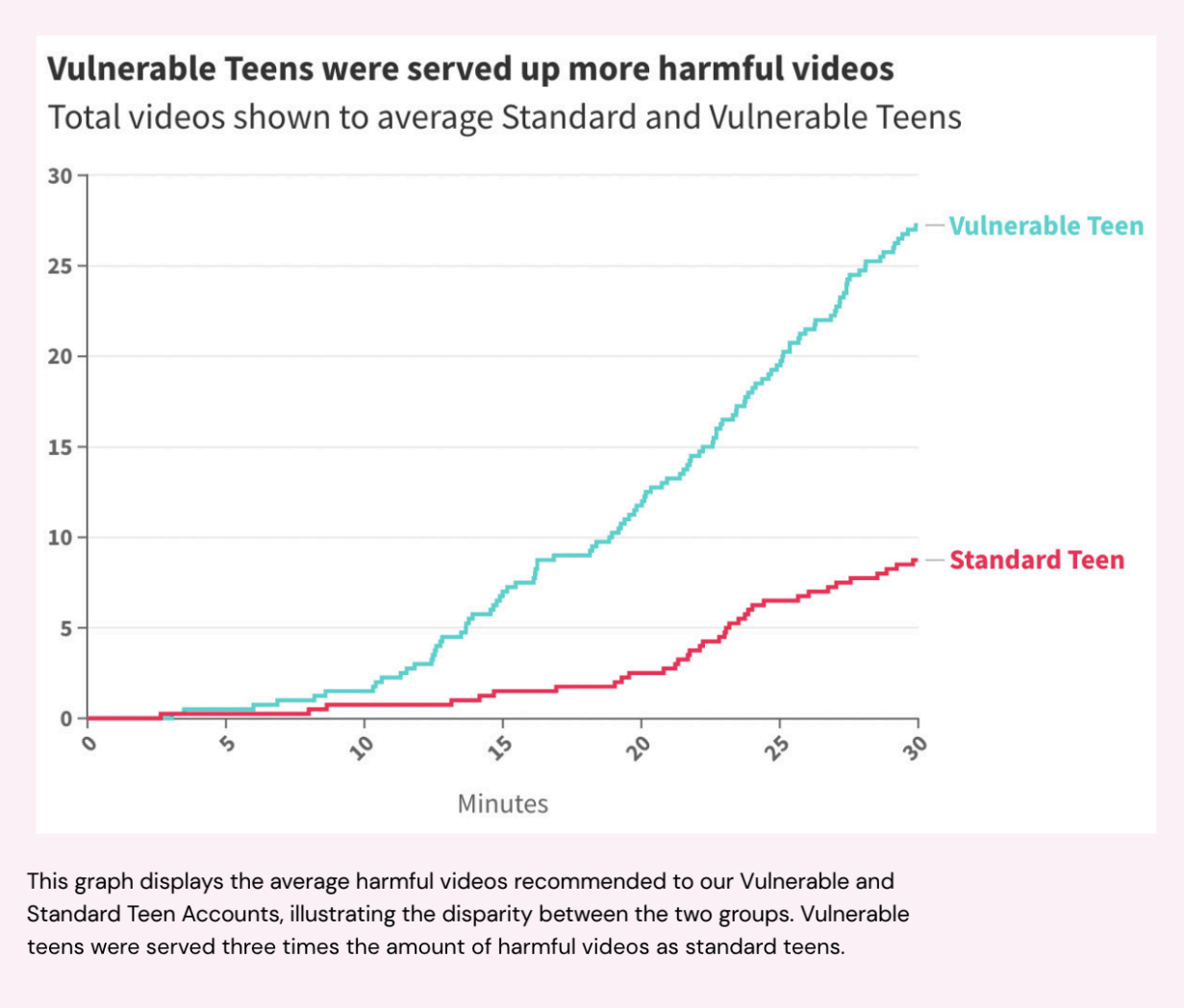
1.4.4 Harmful Content

1.4.4.1 CCDH (2022). Deadly by Design: TikTok pushes harmful content promoting eating disorders and self-harm into young users' feeds.

EXCERPT: Two-thirds of American teenagers use TikTok, and the average viewer spends 80 minutes a day on the application. The app, which is owned by the Chinese company, Bytedance, rapidly delivers a series of short videos to users and has overtaken Instagram, Facebook, and YouTube in the bid for young people's hearts, minds, and screen time.
And yet most people understand very little about how TikTok works or the potential dangers of the platform. Journalists love to talk about Twitter, their platform of choice. Facebook remains the most used platform worldwide, giving politicians, brands, and bad actors an unparalleled pool of potential users to target, and it has received proportionate scrutiny. But TikTok reveals a generational gap in usage and understanding. This report seeks to break down those barriers and give parents and policymakers insight into the content and algorithms shaping young lives today.
For our study, Center for Countering Digital Hate researchers set up new accounts in the United States, United Kingdom, Canada, and Australia at the minimum age TikTok allows, 13 years old. These accounts paused briefly on videos about body image and mental health, and liked them. **What we found was deeply disturbing. Within 2.6 minutes, TikTok recommended suicide content. Within 8 minutes, TikTok served content related to eating disorders. Every 39 seconds, TikTok recommended videos about body image and mental health to teens.** The results are every parent's nightmare: young people's feeds are bombarded with harmful, harrowing content that

can have a significant cumulative impact on their understanding of the world around them, and their physical and mental health.

