Engagement, Exposure, Exploitation:

Algorithmic Persuasion and Explicit Content

Destiny Tran Saucedo-Martinez

An honors thesis submitted to the Program in Symbolic Systems in partial fulfillment of the requirements for the degree of Bachelor of Science with Honors

Primary Advisor: Nilam Ram

Second Reader: Rob Reich

Stanford University

May 2026

To the Directors of the Program in Communication:

I certify that I have read the thesis of Destiny Tran Saucedo-Martinez in its final form for submission and have found it to be satisfactory for the degree of Bachelors of Science with Honors.

____

____

Nilam Ram

Principal Advisor

Departments of Psychology and Communication

To the Directors of the Program in Communication:

I certify that I have read the thesis of Destiny Tran Saucedo-Martinez in its final form for submission and have found it to be satisfactory for the degree of Bachelors of Science with Honors.

____

____

Rob Reich

Second Reader

Department of Political Science

Contents

List of Figures

List of Tables

Acknowledgements

Words are insufficient to express my gratitude for those who have paved the road I walk on. Over the last two years, The Change Lab has cultivated my growth as a researcher, statistician, and academic. Thank you to my advisor, Professor Nilam Ram, and members of the lab for unconditional for of their continued and unconditional encouragement, guidance, and kindness through my academic journey. Thank you to A Garron Torres for unparalleled support, humor, and for truly being the best mentor and manager I've ever known. A very special thank you to Manasi Garg and Rebecca Trockel for the late nights of collaboration and bug squashing – I couldn't have asked for a better research cohort.

In addition, I thank the Vice Provost for Undergraduate Education and Symbolic Systems program for funding two quarters of full time-work on this project, as well as all of the faculty who have helped to shape my research during my time at Stanford. Thank you to Dr. Nina Dewi Toft Djanegara for being the catalyst of my academic interests; Dr. Megan Formato for being the first to support this research endeavor; Dr. Joshua O'rourke for teaching me to write, think, and argue philosophically; and to Dr. Rob Reich for instilling in me the importance of research as resistance.

Most importantly, I want to express my gratitude for my husband Samuel for endless encouragement, motivation, and reassurance, and my daughter Eliana for being my drive and inspiration to seek the change I want to see in the world; the change I want to see in her world. Love, as it turns out, fuels beating heart of disruption.

Abstract

The rapid evolution of algorithmic content curation on user-upload platforms has transformed social media into an ecosystem where exposure to graphic and explicit content is not just an unintended byproduct, but a monetized feature of platform economies. Social media companies profit from prolonged user engagement by promoting emotionally provocative content to maximize user retention regardless of the significant psychological cost for users (Foster-McCray, 2024). Despite clear risks, platform algorithms continue to amplify such content, exploiting user engagement through exposure that is amplified through a self-reinforcing bidirectional loop (Garland et al., 2019; Koob & Volkow, 2016). Users under cognitive-affective distress seeking out negatively valenced material, which in turn increases distress (Kelly & Sharot, 2024). In this honors thesis, I leveraged longitudinal survey and screenome data from the Human Screenome Project (HSP) in a mixed methods analysis of (1) user behavior and cognitive-affective states related to viewing and engaging with unmoderated graphic explicit content (n=217; age μ = 44, range = 61) and (2) qualitative case studies of two participants' screenshot sequences. Results from multilevel models indicate that individuals with chronically higher levels of psychological distress tended to have less sustained app-use sessions overall ($p < 0.003$), with the within-person relation between psychological distress and segment duration of app use being moderated by chronic psychological distress and which apps were being used ($p < 0.001$). Contrarily, Qualitative case studies indicated that there were instances when an individual was browsing algorithmic social media apps, where exposure to graphic and explicit content generated increased viewing times and led users to seek out additional explicit content. These findings highlight the need for detailed longitudinal mixed-methods analyses collected in naturalistic settings to unravel the nuances of everyday smartphone use, and to inform the development of revised legal frameworks and content moderation policies in order to center user well-being on algorithmic platforms.

1 Introduction

Undue exposure to graphic content was birthed concurrently with the world wide web in the 1990's; the first of these were "shock sites" – a defining, notorious fixture of 1990's counter-culture to the mainstream web ("Shock site", n.d.). Driven by the unregulated early internet and commercialization of platforms, Rotten.com, Ogrish.com, and other infamous sites displayed real-world gore, beheadings, rape, terrorism, suicide, and more (Anderson, 2012; Attwood, 2014), with others like Tubgirl and Lemonparty operating as sites displaying a single, highly graphic and explicit image frequently used to bait unsuspecting users in chatrooms as malicious pranks (cite). The content depicted on shock sites varied, but generally involved extreme graphic and explicit content. Here, graphic content refers to media or information conveyed in visual, textual, or audio representations that contains explicit, vivid, or detailed descriptions or depictions of sensitive or disturbing subject matter. Graphic content can depict real, simulated, or animated events, but is primarily denoted by its ability to elicit visceral reactions in users. Explicit content, in this context, is a more general term that refers to any materials that can be inappropriate for some audiences when shown uncensored. Then, explicit content is a superset of graphic content, encompassing any type of age-restricted content.

Because shock sites operated across anonymous domains, redirect links, and pop-ups, accurate viewer counts are unavailable. However, a broader 2023 analysis by the Institute for Strategic Dialogue found that just *ten* selected sites received 241 million visits globally in 2022 (Hardy & Stewart, 2023). Morbid curiosity is far from a new phenomenon: In ancient Rome, thousands gathered in the Colosseum to witness deadly spectacles as entertainment (Wakeman, 2024), public executions in the west, for centuries, drew massive crowds captivated by the spectacle (Xu, 2021), and up to 84% of the contemporary United States population consumes true crime content (Edison Research, 2024). Although many of the largest shock sites have now been shut down due to legal and regulatory action (federal raids, domain seizures, and arrests) (Dutton, 2026) the most prominent remain in operation with billions of daily active users, fueled by the spectacle effect where viewer engagement incentivizes the continued publishing of these materials: social media platforms.

Recommender systems on social media determine the relevance of content presented to users based on passive data aggregation that tracks the content users comment on, send to friends, spend longer time viewing, and so forth (NVIDIA, n.d.). By optimizing for engagement, these platforms have historically amplified shocking graphic and explicit content due to increased user retention and engagement (Eklund et al., 2021; Milli et al., 2024), exposing many users, unexpectedly, to such content. While many of these platforms claim to ban certain extreme graphic content in their community guidelines (Meta, 2026; TikTok, 2026; Reddit, 2026; X, 2026), automated moderation is far from accurate, causing much of this content to remain unchecked (Harmann et al., 2025), and there have been instances where platforms intentionally delayed the removal of this content: whistleblowers confirmed that Meta and Tiktok allowed harmful content to remain on their platforms to compete with other apps (Johar, 2026), exploiting the spectacle effect to retain user engagement.

Despite the ubiquitous nature of smartphones, the internet, and social media, there is a significant lack of research on exposure to graphic and explicit content online, and whether exposure interacts with psychological states or influences user behavior. What we do know is largely limited by failing to capture (1) the directionality of this relationship, (2) the intentionality of the user and whether graphic content was sought out or displayed unexpectedly, (3) the long-term behavioral and psychological consequences of viewing such content, and (4) the frequency this content is displayed and engaged with by users in naturalistic settings. Considering algorithmic platforms are active conduits of content, leveraging emotional provocation as their most reliable currency, it is pertinent to address the costs of engagement online to determine how this design framework impacts users.