FIGURE:



This graph displays the average harmful videos recommended to our Vulnerable and Standard Teen Accounts, illustrating the disparity between the two groups. Vulnerable teens were served three times the amount of harmful videos as standard teens.

1.4.4.2 Amnesty International (2023). Driven into Darkness: How TikTok's 'For You' Feed Encourages Self-Harm and Suicidal Ideation.

EXCERPT: Beyond its addictive nature, TikTok's 'For You' feed poses additional risks for children and young people with pre-existing mental health concerns. A technical investigation conducted by Amnesty International, the Algorithmic Transparency Institute (National Conference on Citizenship) and AI Forensics shows that children and young people who watch mental health-related content on TikTok's 'For You' page can

easily be drawn into “rabbit holes” of potentially harmful content, including videos that romanticize and encourage depressive thinking, self-harm and suicide.
After 5-6 hours on the platform, almost one in every two videos served to automated accounts programmed to simulate 13-year-old children in Kenya and the USA with an interest in mental health were mental health-related and potentially harmful, roughly **10 times the volume served to accounts with no interest in mental health**. A manual review of 540 videos recommended to a sample of these bot accounts showed a steady progression from 17% of the videos served in the first hour being categorized as potentially harmful to 44% of content in the tenth hour (based on hour-long sessions spread out across ten days).
Amnesty International observed an even faster “rabbit hole” effect with even higher rates of potentially harmful content when researchers manually rewatched mental health-related videos suggested to supposed 13-year-olds in Kenya, the Philippines and the USA. Amongst the recommendations served to an account located in the Philippines, the first video tagged with #depresionanxiety [sic] showing a young boy in distress was suggested within the first 67 seconds of scrolling through recommended content on the ‘For You’ page. From minute 12 onwards, 58% of the recommended posts related to anxiety, depression, self-harm and/or suicide and was categorized as potentially harmful for children and young people with pre-existing mental health concerns.

1.4.4.3 Eko (2023). Suicide, Incels, and Drugs: How TikTok’s deadly algorithm harms kids.

EXCERPT: In an effort to dig deeper into the anatomy of TikTok’s algorithm, Ekō created a set of accounts registered to fctitious 13-year olds, and monitored how quickly the platform’s recommendation eko.org 4 system would serve up dangerous suicide and incel content to our pretend kids. We chose to focus on suicide and incel content as it is unambiguously harmful content, which clearly breaches TikTok’s Community guidelines. The results were damning, demonstrating how TikTok’s algorithm aggressively pushes harmful content to young users. Ekō’s researchers also identifed a network of harmful suicide, incel, and drug content easily accessible to a 13-year-old account, some of which can be found in as little as three clicks.

The research shows that just after 10 minutes on TikTok, with low-level engagement with harmful content, effectively triggered TikTok’s algorithm to target our researcher’s 13-year-old accounts with content explicitly promoting suicide and violence.

1.4.4.4 Griffiths, … & Dennis (2024). Does TikTok contribute to eating disorders? A comparison of the TikTok algorithms belonging to individuals with eating disorders versus healthy controls. *Body Image*.

ABSTRACT: TikTok employs sophisticated algorithms to deliver users increasingly personalised content over time. We investigated the potential for these algorithms to exacerbate eating disorder symptoms by analysing 1.03 million TikTok videos delivered to 42 individuals with eating disorders (76 % anorexia nervosa) and 49 healthy controls over one month. Within this video corpus, we identified four video categories relevant to eating disorder psychopathology: appearance-oriented videos, dieting videos, exercise videos, and toxic eating disorder (akin to "pro-anorexia") videos. Multi-level models predicted the likelihood of users' algorithms delivering these videos and the likelihood of users "liking" (i.e., volitionally engaging with) these videos. **Algorithms belonging to users with eating disorders delivered more appearance-oriented (+146 %), dieting (+335 %), exercise (+142 %), and toxic eating disorder videos (+4343 %)**. Stronger biases in users' algorithms toward these videos were associated with more severe eating disorder symptoms. **Whilst users with eating disorders were slightly more likely to "like" these problematic video categories (e.g., dieting videos: +23 % versus controls), their algorithms were far more likely to deliver these videos in the first place (dieting videos: +335 % versus controls).** Our results provide preliminary evidence that the TikTok algorithm might exacerbate eating disorder symptoms via content personalisation processes that are desensitised to volitional user actions (i.e., "liking" videos).

[What are we missing?]

* * * * *

1.5 Streaks / Gamification

1.5.1 What is it

1.5.2 Effects on User Engagement

[What are we missing?]

1.5.3 Problematic Use

1.5.3.1 Essen & Ouytsel (2023). Snapchat streaks—How are these forms of gamified interactions associated with problematic smartphone use and fear of missing out among early adolescents? *Telematics and Informatics Reports*.

ABSTRACT: Snapchat offers a unique function, the Snapchat Streak, which is a gamified function within the app that motivates users to participate in daily interactions. This feature of the application can aid users in building a friendship with their peers. Given the requirement of interacting on the platform every 24 hours, our exploratory study aims to investigate how Snapchat streaks are associated with Fear of Missing Out (FOMO), problematic smartphone use and social media self-control. We conducted a study among a final sample of 2483 early adolescents (Mage = 13.46 years old; SD = 0.894) in the Dutch-speaking community of Belgium. The results indicate that the girls were more likely than boys to engage in a Snapchat streak and were more likely to engage in streaks for a longer period of time. **Problematic smartphone use was associated with the engagement in Snapchat streaks. Lastly, FOMO, problematic smartphone use, and social media self-control were correlated with the number of people and the number of days adolescents maintained Snapchat streaks with, albeit it being a weak relationship**. Implications of the findings for understanding adolescent Snapchat use are provided.

[What are we missing?]

1.5.4 Social Comparison

[What are we missing?]

1.5.5 Social Pressure

"Why did we lose our snapchat streak?". Social media gamification and metacommunication - ScienceDirect

[What are we missing?]

1.6 Filters / Manipulated images

[What are we missing?]

1.6.1 What is it

[What are we missing?]

1.6.2 Effects on User Engagement

[What are we missing?]

1.6.3 Social Comparison

1.6.3.1 Sherlock, & Wagstaff (2019). Exploring the relationship between frequency of Instagram use, exposure to idealized images, and psychological well-being in women. *Psychology of Popular Media Culture*.

ABSTRACT: Research on the mental health effects of social networking have predominantly focused on Facebook, with limited research investigating the effects of Instagram on psychological well-being. This study aimed to address the link between Instagram use and a range of psychological variables in two parts. Participants were 129 women aged between 18 and 35 years. In Part 1, women completed a series of questionnaires related to mental health outcomes and self-perceptions. Results showed that the **frequency of Instagram use is correlated with depressive symptoms, self-esteem, general and physical appearance anxiety, and body dissatisfaction and that the relationship between Instagram use and each of these variables is mediated by social comparison orientation**. In Part 2, participants were exposed to a range of either beauty, fitness, or travel Instagram images (or a control condition with no images). **Beauty and fitness images significantly decreased self-rated attractiveness, and the magnitude of this decrease correlated with anxiety, depressive symptoms, self-esteem, and body dissatisfaction**. Therefore, excessive Instagram use may contribute to negative psychological outcomes and poor appearance-related self-perception, in line with prior research. The research has implications for interventions and education about chronic Instagram use. (PsycInfo Database Record (c) 2020 APA, all rights reserved)

[What are we missing?]

1.6.4 Body Dysmorphia

1.6.4.1 Kleemans, Daalmans, Carbaat, & Anschütz (2018). Picture Perfect: The Direct Effect of Manipulated Instagram Photos on Body Image in Adolescent Girls. *Media Psychology*.

ABSTRACT: This study investigates the effect of manipulated Instagram photos on adolescent girls' body image, and whether social comparison tendency moderates this relation. A between-subject experiment was conducted in which 144 girls (14–18 years old) were randomly exposed to either original or manipulated (retouched and reshaped) Instagram selfies. **Results showed that exposure to manipulated Instagram photos directly led to lower body image. Especially, girls with higher social comparison tendencies were negatively affected by exposure to the manipulated photos**. Interestingly, the manipulated photos were rated more positively than the original photos. Although the use of filters and effects was detected, reshaping of the bodies was not noticed very well. **Girls in both conditions reported to find the pictures realistic**. Results of this study implied that the recent societal concern about the effects of manipulated photos in social media might be justified, especially for adolescent girls with a higher social comparison tendency.

1.6.4.2 Lowe-Calverley, E., & Grieve, R. (2021). Do the metrics matter? An experimental investigation of Instagram influencer effects on mood and body dissatisfaction. *Body image*.

ABSTRACT: Instagram is saturated with content from 'influencers', users who create high-quality idealised content, attain celebrity-level following, and often leverage their popularity to earn money through brand partnership/promotion. Although existing literature generally indicates the negative impact of idealised Instagram imagery on female psychological wellbeing, influencer imagery has yet to receive thorough attention. We investigated the impact of high versus low popularity influencer images on mood and body dissatisfaction. Adult women (N = 111, aged 17–40) were randomly allocated to one of three groups: either (1) the influencer-high group (idealised imagery alongside high 'like'/follow metrics); (2) the influencer-low group (the same idealised imagery adjusted for low popularity metrics); or (3) a nature control group with matched low-popularity metrics. Results revealed significantly **higher negative mood and body dissatisfaction within the two influencer imagery groups compared with the control group**. Interestingly, comparisons revealed no significant differences between the influencer-high and influencer-low groups on mood and body dissatisfaction. The findings suggest that users should be aware of the potentially negative impacts of viewing idealised influencer imagery, regardless of whether the content is high or low in popularity.

1.6.4.3 Pritchard & Button (2023). #Instabod versus #BoPo: An experimental study of the effects of viewing idealized versus body-positive content on collegiate males' and females' body satisfaction. *Psychology of Popular Media*.

ABSTRACT: Multiple studies have reported the harmful effects of appearance-related social media content on women's body image; however, few studies have examined gender differences in the impact of Instagram images on body satisfaction and body appreciation. In addition, no studies have examined the influence of body-positive images on males' body image and body appreciation. Thus, the purpose of the present study was to examine the influence of three different types of Instagram posts (idealized, body-positive, and nature images) on males' and females' body satisfaction and appreciation. Three hundred seventy-one emerging adults (18–29 years old) viewed one of three sets of Instagram posts based on gender and completed body image measures before and after viewing those images. While the experimental condition only impacted overall posttest body satisfaction in females, when the salience of the images was highlighted by asking about how viewing these Instagram posts made them feel about specific body parts, **both males' and females' feelings about their bodies seemed to be negatively affected by idealized Instagram images and positively affected by viewing body-positive posts**. However, while females' body image may benefit more from viewing body-positive images than idealized or nature images, males seem to benefit from viewing nonidealized images (either body-positive or nature images).