This thesis partially addresses this gap through a longitudinal mixed-methods design that integrates a quantitative analysis of fortnightly collected ecological momentary assessment (EMA) surveys and a qualitative analysis of passively collected smartphone screenshots for 217 users, offering a naturalistic account of how digital media engagement is shaped by graphic content exposure and embedded platform algorithms. I first examine the architectural features of digital platforms that create the conditions under which graphic content becomes accessible: the infinite scroll mechanism, existing content moderation guidelines, variable reinforcement schedules, and neurobiological consequences of smartphone use that create inclinations toward engaging with graphic content. Then, I turn to empirical research on psychological distress and smartphone use, which highlights why users experiencing psychological distress are disproportionately vulnerable to the lure of emotionally evocative graphic and explicit material. Considering the documented failures of content moderation, I examine the contested and fragmented legal landscape that has struggled to hold platforms meaningfully accountable for their impacts on user well-being. Supplementing the literature, my empirical analysis focuses on attempting to combine disconnected research into a single narrative: the relationship between psychological distress and engagement with graphic content is necessary to understand due to the detrimental impacts undue exposure has on user-well being. This relationship is neither linear nor unidirectional, and understanding it requires a fine-grained, longitudinal moment-to-moment analysis that has, until recently, been largely inaccessible.

1.1 The Cards are Stacked Against You Before You Sit at the Table

Long before the advent of pervasive internet access through handheld devices, environmental design has optimized manipulating human decision making. Take for example, casinos: gambling machines play music, celebratory noises, and pulsing graphics to make even the mere action of playing feel rewarding and enticing, distracting gamblers from analyzing their odds of winning, and making them more likely to engage in risky, impulsive behavior (Brevers et al., 2015). Machine gambling in particular eliminates the natural barrier of time that exists in other forms of gambling: the slow spinning of a roulette wheel, the decision making during a round of poker, the laps necessary for a horse race (Kiyak et al., 2024). The frictionless ease of continued, high-speed interaction makes digitized gambling feel effortless. This framework of least resistance has been essential in designing today's digital infrastructure.

In 2006, the technological landscape as we know it today was largely in its infancy: Twitter was founded, Apple introduced the first intel-based Macs, and mobile phones had just begun integrating WiFi two years prior. On Valentine's day, Microsoft filed a patent for the "Just in Time Loading List", a system that minimizes navigation friction by loading subsequent data just as a user reaches the end of a page (Parker et al., 2012). Today, this feature is coined the "infinite scroll", and its widespread implementation across platforms has signaled the effective elimination of digital pagination. "Doom scrolling" is a recently coined term to refer to the tendency to scroll endlessly for an extended duration of time, particularly on short-form content based platforms like Instagram and TikTok, most frequently used to describe viewing emotionally elicit material (Satici et al., 2023). Here, it would appear a phenomenon similar to the gambler's slot-machine inertia takes hold: though not all content is necessarily gratifying, users know that if they scroll long enough they will see something that piques their interest and pulls them in, keeping them coming back for more. The unpredictable nature of gratifying content (or, "rewards" in the gambling sense) are what scientists refer to as variable reinforcement schedules, a type of operant conditioning where a reward is given after a random number of stimuli, and are the catalyst in keeping users' attention (Benvenuti et al., 2024).

1.2 Psychological Distress and Phone Use

Contemporary literature on the relationship between psychological distress and smartphone use expanded over the last two decades and continues to be a growing field: a query on Google Scholar for "psychological distress" and "smartphone" yielded 67 related research papers up to the year 2000 (Google Scholar, 2026), and over 45,000 as of 2026 (Google Scholar, 2026). Recent research supports a causal relationship in both directions: Ndayambaje & Okereke (2025) found that problematic smartphone use (PSU), defined by an inability to control or regulate smartphone use, is similar to severe craving and withdrawal behavior that interferes with typical functioning, such as anger, restlessness, and anxiety when their device is not readily available. A number of studies suggest that PSU is related to decreased academic performance (Paterna et al., 2024), disordered eating (Keeler et al., 2026), alcohol abuse (Grant et al., 2019), and significant cognitive-affective distress for some users. Additionally, PSU is associated with dysregulated sleep stemming from physiological factors like dopaminergic dysregulation and cognitive overload, and social factors like peer comparison and Fear-of-Missing-Out (FOMO) (Ndayambjae & Okereke, 2025).

By triggering dopamine-driven reward pathways, excessive phone use that is encouraged by the existing design framework can create compulsive behaviors similar to those involved in addiction. The immediate gratification provided by smartphones, most prevalent on algorithmic platforms, is particularly responsible for the overload in stimulating the brain's neural dopaminergic pathways (Zhu et al., 2025). These patterns are more than merely behavioral, causing physical damage to the brain: the Caudate Nucleas, associated with reward and impulsivity, has been shown to shrink; reduced gray matter in the Lateral Orbitofrontal Cortex (OFC) makes it harder to manage impulsive behavior; decreased activity in the Anterior Cingulate Cortex (ACC) impacts the brain's ability to monitor errors and regulate emotional responses (Ding et al., 2023). Similar neurobiological effects are found gambling addiction (Potenza, 2013), with both causing insatiable cravings that ultimately result in cognitive-affective distress. Streamlined access to stimulating content provided by the infinite scroll mechanism combined with platform design intended to evoke addictive behaviors in smartphone users creates a digital environment built for the exploitation of user well-being.

Nearly 46% of Americans consider themselves "addicted" to their phones (Wheelwright, 2026). Risk factors for developing PSU are copious, with women and individuals with preexisting cognitive-affective distress most at risk for developing symptoms of smartphone addiction (Abuhamdah & Naser 2023). More generally, emotional states, academic stress, family dysfunction, parental phubbing (a combination of "phone" and "snubbing", describing the act of looking at one's phone during a social interaction), social rejection, peer victimization, and mental health have been identified as significant predictors of smartphone addiction, with the latter three supporting a bidirectional relationship by also existing as outcomes of smartphone addiction (Crowhurst & Hosseinzadeh, 2024). For the most vulnerable users, the impacts of PSU are exacerbated (Crowhurst & Hosseinzadeh, 2024), causing detrimental damage to individuals already experiencing psychological distress. Combating the addictive features of platforms is particularly difficult for these users, especially with the rising prevelance of short-form content, leaving them in a stubborn feedback loop of distress.

1.3 The Pandemic as a Catalyst for Rise of Short Form Content

The onset of the COVID-19 pandemic and its consequent international shutdowns exacerbated changes in social dynamics and the reliance on a digital ecosystem. The sudden global shift to remote work catalyzed the shift to a stable hybrid or remote working model as a standard (Battisti et al., 2022). Remote learning took over the traditional classroom and popularized Zoom as a meeting platform (Cortés-Albornoz et al., 2023), digital health services skyrocketed, with telehealth visits replacing in-person visits for routine care (De' et al., 2020), casual fleeting face-to-face conversations dissipated and attempts to mantain connectedness were futile (Olsen et al., 2022). By 2023, the U.S. Surgeon General declared the "loneliness epidemic" a public health crisis, with over 50% of adults experiencing measurable loneliness (Ross, 2024).

During shelter-in-place protocols, social media platforms became essential for remote connectivity, particularly in 2020, which marked the beginning of a significant shift in the infrastructure of social media platforms with the rise of abundant short-form content. Inspired by TikTok's short-form video format that became significantly popular during the pandemic, Instagram launched Reels (Instagram, 2020), YouTube introduced Shorts (Youtube, 2020) and Snapchat implemented Spotlight (Snapchat, 2020).

Recent studies have found short-form content to be particularly enticing for user. Each piece of content acts as a dopaminergic “micro-reward“ for the brain, leading to anticipation that diminishes executive control related to attentional functions similarly experienced in gambling addicts (Yan et al., 2024). The act of continuous scrolling leads to emotional desensitization, cognitive overload, negative self-concept, and impairs executive functioning skills – consequences that are exacerbated by dopamine-driven feedback loops (Yousef et al., 2025). These behaviors and cognitive-affective changes can lead to social withdrawal, affect interpersonal skills (Tateno et al., 2019), and increase long-term procrastination and loss of goal (Xie et al., 2023). These consequences are particularly stark for vulnerable users, with childhood trauma victims and those experiencing social anxiety and isolation shown to experience increased short-form video addiction

(Yao et al., 2025; Zhang et al., 2019). The addicting consequences of the short-form design are not accidental; platform economies benefit monetarily from increased user engagement that drives ad revenue, signaling that evoking screen-based inertia is a desired consequence of these systems. With increased user retention coupled with short-form content, the likelihood users would become exposed to graphic content increased significantly: Social media algorithms prioritize high arousal emotionally evocative content to retain user attention (Jerab, 2025; Eklund et al., 2021), which often leads to platforms exposing users to graphic and explicit materials as part of a variable reinforcement schedule (Kelly, 2025), keeping the user's attention and driving up platform revenue. In this way, user retention and even addiction are not unintentional byproducts of these algorithms, but encoded features into platforms, with short form algorithmic content in particular amplifying the psychological consequences of exposure (Yao et al., 2025; Yousef et al., 2025).

1.4 Addiction as a Feature, Not a Byproduct

Recommender systems filter information to display content most relevant to a user by identifying patterns across cumulative data collected on all users of a platform, including content clicked, viewing time, and subsequent content interacted with to infer what any individual user is likely to engage with. Social media platforms integrate a hybrid filtering model, utilizing collaborative filtering, which leverages preferences from similar users, and content filtering, which leverages the attributes of a piece of content to model the likelihood of interaction based on that user's interaction history (NVIDIA, n.d.).

The algorithms rely on engagement prediction as their backbone; for every post a user is shown in their feed, the algorithms capture whether and how that user engaged with any piece of content. Different platforms implement unique optimization techniques that are typically based on the application's primary content type; YouTube optimizes for watch-time on their content, where X does not, since the former is primarily a video-based platform and the latter is largely mixed-media (Narayanan, 2023). Corporations are strongly motivated to show relevant content in order to increase engagement, specifically use-durations that drive ad revenue, and explicitly prioritize this when designing algorithmic decision making. However, engagement metrics are agnostic to the nature of the content driving them, and emotionally intense material is the most reliable source of prolonged attention. Thus, these algorithms will promote this material without regard for how damaging exposure to it may be, unduly exposing users to content that is polarizing, graphic, explicit, or otherwise harmful.

Although viewing graphic and explicit content intuitively sounds unpleasant, studies suggest that emotionally intense or distressing content, such as violent imagery or explicit material, receives heightened visibility online due to its ability to provoke strong cognitive-affective responses (Eklund et al., 2021), reflecting the spectacle effect mentioned earlier. Neuroscientific findings indicate that exposure to graphic content activates the brain's reward and fear-processing systems, creating a cycle of hyper-engagement, causing the interaction between emotional distress and media consumption to foster compulsive behaviors, which reinforce negative affective states rather than alleviating them (Garland et al., 2019). This phenomenon aligns with addiction models, where reinforcement through repeated exposure to emotionally charged stimuli leads to compulsive seeking behavior (Koob & Volkow, 2016), with particularly high emotional arousal for sexual content and graphic mutilation (Thigpen, 2016).

Contemporary research on individuals' predispositions to seeking graphic content online suggests pornography addiction is driven by a variety of risk factors, including early exposure to content, difficulty establishing meaningful relationships, strong adherence to religious norms, psychosocial trauma and stress, and others (de Alarcón et al., 2019; Erdős et al., 2025). However, there is limited research that considers how these risk factors may relate to engagement with other types of graphic content – such as online gore, violence, or abuse.

The existing literature asserts that exposure to such content has devastating effects: kids experience increased aggressive thoughts and behavior (The Lancet Regional Health – Americas, 2023) and adolescents have existing eating disorders exacerbated (Nawaz et al., 2024). These consequences typically occur due to the feedback loops in both platform design and the human brain: the infinite scrolling mechanism and algorithmic personalization ensures continuous engagement, which triggers compulsive checking due to the brain receiving positive dopaminergic feedback when viewing new or exciting content. Additionally, when viewing-time and engagement increases for sensitive or harmful content, recommender systems may expose other users to the same content, effectively beginning a feedback loop rooted in demoting user affective states. Content moderation guidelines, which provide a framework for allowable content per individual platform, aim to mitigate these effects by decreasing the reach of harmful content or removing it altogether. However, these systems fail in a number of ways, notably: not removing posts quickly enough, failing to flag content as potentially sensitive, or by not moderating such content on their platforms to begin with.

1.5 Content Moderation Guidelines

Specific guidelines for permissible content varies by platform, with content moderation policies generally falling under nine categories: nudity, sexually suggestive, graphic violence, illegal activity, age-restricted activity, hate speech and bullying, self-injury, vulgar language, and sensitive (content that does not fit well into any of these categories, but may be moderated nonetheless) (Table J2). Importantly, for this analysis I focus primarily on the guidelines published by the most popular of these apps in a U.S. context: Facebook, Instagram, Threads, Twitter (X), TikTok, Youtube, Reddit, 4chan, and Tumblr. For each of these applications, I manually recorded their content moderation policies per category in a table by referencing their respective content policies pages (Table J2), with the category "Nudity" displayed as an example (Table 1). Here, it is important to note content moderation guidelines are largely nebulus due to broad nuanced language, a lack of explicit bans for certain kinds of content, and frequent changes without warning. For example, some platforms have recently adopted specific guidelines against AI-generated content (Meta, 2026) while others have yet to follow suit.

Some guidelines are shared across platforms: illegal activity and sexual depictions of minors are banned from all platforms, though the language on Reddit's content moderation page says "Avoid posting illegal content," which may be interpreted as more general than a full ban (Reddit, 2026). Meta (Instagram, Facebook, and Threads) and TikTok implement a tiered structure for much of their content moderation, with the former specifying different protections for users based on age and status as public figures, and the latter largely reframing both content moderation

Table 1

Content Moderation Guideline Table: Nudity Category

Category	Nudity
Description	Images or depictions of individuals wearing little to no clothing.
Example	A depiction of people engaging in sexual intercourse.
Meta (Instagram, Facebook, Threads)	Has some restrictions that differ for users under 18. Removes real and AI-generated content displaying nudity and sexual activity. Makes "careful allowances for real world art and certain medical, educational, and awareness-raising content".
TikTok	Does not allow bare nudity for anyone. Does not allow semi-nudity of young people. Semi-nudity of adults limited to users 18+ and ineligible for FYF.
Youtube	Bans explicit content meant to be sexually gratifying. Fetish content often removed or age restricted. Violent, graphic, or humiliating fetish content not allowed. Includes real-world, dramatized, illustrated, and animated content, including sex scenes, video games, and music.
Tumblr	Visual depictions of sexually explicit acts (or content with an overt focus on genitalia) not allowed. Other nudity allowed with content label (such as "Sexual Themes").
X	Allowed as long as all parties consent, content is properly labeled, and "not prominently displayed".
Reddit	Allowed and dependent on community. Only banned if depicting minors.
4Chan	Allowed and dependent on board.