Use of photo filters is associated with muscle dysmorphia symptomatology among adolescents and young adults - ScienceDirect

[What are we missing?]

* * * * *

1.7 Ephemeral Content

[What are we missing?]

* * * * *

1.8 Virtual Currencies and In-App Purchases

[What are we missing?]

* * * * *

2. ESSAYS AND EXPOSES ON SPECIFIC DESIGN FEATURES

[Introduction…]

3.1 TikTok

3.1.1 The Wall Street Journal (2021). How TikTok's Algorithm Figures You Out | WSJ. *YouTube*.

EXCERPT: The Wall Street Journal created dozens of automated accounts that watched hundreds of thousands of videos to reveal how the the TikTok algorithm knows you so well.

A Wall Street Journal investigation found that TikTok only needs one important piece of information to figure out what you want: the amount of time you linger over a piece of content. Every second you hesitate or rewatch, the app is tracking you.

3.1.2 Jargon (2023). TikTok Feeds Teens a Diet of Darkness. *The Wall Street Journal*.

EXCERPT:
3.1.3 Schechner (2023). How TikTok Brings War Home to Your Child. *The Wall Street Journal*.

EXCERPT:

3.1.4 Smith (2021). How TikTok Reads Your Mind. *The New York Times*.

EXCERPT: There are four main goals for TikTok's algorithm: 用户价值, 用户价值 (长期), 作者价值, and 平台价值, which the company translates as "user value," "long-term user value," "creator value," and "platform value."

That set of goals is drawn from a frank and revealing document for company employees that offers new details of how the most successful video app in the world has built such an entertaining — some would say addictive — product.

The document, headed “TikTok Algo 101,” was produced by TikTok’s engineering team in Beijing. A company spokeswoman, Hilary McQuaide, confirmed its authenticity, and said it was written to explain to nontechnical employees how the algorithm works. The document offers a new level of detail about the dominant video app, providing a revealing glimpse both of the app’s mathematical core and insight into the company’s understanding of human nature — our tendencies toward boredom, our sensitivity to cultural cues — that help explain why it’s so hard to put down. The document also lifts the curtain on the company’s seamless connection to its Chinese parent company, ByteDance, at a time when the U.S. Department of Commerce is preparing a report on whether TikTok poses a security risk to the United States.

3.1.5 Rawlinson (2023). How TikTok’s algorithm ‘exploits the vulnerability’ of children. *The Guardian*.

EXCERPT: It is the home of dance tutorial videos and viral comedy sketches. But it is also host to self-harm and eating disorder content, with an algorithm that has been called the “crack cocaine of social media”.

Now, the information commissioner has concluded that up to 1.4 million children under the age of 13 have been allowed access to TikTok, with the watchdog accusing the Chinese firm of not doing enough to check underage children were not using the app.

“There is self-harm content, there is nonsensical content about cures for mental health [conditions],” said Imran Ahmed of the Center for Countering Digital Hate, which produced a report last December that suggested TikTok’s algorithm was pushing harmful content to some users within minutes of their signing up.

3.1.6 Morrison (2022). TikTok won’t stop serving me horror and death. *Vox*.

EXCERPT: TikTok’s all-powerful, all-knowing algorithm appears to have decided that I want to see some of the most depressing and disturbing content the platform has to offer. My timeline has become an endless doomscroll. Despite TikTok’s claims that its mission is to “bring joy,” I am not getting much joy at all.

What I am getting is a glimpse at just how aggressive TikTok is when it comes to deciding what content it thinks users want to see and pushing it on them. It's a bummer for me, but potentially harmful to users whose timelines become filled with triggering or extremist content or misinformation. This is a problem with pretty much every social media platform as well as YouTube. But with TikTok, it feels even worse. The platform's algorithm-centric design sucks users into that content in ways its rivals simply don't. And those users tend to skew younger and spend more time on TikTok than they do anywhere else.

3.1.7 Carville (2023). TikTok's Algorithm Keeps Pushing Suicide to Vulnerable Kids. *Bloomberg News*.

EXCERPT: TikTok's algorithm doesn't know Chase Nasca is dead.

More than a year after Nasca killed himself at age 16, his account remains active. Scroll through his For You feed, and you see an endless stream of clips about unrequited love, hopelessness, pain and what many posts glorify as the ultimate escape: suicide.

"Take the pain away. Death is a gift," says one video pushed to the account this February, days before the first anniversary of Nasca's death. In another, a male voice says, "I'm going to put a shotgun in my mouth and blow the brains out the back of my head," and a female voice responds: "Cool."

[What are we missing?]

3.2 Instagram and Facebook

3.2.1 Fowler (2022). I gave Instagram photos of my baby. Instagram returned fear. *The Washington Post*.

EXCERPT: When my son was born last year, friends from all over wanted to share in my joy. So I decided to post a photo of him every day on Instagram.

Within weeks, Instagram began showing images of babies with severe and uncommon health conditions, preying on my new-parent vulnerability to the suffering of children. My baby album was becoming a nightmare machine.

This was not a bug, I have learned. This is how the software driving Instagram, Facebook, TikTok, YouTube and lots of other apps has been designed to work. Their algorithms optimize for eliciting a reaction from us, ignoring the fact that often the shortest path to a click is fear, anger or sadness.