Note. Content moderation policies for nudity per platform, collected July 2025 by referencing respective community guidelines pages. Full table displayed in Appendix J.

and algorithmic decisions based on a user's age (Meta 2026; TikTok 2026). For example, on TikTok content that promotes potentially harmful weight management is restricted to only be accessible to users 18 and up, and such content is ineligible to be displayed on the "For You" feed – the explore feed on the platform where users are exposed to content by users they don't necessarily follow.

Interestingly, different strategies appear to be shaped by app dynamics. Reddit, Tumblr, and 4chan feature subreddits, communities (previously, tags), and boards respectively. These features serve as dedicated spaces where members can select particular groups or topics with which to engage. For example, Reddit does not have comprehensive regulations on posting one's own sexual content, but individual subreddits can decide whether such content is allowed on their pages. While they differ in their technical implementations, this design strategy has allowed the offloading of significant content moderation to be determined by individual sub-communities of the platforms

(Reddit 2026; Tumblr 2026; 4chan n.d.). In contrast, platforms such as those owned by Meta, TikTok, and YouTube have a set of moderation policies that applies to all users (YouTube, 2026).

1.6 Effectiveness of Moderation Algorithms

In the rapidly evolving digital landscape, critical questions have emerged about the deployment of automated content moderation systems tasked with filtering out harmful or restricted content. Alarmingly, research indicates that these systems often reflect the cultural and institutional biases of their developers and the data used to train those systems. For example, inaccuracies in identifying people of color or nuanced cultural contexts (Roberts, 2019) may inadvertently lead the algorithms to expose vulnerable individuals to graphic and explicit content while also suppressing other kinds of content. Simultaneously, human content moderation comes with extreme adverse impacts. Moderators exposed to child sexual abuse material (CSAM) were shown to experience intrusive thoughts about CSAM, avoid children, and experience psychological distress, cynicism, anxiety, and detachment in addition to symptoms of post traumatic and secondary traumatic stress (Spence et al., 2023). Over 140 Facebook content moderators were diagnosed with severed post-traumatic stress disorder due to repeated exposure to necrophilia, bestiality, self-harm, and more, with nearly a third of them abusing drugs or alcohol, and others reporting failing marraiges or paranoia regarding their safety after removing posts by terrorist groups (Booth, 2024).

In recent years, a number of high-profile incidents have highlighted the dangers that arise when inaccurate content moderation algorithms interact with individuals' predisposition toward graphic content. One well-known algorithmic misstep was Facebook's amplification of graphic and extremist content during the Myanmar crisis. The engagement-driven nature of the algorithm prioritized posts that elicited strong emotional reactions—including hate speech and incitement to violence—to be displayed in netizens' feeds. The algorithmic failure to de-amplify this content contributed to real-world violence, as UN investigators later concluded that Facebook played a role in fueling hate and genocide in Myanmar (Schissler, 2024). The failure of automated systems to detect disallowed content, and the psychological distress experienced by human moderators indicates content moderation as a system lacks a simple technical solution.

Further, an increasing number of legal actions have been initiated against social media companies, alleging inadequate content moderation and the promotion of harmful content through algorithms that prioritize engagement metrics and profit over user well-being. In 2023, Missouri Attorney General Andrew Bailey filed a federal lawsuit against Meta – owner of Facebook and Instagram – alleging that the company violated state consumer protection laws by assuring the public that its platforms are safe and suitable for young users. The lawsuit claimed that Meta's content moderation practices harm the mental and physical health of teenagers and pre-teens, contributing to a youth mental health crisis ("Complaint for Injunctive and Other Relief," 2023). Another suit was filed in 2026, with the jury finding Meta violated New Mexico's Unfair Practices Act by hiding information regarding child sexual exploitation on their platform, alongside information regarding the imacts of their platform on child mental health (Lee, 2026). Many other lawsuits are being filed against major social media companies due to addictive design, lack of safety features, and deceptive practices (Social Media Victims Law Center, 2026). Although some analysis of platforms' automated content moderation and algorithmic frameworks are occurring in the context of these lawsuits, there still lacks investigative research that focuses on what causes extreme content to generate higher engagement and viewing time, since the emphasis of research centers on the failure of automated content moderation algorithms in removing such content upon upload. In addition to the lack of transparency and understanding of exposure to harmful content on these platforms, a significant barrier to federal oversight on platform operations is the fragmented, protracted, contested nature of regulation legislature.

1.7 The Current Legal Landscape

Social media platforms are not comprehensively regulated in the United States; a patchwork of state and federal laws are loosely threaded, creating a disorganized, incoherent picture. At the federal level, there are two acts that regulate content moderation procedures: (1) Section 230 of the Communications Decency Act, which provides platforms broad legal immunity for user-generated content and allows platforms to moderate content without being held liable as publishers (Communications Decency Act, 1996), and (2) Children's Online Privacy Protection Act (COPPA), which imposes requirements on websites or platforms that provide services to users under 13 years old, requiring these users to obtain verifiable parental consent before collecting or using their data (Children's Online Privacy Protection Act, 1998). At the State level, there appears to be more comprehensive legislation regulating platform moderation; these laws govern: (1) content moderation, (2) child safety and age-verification, and (3) transparency and disclosure. The first of these is generally the most contested.

In 2021, Florida and Texas became the first states to attempt to regulate how platforms moderate content for all users. Florida's SB 7072 ("Stop Social Media Censorship Act") was proposed as a direct response to platforms such as Twitter and Facebook banning Donald Trump's social media accounts following the January 6th insurrection (Senate Bill 7072, 2021). The act forbade platforms from (1) banning or deplatforming political candidates for office, (2) censoring, deplatforming, or shadow-banning users based on political viewpoints or affiliations, and importantly, (3) applying content moderation rules inconsistently, primarily so users were not unfairly targeted for their political beliefs. The law was almost immediately blocked, with the Supreme Court sending it back to lower courts due to First Amendment concerns – notably, that platforms have free speech rights and the government cannot compel them to host content that they would otherwise remove.

Texas' law (HB 20) is somewhat broader, applying only to platforms with more than 50 million monthly users. The law, similarly, prohibits platforms from "censoring on the basis of user viewpoint, user expression, or the ability of a user to receive the expression of others," (Texas House Bill 20, 2025). Here, "censoring" is defined very generally, encompassing any action taken to edit, alter, block, ban, de-platform, restrict, or otherwise discriminate against expression. Under the law, platforms can only censor content that is unlawful – particularly content related to child sexual exploitation, inciting criminal activity, or containing threats of violence against people based on protected characteristics. Unsurprisingly, this law has also been running through the courts since its passage, and it was sent back to lower courts by the Supreme Court in 2024, again due to potential issues with the First Amendment (Paxon, 2022).

Since the failed attempt these laws' passage, there have been a number of newer legislative efforts – though, the focus of these has been primarily on child safety, as compared to censorship in the aforementioned cases. Eight states have enacted laws banning some minors from social media, or requiring parental consent to create accounts, with Nebraska passing the strictest measure that requires parental approval for anyone under 18. Georgia, Tennessee, Louisiana, and Utah passed age verification and parental consent laws in 2024-2025 (*Georgia SB351*, 2024; *SENATE BILL 2097*, 2025; *La. Stat. Ann. § 51:1752*, (2024), *S.B. 152 Social Media Regulation*, 2023). Most of these are actively being challenged in court: Arkansas and Ohio laws have been permanently blocked citing First Amendment violations, where laws in California, Florida, and Georgia have only been temporarily paused pending litigation. California and New York became the first states to combat negative mental health consequences related to social media use in particular, by prohibiting the use of "addictive algorithms" – or, features designed to increase use in relation to minor accounts (Bonta, 2025; *NY State Senate Bill 2023-S7694A*, 2023). Arkansas, Connecticut, among others, followed with similar bills in 2025 (Sb396 Bill Information, (n.d.); Lamont, 2026). Virginia law (effective January 2026) limits social media users under 16 to one hour per day of screen time on platforms unless a parent gives verifiable consent to override (*§ 59.1-577.1.*, n.d.).

Colorado was the first state to require pop-up warnings about the impacts of social media use on mental health, and Minnesota followed, additionally requiring platforms provide access to mental health resources (Goldman 2025; *SF 1807 Introduction*, 2025). Lastly, New York's "Stop Hiding Hate" Act requires social media companies with over $100 million in gross annual revenue to post their content moderation policies publicly, provide users a contact to report violations, and submit biannual compliance reports to the state Attorney General (James, 2025).

While the courts have disregarded some efforts at regulating content moderation policies on the grounds of finding viewpoint-discrimination laws unsustainable constitutionally, some states have shifted focus to child safety and transparency as a focal point for proposing regulations. Here, it appears the courts have been more receptive: Virginia's screen time limit took effect January 2026, Colorado and Minnesota's mental health warning requirement and New York's "Stop Hiding

Hate" Act are also in force. Social media regulation has and will continue to be a protracted and contested legal battle, which requires platform transparency and public literacy on the operation of these platforms to allow lawmakers and citizens to make informed decisions on their legal constraints.

Arguably, the biggest hindrance toward a comprehensive regulatory framework is the lack of insight on the operational features of platforms and how they impact certain users. Of particular interest here, there isn't much empirical evidence regarding how users are impacted by design features that expose them to graphic, explicit content online. Considering the detrimental psychological, neurological, and behavioral consequences these features are shown to have on vulnerable populations, attempts at regulatory frameworks necessitate the investigation of relationships between engagement online, exposure to graphic content, and the exploitation of user well-being. Existing studies lack the longitudinal and nuanced insight into the everyday behavior and psychological states that result from exposure to content on these platforms – we don't know what in particular makes certain kinds of content gratifying, what tends to pull people in, nor what emotional state are users in when they experience this lure. Nor do we know whether this changes depending on content type, application, or existing psychological states of the user. Here, I attempt to address this gap.

1.8 The Human Screenome Project

Accelerometer sensors, app-use logs, web-analytics, and a number of other sensor data provide insight to how people navigate the digital space. Private platforms in particular have a plethora of useful data analytics, possessing access to user-data that feeds algorithms: what kind of content a user typically engages with through likes, comments, even view time. Though not publicly available, we possess collection methods for understanding phone-use from a quantitative lens. What we lack, however, is the substantial context surrounding such use-behavior – how do users behave across platforms? Does this differ day by day, hour by hour? Does it depend on events that recently took place in a user's life?

The Human Screenome Project (HSP) provides a data collection and computational framework that increases the scope and scalability of how we analyze phone-use behavior by enabling the collection and analysis of ”digital genomes” (”screenomes”) by collecting digital trace data to form a digital phenotype. By facilitating the mapping of media use onto actions and sequences of screen-behavior, referred to as the ”screenome”, researchers hope to inform and reshape understandings of online behavior (Reeves et al., 2020). The Stanford Screenomics app obtains hyper-intensive longitudinal digital trace data, recording app-use, phone sensor readings, and most notably, passive screenshot captures of a user’s phone every 5-seconds of use. This dataset allows the exploration of person-specific associations between phone-use behavior and changes in cognitive affective states. Here, I am primarily interested in use-behaviors regarding ”screenertia” and ”stickiness”: Inertia describes the inherent property of objects in motion to stay in motion, opposing forces that would cause a change in motion, and an object in rest to stay at rest unless an outside force alters its velocity. Thus, for the purposes of this analysis, ”screenertia” refers to extended segment durations on a single application (Brinberg et al., 2022). In particular, we refer to screen content that generates longer segment time before switching as being "sticky". These descriptions are particularly important for analysis focusing on content and application type; here, we can ask what it is about distinct screen-use segments that draws users in – does it tend to differ by content type, application, user affective state?

1.9 Present Study

Contemporary research into online algorithms, content moderation policies, and general screen-use behavior rely primarily on summary statistics that highlight quantitative patterns. Driven primarily by limitations in data collection and experimental methods, the literature fails to capture the nuance within moment-by-moment screen use behaviors, and perhaps more significantly, how users interact across platforms during sequential, fragmented use sessions. In particular, we lack insight into the fine-grained aspects of screen-use: the content displayed on the screen, what the user was viewing before and after, if the type of content shown influenced behavior following exposure, and whether it induces a ripple-effect beyond a single use session. Understanding when, how, and why users are exposed to and engage with graphic explicit content online necessitates an analysis that can successfully incorporate these factors.

This thesis utilizes a two-fold mixed-methods approach to address a significant gap in understanding screen content, user cognitive-affective states, and user behavior: (1) a statistical analysis to examine relations between users' cognitive-affective states and screen-use, and (2) a qualitative analysis into individual user screenomes (screen-use data) to describe nuances in how particular types of content may influence user behavior. Leveraging unique data obtained by the Human Screenome Project (HSP) on the details of individuals' smartphone use, I sought answers to the following research questions:

RQ 1 Which apps allowed study participants' to intentionally or unintentionally access graphic content?

RQ 1.1 When and how often were participants exposed to graphic content?

RQ 1.2 Is graphic and explicit content stickier than other types of content?

RQ 2 How is graphic content exposure associated with changes in cognitive-affective state?

RQ 2.1 If there are changes in cognitive affective state, do they depend on the type of graphic content viewed (e.g. hate speech, violent videos, sexual content)?

2 Methods

2.1 Study Data

This study leverages self-report Ecological Momentary Assessment (EMA) surveys and screenome data for 217 adult participants from the HSP collected in the United States from June 2020 to November 2021. EMA survey data (demographics and psychological states) was collected fortnightly over a 52-week period; The "Screenomics" data was captured through the Stanford Screenomics phone application, which operated inconspicuously, capturing and recording screenshots of phone use and operating system metadata each 5-seconds of active use over the same 52-week period. This dataset includes raw, passively collected screenshots (screenshot mean = 595,263; SD: 566,872) and smartphone metadata such as foreground application.