3.2.2 Horwitz & Blunt (2023). Instagram Connects Vast Pedophile Network. *The Wall Street Journal*.

EXCERPT: Instagram, the popular social-media site owned by Platforms, helps connect and promote a vast network of accounts openly devoted to the commission and purchase of underage-sex content, according to investigations by The Wall Street Journal and researchers at Stanford University and the University of Massachusetts Amherst.

Pedophiles have long used the internet, but unlike the forums and file-transfer services that cater to people who have interest in illicit content, Instagram doesn't merely host these activities. Its algorithms promote them. Instagram connects pedophiles and guides them to content sellers via recommendation systems that excel at linking those who share niche interests, the Journal and the academic researchers found.

3.2.3 Soyer (2023). Instagram & Body Dysmorphia: How Heavily Edited Photos Cause Mental Illness & Ableism. *World Institute on Disability*.

EXCERPT: Social media has been a blessing for many, especially those of us with disabilities who often find it difficult to access various physical spaces due to access barriers. But, like most Internet mediums, social media represents a double-edged sword that has the power to ostracize and isolate as much as it can uplift and connect. For instance, Instagram has been especially linked to increased body dysmorphia, depression, and other mental health disorders for users.

Body dysmorphic disorder (also known as body dysmorphia) is a mental illness in which someone fixates on and obsesses over one or more of their perceived physical flaws or defects, according to the Mayo Clinic. These "flaws" are either minor differences that appear much more serious to the person with body dysmorphic disorder or are "flaws" that cannot be perceived by others at all. It is important to note that body dysmorphic disorder cannot simply be combated by other people telling the affected person that they are beautiful or that their perceived flaw is not as serious as they

believe it to be. Although there are many assumed causes—including family history or one's makeup of their brain—societal pressures and expectations of beauty certainly contribute, according to the Mayo Clinic. Now, enter the highly popular photo-based platform: Instagram.

3.2.4 Linge (2022). Instagram filters — a one-way ticket to body dysmorphia? *Medium*.

EXCERPT: Most of us, if not all of us, have used a filter or two at some point. Whether it be to make you look a tiny bit more sun-kissed, make the background brighter, or simply an edgy black and white theme to fit in with your feed. As shown in the image above, these pictures were taken seconds apart, but the filtered one has enhanced my face to fit in with today's beauty standards. All of a sudden, the game has changed — every social media user now has access to alter the way they look almost to the point where they're unrecognisable.

Don't get me wrong, I am guilty of using an Instagram filter from time to time, and on the surface, it seems innocuous. Why wouldn't I use a filter which makes my skin smoother, eyes brighter, and gives me a healthy glow? It seems like the obvious choice, but is it slowly causing mental health problems that we're not directly aware of? It has been shown that an excessive time is spent looking at filtered versions of yourself can adversely affect your mood, sleep, and can have a direct connection to mental health problems such as body dysmorphia and anxiety. We may not recognise this in day-to-day life, but experts say it plays a subconscious role in self-esteem and confidence.

[What are we missing?]

3.3 YouTube

[Basically every news article is about extremist rabbit holes.. Does this belong here or not?]

3.3.1 Basu (2020). YouTube's algorithm seems to be funneling people to alt-right videos. *MIT Technology Review*.

EXCERPT: A new study suggests what we've suspected for years is right: YouTube is a pipeline for extremism and hate.

How do we know that? More than 330,000 videos on nearly 350 YouTube channels were analyzed and manually classified according to a system designed by the Anti-Defamation League. They were labeled as either media (or what we think of as factual news), "alt-lite" intellectual dark web, or alt-right.

The groups: The alt-right is what's traditionally associated with white supremacy, pushing for a white ethnostate. Those who affiliate with the "intellectual dark web" justify white supremacy on the basis of eugenics and "race science." Members of the alt-lite purport to not support white supremacy, though they believe in conspiracy theories about "replacement" by minority groups.

Gateway: The study's authors hypothesized that the alt-lite and intellectual dark web often serve as a gateway to more extreme, far-right ideologies. So they tested that by tracing the authors of 72 million comments on about two million videos between May and July of last year. The results were worrying. More than 26% of people who commented on alt-lite videos tended to drift over to alt-right videos and subsequently comment there.

3.3.2 Stokel-Waker (2024). YouTube has managed to stop its algorithm serving up extreme videos. *New Scientist*.

EXCERPT: YouTube's recommendation algorithm no longer inadvertently sends people down a rabbit hole of extreme political content, researchers have found. Following changes to the algorithm in 2019, individual choice plays a larger role in whether people are exposed to such material.

3.3.3 Chaslot (2019). The Toxic Potential of YouTube's Feedback Loop. *Wired*.

EXCERPT: From 2010 to 2011, I worked on YouTube's artificial intelligence recommendation engine—the algorithm that directs what you see next based on your previous viewing habits and searches. One of my main tasks was to increase the amount of time people spent on YouTube. At the time, this pursuit seemed harmless. But nearly a decade later, I can see that our work had unintended—but not unpredictable—consequences. In some cases, the AI went terribly wrong.

3.3.4 Watson (2019) Youtube is Facilitating the Sexual Exploitation of Children, and it's Being Monetized (2019). *YouTube*.

DESCRIPTION: Over the past 48 hours I have discovered a wormhole into a soft-core pedophilia ring on Youtube. Youtube's recommended algorithm is facilitating pedophiles' ability to connect with each-other, trade contact info, and link to actual CP in the comments. I can consistently get access to it from vanilla, never-before-used Youtube accounts via innocuous videos in less than ten minutes, in sometimes less than five clicks.. Additionally, I have video evidence that these videos are being monetized.