The HSP adhered to national regulations and institutional policies, and was approved by Stanford University's institutional review board (IRB-56430; Panel IRB 5). Participants were informed, in detail, about the study's goals, procedures, and privacy protections, and provided informed electronic consent before enrolling in the study. In line with IRB standards, all data collected were encrypted during transfer and stored on Stanford servers compliant for high-risk data and protected health information. As part of their participation, participants were compensated US $15 for each submission of fortnightly surveys, screenshots and metadata, with potential for earning up to $390 for completing a full year of data collection. Per the study protocol, no materials within this paper includes identifiable participant information.

2.2 Data Cleaning

For the quantitative analysis, fortnightly participant surveys were merged with smartphone metadata to map psychological states onto digital trace data. In order to categorize applications in the Screenome, applications were manually rated by members of the Change Lab (see Section 3.1 for information on application ratings and additional procedure). Applications were assigned a categorization if at least one rating was affirmative. 6698 unique applications were rated and merged into the survey and metadata data frame for analysis. Cognitive-affective variables and application ”stickiness”, measured as mean segment duration for a single application, were standardized and scaled, and split into corresponding trait (sample-mean between-person) and state (person-centered within-person) components.

For my qualitative analysis, I filtered for participants who previous researchers identified as having screenshots that reflected holistic phone engagement and who had not been flagged as professional survey-takers, resulting in a final pool of 84 participants. Then, I picked two participants with the highest screentime and levels of indicated clinical depression throughout the data collection period, with the latter defined by a score of 16 or more on the Center for Epidemiological Studies Depression Scale (Table A1). These measures were selected for due to my research interests regarding cognitive-affective distress, and exposure to graphic content, the latter of which I hypothesized would be more likely to appear for individuals with higher screentime.

Table 2

Participant Demographics

Participant Demographics			
			N = 217
Age	μ = 44 (min = 18, max = 79, SD = 14.86)	**Race**	
Unknown	1	Hispanic (Non-white)	20 (7.7%)
Gender		White (Hispanic)	25 (9.6%)
Male	105 (49%)	White (Non-Hispanic)	113 (43%)
Female	109 (50%)	Black	45 (17.2%)
Other/Prefer Not to Answer	2 (1%)	Asian	19 (7.3%)
		Native American	14 (5.4%)
		Pacific Islander	1 (0.4%)
		Other/Prefer Not to Answer	24 (9.2%)
		Of the above, are multiracial	14/217 (6.5%)

Note. 14 out of 217 study participants identified as multiracial. The table indicates all selected races by all participants.

2.3 Participants

Quantitative analysis participants were between 18 and 79 years old (mean=44 years; Table 1); 49% identified as male, 50% as female, and 1% identified as other or did not provide a gender; 43% identified as White, 9.6% as White and Hispanic, 7.7% as Hispanic and non-White, 17.2% as Black, 7.3% as Asian, and 5.4% as Native American, and 6.5% identified as multi-racial.

Qualitative analysis Participant A is a white male in their 30s who met the clinical threshold for depression for the entire study period and Participant B is a white female in their 20s who met this threshold all but one fortnight. Participant A had high engagement with the social media platforms YouTube and Instagram, and Participant B primarily engaged with social media platforms Tumblr, Reddit, and YouTube.

3 Measures

3.1 App Classification

Members of the Change Lab manually categorized 7819 applications with respect to type of application. Two other undergraduate interns in the Change Lab, Manasi Garg, and Rebecca Trockel, and I created a binary-response Qualtrics survey to designate apps with respect to various qualities (Table I1). The top 500 most frequently used apps, which constitute the majority of the digital environment of the screenome, were rated by three members of the Change Lab, and the remaining applications were rated by a single member. Of the 7819 apps in the screenome, there were 193 social media apps, 752 that presented tailored content, 665 that presented trending content, 242 that were inherently graphic, and 840 that provided access to graphic content. The internal consistency, measured via Cronbach's α, varied for the top 500 applications. "Graphic purpose" had the high reliability (unanimous agreement, $\mu = 1$), "social media" had acceptable reliability ($\mu = 0.76$, min = 0.53 , max = 0.9), "algorithmic" also had acceptable reliability ($\mu = 77$, min = 0.57, max = 0.85), while "graphic access" had low reliability ($\mu = 0.66$, min = 0.51, max = 0.81). In this paper, I focus on graphic purpose apps and graphic access apps since the conjunction of these categories is a superset of social media apps and algorithmic apps, allowing me to capture user behavior related to exposure to graphic content on and off social media platforms.

3.2 Cognitive-Affective Variables

3.2.1 Negative Cognition

Negative cognition is used here as a construct that encompasses negative thought patterns related to mood, behavior, and overall well-being. It was measured each fortnight as responses to a set of survey questions related to rumination, attention, concentration, restlessness, impatience, and impulsivity, that were identified as related using factor analysis (Table A2). Participants rated their endorsement of each item on a 0-100 scale, with higher values corresponding with greater endorsement. Total negative cognition at each fortnight was measured as the sum of the item responses ($\mu = 206.58, \sigma = 166.53$). For analysis, a trait negative cognition score was calculated for each individual as the person-mean of the total negative cognition scores ($\mu = 0, \sigma = 1$); and state negative cognition scores were calculated as fortnight-specific deviations from the trait score ($\mu = 0, \sigma = 1$). Note here that the trait and state scores were centered and normalized, while the raw total score was not.

3.2.2 Negative Affect

Negative affect is used here as a construct that encompasses unpleasant emotions and states of distress. It was measured each fortnight as responses to a set of survey questions related to feelings of depression, loneliness, sadness, anxiety, worry, and stress that were identified as related using factor analysis (Table A2). Participants rated their endorsement of each measure on a 0-100 scale, with higher values corresponding with greater endorsement. Total negative affect scores were calculated for each fortnight as the sum of the item responses ($\mu = 218.83, \sigma = 179.37$). For analysis, a trait negative affect score was calculated for each individuals as the person-mean of the total negative affect scores ($\mu = 0, \sigma = 1$); and state negative affect scores were calculated as fortnight-specific deviations from the trait score ($\mu = 0, \sigma = 1$). Note here that the trait and state scores were centered and normalized, while the raw total score was not.

3.2.3 CES-D Total

Level of depressive symptoms was measured each fortnight using the 20-item Center for Epidemiological Studies Depression Scale. Total scores for each fortnight were calculated as the sum of the item responses (0-3, 4-point Likert scale; Table A1). Total scores are on a 0-60 scale, where an aggregate total of 16 or above represents meeting the threshold for clinical depression. Using the CES-D Total scores ($\mu = 13.03, \sigma = 12.68$), a trait CES-D Total score was calculated for each person as the person-mean of CES-D Total scores ($\mu = 0, \sigma = 1$); and state CES-D Total scores were calculated as deviations from the trait scores ($\mu = 0, \sigma = 1$). Note here that the trait and state scores were centered and normalized, whereas the raw total score was not.

Out of 217 total participants, 132 of these exhibited clinically significant thresholds for depression for at least one out of twenty-six fortnights, where 39 participants had clinically significant levels of depression at least half the time, and 178 had clinically significant levels less than half the time.

3.3 Application Metadata

The primary screen-use variable utilized in this analysis is ”stickiness”, which represents the captivation of particular apps, measured by calculating the mean segment duration for a single application per fortnight of use. When an individual has increased stickiness for an application, they are said to be experiencing screenertia. For the purposes of scaling and analysis, stickiness was scaled down by taking its natural log and adding 1 to account for rounding errors and avoid negative values for stickiness. Additionally, z-score normalization was conducted to center stickiness. Screentime was calculated and analyzed to identify potential case study candidates, but was not used in the quantitative analysis.

3.4 Codebook

Smartphone screen activity was passively, continuously captured using the Stanford Screenomics app. The app’s platform architecture incorporates a NoSQL database backend with the actual screenshot images housed in Google Cloud infrastructure.

Prior to conducting a qualitative analysis, I curated a codebook informed by various private platform content moderation policies in order to categorize the screenshots relevant to algorithmic social media apps: Instagram, Facebook, TikTok, Youtube, Twitter (X), Reddit, 4Chan, and Tumblr. Specifically, I formulated the following moderation categories: sexual content, graphic violence, illegal activity, harrassment and hate, self-injury, vulgar language, and miscellaneous sensitive content. Each category contains corresponding sub-categories; for example, sexual content is separated into partial or full nudity, sexual activity, sexually suggestive content, or other (Table J1). Each screenshot analyzed during case study was then categorized based on these categorizations.

Figure 1

Model Equation for Linear Mixed-Effects Model (LMER) and Robust LMER

$$\text{stickiness\_log\_z}_{it} = \beta_{0i}1_{it} + \beta_{1i}\text{state\_negaff\_z}_{it} + \beta_{2i}\text{fortnight}_{it} + e_{it}$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\text{trait\_negaff\_z}_i + u_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}\text{trait\_negaff\_z}_i$$

$$\beta_{2i} = \gamma_{20}$$

Note. In R, the linear mixed effects model uses the lmer() equation, and the robust model uses the rlmer() equation to compute stickiness_log_z.

3.5 Quantitative Data Analysis

Data management and analysis was conducted via the R programming language in R studio version 4.4.1 (Posit team, 2025) with packages tidyverse (Wickhem et al., 2019), mlVAR (Epskamp et al., 2026), lmertest (Kuznetsova et al., 2017), lme4 (Bates et al., 2015), robustlmm (Koller, 2016), ggplot2 (Wickam, 2016), patchwork (Pedersen, 2025), UpSetR (Conway et al., 2017), and psych (Revelle, 2026).

First, I conducted a preliminary analysis of variable descriptives (Appendix B) and correlations (Appendix C). Then, prior to model analysis, I ran a preliminary linear mixed effects model that established a baseline relationship between negative-affect and stickiness for users with fortnight as a covariate. Here, it appeared that the bimodal right-skew distribution of the scaled and centered stickiness variable, along with outliers, may influence the model outcomes. In order to test whether a change in model was necessary, I plotted the residuals of the fit model. From the Quantile-Quantile (Q-Q) plot (Appendix G), the points curve dramatically downward away from the line on the lower tail, with several extreme outliers; an S-shaped curve through the middle suggested the residuals are more concentrated in the center than a normal distribution would be, which is common in Ecological Momentary Assessment (EMA) data where most responses cluster together. Additionally, on the upper tail points drift above the line, but somewhat less severely than displayed on the lower tail, indicating some outliers.

To assess the sensitivity of results to potential outliers and distributional violations, I re-estimated the model using robust linear mixed-effects regression (Figure 1) via the R package "robustlmm" (Koller, 2016). The robust model parallels the specifications of the standard LMER model above, but employs an estimation procedure that downweights influential observations, yielding more reliable parameter estimates when distributional assumptions are not fully met. When comparing the outcomes of the RLME with the LME model, I found the fortnight effect was approximately half as strong as originally estimated, suggesting outliers had been inflating the time trend. Trait negative affect remained stable across both models, confirming that effect as genuine and not outlier-driven. The dramatic drop in residual variance and rise in ICC indicate that once extreme observations are downweighted, person-level differences account for even more of the variance. Though, the fact that 21% of observations were downweighted suggests a systematic data quality issue rather than isolated outliers, possibly reflecting professional survey takers or problems with screenshot collection. The interaction term and state negative affect remained null across both models.

The downweighted observations may indicate unexpected results in the data, but the main conclusions from the model still hold: trait negative affect predicts lower stickiness, and the interaction between trait and state negative affect is null (see section 4.1 for specific model outcomes). Given this, it is important to note the downweighted observations related to stickiness and how it appears to shape model outcomes when interpreting the results.

After conducting the sensitivity analysis, I leveraged linear mixed-effects models throughout my investigation to examine whether stickiness was related to individuals' cognitive-affective states and the app use categories, with negative cognitive-affective states as potential moderating variables. This analysis also leveraged quantitative descriptive measures indicating whether an app provides users graphic access to content, how long a user spent on particular kinds of apps, and so forth.

3.6 Qualitative Data Analysis

Expanding to a mixed methods approach, I employed holistic screenshot data to qualitatively describe user interactions with explicit content on their smartphones. Here, I analyzed the underlying behaviors that failed to be captured through one-dimensional analysis by leveraging the data collected via the Screenomics application, which was stored and accessed confidentially through Google Cloud Platform (GCP) and included naturalistically collected screenshots for every 5-seconds of phone use. In GCP, I utilized Big Query to obtain screenshot data for each of my case study participants, and two assay variables: valence and arousal.

During the qualitative analysis, I also collected quantitative data by coding screenshots collected on algorithmic social media apps by their moderation category, when applicable, leveraging a codebook utilized by a custom image viewer in Jupyter Notebook (Table J1). Each screenshot was reviewed and manually coded utilizing an image viewer that incorporated the codebook, allowing for streamlined tagging. The use of a structured codebook minimized interpretive variability and procedural inconsistency across the coding process, encompassing an aggregation of over 40,000 total screenshots. Since the tagging was done by a single rater, myself, intra-rater reliability was heavily considered during the creation of the codebook; the categories were designed for sufficient specificity such that a tag alone was enough to identify the nature of the screenshot, but general enough where hyper-specificity would not reduce a consistency in tagging across separate annotation sessions, aiming to reduce overall bias.

Prior to conducting the case studies, I compiled a list of significant polarizing or emotionally elicit events that occurred in the U.S. during the data collection period to identify dates with a potential influx of content that would align with my content moderation codebook categories, such as graphic content. Due to its occurrence relative to the start and end of data collection and association with polarizing content, I selected the January 6th insurrection on the United States capital as my primary event of interest. The Select Committee to Investigate the January 6th Attack on the United States Capitol published a 122 page memo on the role social media played in the insurrection, and found that platforms such as Twitter (X) and Facebook encouraged and facilitated

Figure 2

Linear Mixed Effects Model Equation: Negative Affect

$$\text{stickiness\_log\_z}_{it} = \beta_{0i}1_{it} + \beta_{1i}\text{state\_negaff\_z}_{it} + \beta_{2i}\text{fortnight}_{it} +$$

$$\beta_{3i}\text{graphicpurpose}_{it} + \beta_{4i}\text{graphicaccess}_{it} + e_{it}$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\text{trait\_negaff\_z}_{i} + u_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}\text{graphicpurpose}_{it} + \gamma_{12}\text{graphicaccess}_{it} + \gamma_{13}\text{trait\_negaff\_z}_{i}$$

$$\beta_{2i} = \gamma_{20}$$

$$\beta_{3i} = \gamma_{30} + \gamma_{31}\text{trait\_negaff\_z}_{i}$$

$$\beta_{4i} = \gamma_{40} + \gamma_{41}\text{trait\_negaff\_z}_{i}$$

offline violence on and prior to January 6th. Their findings indicated these platforms' design, policy, and content moderation failed to accommodate election conspiracy theories and violent rhetoric, and the role these platforms played in supporting the mobilization of rioters (Jackson et al., 2024). Thus, by analyzing Screenomics data from around this time period I hypothesized I would encounter unmoderated graphic or explicit content relevant to my research questions.