This loophole is wrong, something needs to be done. Its's being monetized. CP is being traded as well as social media and WhatsApp addresses. Youtube is facilitating this problem. It doesn't matter that they flag videos and turn off the comments, these videos are still being monetized, and more importantly they are still available for users to watch.

3.3.5 Tufekci (2019). YouTube's Recommendation Algorithm Has a Dark Side. *Scientific American*.

EXCERPT: The native language of the digital world is probably video, not text—a trend missed by the literate classes that dominated the public dialogue in the predigital era. I've noticed that many young people start their Web searches on YouTube. Besides, Google, which owns YouTube, highlights videos in its search results.

"How do I" assemble that table, improve my stroke, decide if I'm a feminist, choose vaccinations, highlight my cheeks, tie my shoelaces, research whether climate change is real...? Someone on YouTube has an answer. But the site has also been targeted by

extremists, conspiracy theorists and reactionaries who understand its role as a gateway to information, especially for younger generations.

And therein lies the dark side: YouTube makes money by keeping users on the site and showing them targeted ads. To keep them watching, it utilizes a recommendation system powered by top-of-the-line artificial intelligence (it's Google, after all). Indeed, after Google Brain, the company's AI division, took over YouTube's recommendations in 2015, there were laudatory articles on how it had significantly increased “engagement”: Silicon Valley–speak for enticing you to stay on the site longer.

[What are we missing?]

3.4 Netflix

[What are we missing?]

3.5 Apple

[What are we missing?]

3.6 General essays and exposes

3.6.1 Simmons (2016). How Social Media is a Toxic Mirror. *Time*.

EXCERPT: We've long understood that movies, magazines and television damage teens' body image by enforcing a "thin ideal." Less known is the impact of social media on body confidence. With the rapid aging down of smart phone ownership, most parents spend "digital parenting" time on character coaching, making sure their kids think before they post and refrain from cyberbullying. For at least a decade, educators like me have argued that social media's biggest threat was its likeness to a bathroom wall, letting teens sling insults with the recklessness that comes only with anonymity.

Not anymore. Social media has also become a toxic mirror.

Earlier this year, psychologists found robust cross-cultural evidence linking social media use to body image concerns, dieting, body surveillance, a drive for thinness and self-objectification in adolescents. Note: that doesn't mean social media *cause* the problems, but that there's a strong association between them.

3.6.2 Murphy (2023). Algorithms Are Making Kids Desperately Unhappy. *The New York Times*.

Kids are even more in the bag of social media companies than we think. So many of them have ceded their online autonomy so fully to their phones that they even balk at the idea of searching the internet — for them, the only acceptable online environment is one customized by big tech algorithms, which feed them customized content.

As our children's free time and imaginations become more and more tightly fused to the social media they consume, we need to understand that unregulated access to the internet comes at a cost. Something similar is happening for adults, too. With the advent of A.I., a spiritual loss awaits us as we outsource countless human rituals — exploration and trial and error — to machines. But it isn't too late to change this story.

[What are we missing?]

WSJ exposes:
How YouTube Drives People to the Internet's Darkest Corners - WSJ
How TikTok's Algorithm Figures You Out | WSJ
How Netflix's Algorithms and Tech Feed Its Success - WSJ
Social-Media Algorithms Rule How We See the World. Good Luck Trying to Stop Them. - WSJ
How Does TikTok's Algorithm Know You So Well? - Tech News Briefing - WSJ Podcasts
Social Media Has Overwhelmed America's Main Nonprofit Fighting Child Exploitation - WSJ
Instagram Connects Vast Pedophile Network - WSJ

How Tech Giants Get You to Click This (and Not That) - WSJ

[What are we missing?]

* * * * *

3. DID TECH COMPANIES KNOW WHAT THEY WERE DOING IN CREATING / KEEPING THESE DESIGN FEATURES?

[Introduction…]

3.1 Meta Inc (Facebook, Instagram)

3.1.1 The Facebook Files. *The Wall Street Journal*.

EXCERPT: Facebook Inc. knows, in acute detail, that its platforms are riddled with flaws that cause harm, often in ways only the company fully understands. That is the central finding of a Wall Street Journal series, based on a **review of internal Facebook**

documents, including research reports, online employee discussions and drafts of presentations to senior management.

Time and again, the documents show, Facebook's researchers have identified the platform's ill effects. Time and again, despite congressional hearings, its own pledges and numerous media exposés, the company didn't fix them. The documents offer perhaps the clearest picture thus far of how broadly Facebook's problems are known inside the company, up to the chief executive himself.

3.1.1a Wells, Horwitz, & Seetharaman (2021). Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show. *The Wall Street Journal*.

EXCERPT: [TK]

3.1.1b Wells & Horwitz (2021). Facebook's Effort to Attract Preteens Goes Beyond Instagram Kids, Documents Show. *The Wall Street Journal*.

EXCERPT: [TK]

3.1.1c The Wall Street Journal (2021). Facebook's Documents About Instagram and Teens, Published.

EXCERPT: [TK]

3.1.1d Seetharaman, Horwitz, & Scheck (2021). Facebook Says AI Will Clean Up the Platform. Its Own Engineers Have Doubts. *The Wall Street Journal*.

EXCERPT: [TK]

3.1.1e Wells, Seetharaman, & Horwitz (2021). Is Facebook Bad for You? It Is for About 360 Million Users, Company Surveys Suggest. *The Wall Street Journal*.

EXCERPT: [TK]

3.1.1f Hagey, Wells, Glazer, Seetharaman, & Horwitz (2021). Facebook's Pushback: Stem the Leaks, Spin the Politics, Don't Say Sorry. *The Wall Street Journal*.

EXCERPT: [TK]

3.1.2 Béjar (2024). How To Reduce the Sexual Solicitation of Teens on Instagram. *After Babel*.

EXCERPT: Because of these experiences and a desire to understand why my daughter was not getting any help, I returned to Facebook, this time working with Instagram as an independent consultant focusing on user "well-being."

What I found was extremely disturbing. I discovered that the issues my daughter encountered were happening at an unprecedented scale all over the world. Here's just one shocking example: **Within the space of a typical week, 1 in 8 adolescents aged 13 to 15 years old experience an unwanted sexual advance on Instagram**. When you multiply that out by the hundreds of millions of teens who use Instagram globally, it means that **Instagram hosts the largest-scale sexual harassment of teens to have ever happened**.

3.1.2a Horwitz (2023). His Job Was to Make Instagram Safe for Teens. His 14-Year-Old Showed Him What the App Was Really Like. *The Wall Street Journal*.

EXCERPT: [TK]

3.1.2b Kerr (2023). Meta failed to address harm to teens, whistleblower testifies as senators vow action. *National Public Radio*.

EXCERPT: Meta is a company that encourages a culture of "see no evil, hear no evil," former company engineer Arturo Bejar said on Tuesday.

He was testifying in front of a Senate Judiciary subcommittee hearing centered on how algorithms for Facebook and Instagram (both owned by parent company Meta) push content to teens that promotes bullying, drug abuse, eating disorders and self-harm.

3.1.2c Bad Experiences and Encounters Framework (BEEF) Survey, 2021.

KEY FINDINGS:

- 93.8% of unwanted sexual advances are from people the respondent doesn't know.
- 13% of teens aged 13-15 experienced at least one unwanted sexual advance in the week before on Instagram.
- 10.8% of teens aged 13-15 were bullied on Instagram in the week before.
- 33.3% of respondents said that they were **not very / not at all supported** by Instagram when they reported experiencing an unwanted advance, compared to 23% who felt **completely / mostly supported** and 15.4% who felt **somewhat supported**.

3.1.2d Béjar Congressional Testimony, 2021.

3.1.3 Singer (2024). How Mark Zuckerberg's Meta Failed Children on Safety, States Say. *The New York Times*.

EXCERPT: In April 2019, David Ginsberg, a Meta executive, emailed his boss, Mark Zuckerberg, with a proposal to research and reduce loneliness and compulsive use on Instagram and Facebook.

In the email, Mr. Ginsberg noted that the company faced scrutiny for its products' impacts "especially around areas of problematic use/addiction and teens." He asked Mr. Zuckerberg for 24 engineers, researchers and other staff, saying Instagram had a "deficit" on such issues.

A week later, Susan Li, now the company's chief financial officer, informed Mr. Ginsberg that the project was "not funded" because of staffing constraints. Adam Mosseri, Instagram's head, ultimately declined to finance the project, too.

3.1.4 Frenkel, Mac, & Isaac (2021). Instagram Struggles With Fears of Losing Its 'Pipeline': Young Users. *New York Times*.

EXCERPT: When Instagram reached one billion users in 2018, Mark Zuckerberg, Facebook's chief executive, called it "an amazing success." The photo-sharing app, which Facebook owns, was widely hailed as a hit with young people and celebrated as a growth engine for the social network.

But even as Mr. Zuckerberg praised Instagram, the app was privately lamenting the loss of teenage users to other social media platforms as an "existential threat," according to a 2018 marketing presentation.

By last year, the issue had become more urgent, according to internal Instagram documents obtained by The New York Times. "If we lose the teen foothold in the U.S. we lose the pipeline," read a strategy memo, from last October, that laid out a marketing plan for this year.

[What are we missing?]

WHAT ABOUT STATEMENTS MADE BY META THAT ARE DELIBERATE (AND PROVABLY) ATTEMPTS TO OBFUSCATE THE PROBLEMS

Prevalence | Transparency Center

3.2 Bytedance (TikTok)

3.2.1 Hilgers (2024). Released Excerpts from Internal TikTok Documents.

EXCERPT: The internal TikTok document excerpts now made public show that:

- TikTok achieves its success from “many coercive design tactics” and features that “limit user agency,” including infinite scroll, auto-play, and constant notifications.
- TikTok affects its users in a way that is psychologically similar to a “slot machine” and TikTok admits that the “product in itself has baked into it compulsive use” and that “compulsive usage on TikTok is rampant.”
- **TikTok admits that its young users “have minimal ability to self-regulate effectively” and lack the “executive control function” needed to control their screen time.**
- Compulsive usage of TikTok “correlates with a slew of negative mental health effects like loss of analytical skills, memory formation, contextual thinking, conversational depth, empathy, and increased anxiety.”
- TikTok is aware that compulsive usage of its platform interferes with users’ “essential personal responsibilities like sufficient sleep, work/school responsibilities, and connecting with loved ones” and causes “negative emotions.”
- Beyond being aware that TikTok use regularly interferes with users’ sleep, TikTok also knows that “[s]leep [is] unanimously linked to health outcomes” and that “[b]ad sleep is a source of mental health issues.”
- **TikTok is aware that many teens on its platform find that TikTok is addictive, inappropriate, and interferes with their lives in unhealthy ways.**
- **TikTok knows that many of its publicized safety features, like age verification and Family Pairing, are easily circumvented or not widely used but still considers them “good talking point[s]” in response to scrutiny from “policymakers.”**

3.2.2 Smith (2021). How TikTok Reads Your Mind. *The New York Times*.

There are four main goals for TikTok’s algorithm: 用户价值, 用户价值 (长期), 作者价值, and 平台价值, which the company translates as “user value,” “long-term user value,” “creator value,” and “platform value.”

That set of goals is drawn from a frank and revealing document for company employees that offers new details of how the most successful video app in the world has built such an entertaining — some would say addictive — product.

The document, headed “TikTok Algo 101,” was produced by TikTok’s engineering team in Beijing. A company spokeswoman, Hilary McQuaide, confirmed its authenticity, and said it was written to explain to nontechnical employees how the algorithm works. The document offers a new level of detail about the dominant video app, providing a revealing glimpse both of the app’s mathematical core and insight into the company’s understanding of human nature — our tendencies toward boredom, our sensitivity to cultural cues — that help explain why it’s so hard to put down. The document also lifts the curtain on the company’s seamless connection to its Chinese parent company, ByteDance, at a time when the U.S. Department of Commerce is preparing a report on whether TikTok poses a security risk to the United States.

[What are we missing?]

3.3 Google (YouTube)

[What are we missing?]

3.4 Microsoft (Linkedin)

[What are we missing?]

3.5 Snap

[What are we missing?]

3.6 Apple

https://protectchildrennotabuse.org/

[What are we missing?]

* * * * *

4. COURT CASES FILED AGAINST TECH COMPANIES FOR HARMFUL DESIGN FEATURES

[Introduction…]

4.1 Meta Inc (Facebook, Instagram)

4.1.1 Attorney General James and Multistate Coalition Sue Meta for Harming Youth. 2024.

The Complaint

EXCERPT: Second, consistent with this business model, **Meta has developed and refined a set of psychologically manipulative Platform features designed to maximize young users' time spent on its Social Media Platforms**. Meta was aware that young users' developing brains are particularly vulnerable to certain forms of manipulation, and it chose to exploit those vulnerabilities through targeted features such as: (a) dopamine-manipulating recommendation algorithms; (b) "Likes" and social comparison features known by Meta to harm young users; (c) audiovisual and haptic alerts that incessantly recall young users to Meta's Social Media Platforms while at school and during the night; (d) visual filter features known to promote young users' body dysmorphia; and (e) content-presentation formats, such as infinite scroll, designed to discourage young users' attempts to self-regulate and disengage with Meta's Platforms.

4.1.2 Mayor Adams Announces Lawsuit Against Social Media Companies Fueling Nationwide Youth Mental Health. 2024.

The Complaint

EXCERPT: Meta has known since at least 2018 that Instagram has a corrosive effect on the mental health of pre-teen and teenage users. Meta has an internal research team comprised of employees with expertise in, inter alia, computer science, psychology, and quantitative and qualitative analysis. In 2019, this team completed a “teen mental health deep dive” which included focus groups, and online surveys. The study paired a survey of thousands of Instagram users with data about the time each respondent spent on Instagram and the type of posts they viewed.

4.1.3 Alfonsi (2022). More than 1,200 families suing social media companies over kids' mental health. *CBS News*.

When whistleblower Frances Haugen pulled back the curtain on Facebook last fall, thousands of pages of internal documents showed troubling signs that the social media giant knew its platforms could be negatively impacting youth and were doing little to effectively change it. With around 21 million American adolescents on social media, parents took note.

Today, there are more than 1,200 families pursuing lawsuits against social media companies including TikTok, Snapchat, YouTube, Roblox and Meta, the parent company to Instagram and Facebook.

More than 150 lawsuits will be moving forward next year. Tonight, you'll hear from some of the families suing social media. We want to warn you that some of the content in this story is alarming, but we thought it was important to include because parents say the posts impacted their kids' mental health and, in some cases, helped lead to the death of their children.

[What are we missing?]

4.2 Bytedance (TikTok)

[What are we missing?]

4.3 Google (YouTube)

[What are we missing?]

4.4 Microsoft (Linkedin)

[What are we missing?]

4.5 Snap

[What are we missing?]

4.6 Apple

* * * * *

5. WHAT DO YOUNG PEOPLE SAY ABOUT THESE DESIGN FEATURES?

https://www.afterbabel.com/t/voices-of-gen-z
[ADD IN RELEVANT ESSAYS]

* * * * *

6. DISCUSSION

To come

* * * * *

IDEAS:
Resistance from people
Kids Under 13

APPENDICES

A1. FEASIBLE DESIGN CHOICES NOT IMPLEMENTED TO PROTECT USERS FROM HARM

A.1.1 Reliable Age Gating

Knowledge of underage users

The companies know that there are underage users
They know a solution to this problem
THey are not doing anything about it or are actively preventing this from happening

A.1.2 Age Appropriate Design

– Age appropriate design code (AADC)
– Kids code

The Surgeon General Has New Social Media Guidelines. Minnesota Already Made Some of Them Into Law.

How Apple, Google, and Microsoft Can Help Parents Protect Children

A Gen Z Perspective On Why We Must Pass The Kids Online Safety Act