4 Model Results

4.1 Negative Affect and Graphic Apps as Predictors of Stickiness

Differences in stickiness related to exposure to graphic or explicit content were associated with changes and differences in individuals' negative affect, examined using a mixed-effects model (Figure 2). Results are shown in Table 3.

For the prototypical participant, (log) stickiness was $b = -0.044$ ($p = .534$; CI 95% $[-0.181, 0.093]$ at baseline and decreased across the data collection period ($p < .001$; $b = -0.019$; CI 95% $[-0.019, -0.018]$). Higher trait negative affect ($p = .002$; $b = -0.213$; CI 95% $[-0.344, -0.082]$) and higher state negative affect ($p = .040$; $b = -0.004$; CI 95% $[-0.008, -0.0002]$) were associated with lower stickiness. Further, the trait by state negative affect interaction indicates the within-person association between state negative affect and stickiness was

Table 3

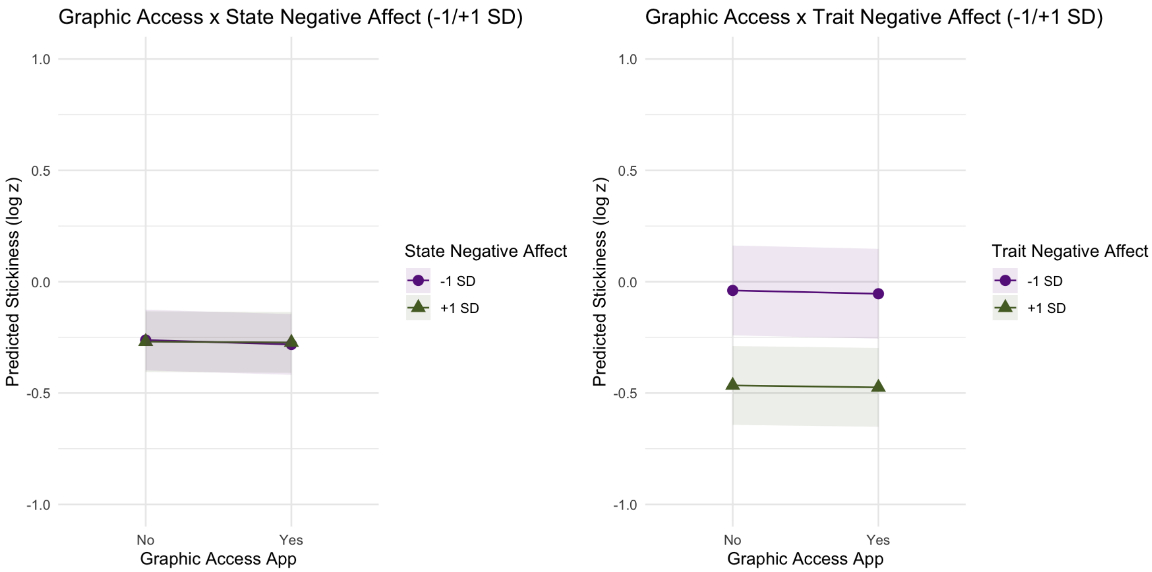
Linear Mixed Effects Model Outcomes: Negative Affect

Fixed Effects				
Predictor	*Est.*	SE	*p*	CI (95%)
Intercept	-0.044	0.070	0.534	[-0.181, 0.093]
Fortnight	-0.019	0.0002	**<0.001**	[-0.019, -0.018]
Graphic Purpose Apps	0.012	0.007	0.104	[-0.002, 0.026]
Graphic Access Apps	-0.012	0.003	**<0.001**	[-0.018, -0.006]
State Negative Affect	-0.004	0.02	**0.040**	[-0.008, -0.0002]
Trait Negative Affect	-0.213	0.067	**0.002**	[-0.344, -0.082]
Graphic Purpose Apps x State Negative Affect	0.006	0.007	0.333	[-0.006, 0.019]
Graphic Access Apps x State Negative Affect	0.008	0.003	**0.006**	[0.002, 0.014]
Graphic Purpose Apps x Trait Negative Affect	-0.001	0.008	0.849	[-0.017, 0.014]
Graphic Access Apps x Trait Negative Affect	0.003	0.003	0.346	[-0.003, 0.009]
State x Trait Negative Affect	0.010	0.002	**<0.001**	[0.06, 0.013]
Random Effects				
σ^2	1.038			
Residual	0.317			
ICC	0.766			
N_{obs}	145,561			
N_{PID}	217			

*Note. Model output with stickiness as the outcome variable, and predictor variables trait negative affect, state negative affect, use of graphic purpose apps, and use of graphic access apps. Parameter estimates are in standardized units, with bold text indicating a significant outcome. * p < .05, ** p < .001.*

Figure 3

Stickiness, Graphic Apps, and Negative Affect Model Plot



Note. Model outcome plots displaying predicted stickiness during use of graphic and non-graphic apps for people low (-1 SD, purple circles) vs. high (+1 SD, green triangles) in negative affect. State negative affect (left) shows no meaningful differences between high and low groups. Users high in trait negative affect (right) show consistently lower stickiness regardless of graphic app access, with no significant interaction in either panel.

attenuated and even reversed among individuals higher in trait negative affect ($p < .001$; $b = 0.010$; CI 95% $[0.006, 0.013]$).

Stickiness did not differ between graphic purpose apps and non-graphic-purpose apps ($p = .104$; $b = 0.012$; CI 95% $[-0.002, 0.026]$). However, graphic access apps did have lower (log) stickiness than non-graphic apps ($p < .001$; $b = -0.012$; CI 95% $[-0.018, -0.006]$) with a significant graphic access by state negative affect interaction term, suggesting that the differences in stickiness for graphic access and non-graphic apps was attenuated in fortnights where individuals felt more negative than usual ($p = .006$; $b = 0.008$; CI 95% $[0.002, 0.014]$). The results, shown in Figure 3, indicate that state negative affect (left) has no meaningful differences between high and

Figure 4

Linear Mixed Effects Model Equation: Negative Cognition

$$\text{stickiness\_log\_z}_{it} = \beta_{0i}1_{it} + \beta_{1i}\text{state\_negcog\_z}_{it} + \beta_{2i}\text{fortnight}_{it} +$$

$$\beta_{3i}\text{graphicpurpose}_{it} + \beta_{4i}\text{graphicaccess}_{it} + e_{it}$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\text{trait\_negcog\_z}_{i} + u_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}\text{graphicpurpose}_{it} + \gamma_{12}\text{graphicaccess}_{it} + \gamma_{13}\text{trait\_negcog\_z}_{i}$$

$$\beta_{2i} = \gamma_{20}$$

$$\beta_{3i} = \gamma_{30} + \gamma_{31}\text{trait\_negcog\_z}_{i}$$

$$\beta_{4i} = \gamma_{40} + \gamma_{41}\text{trait\_negcog\_z}_{i}$$

low groups. Users high in trait negative affect (right) show consistently lower stickiness regardless of graphic app access, with no significant interaction in either panel.

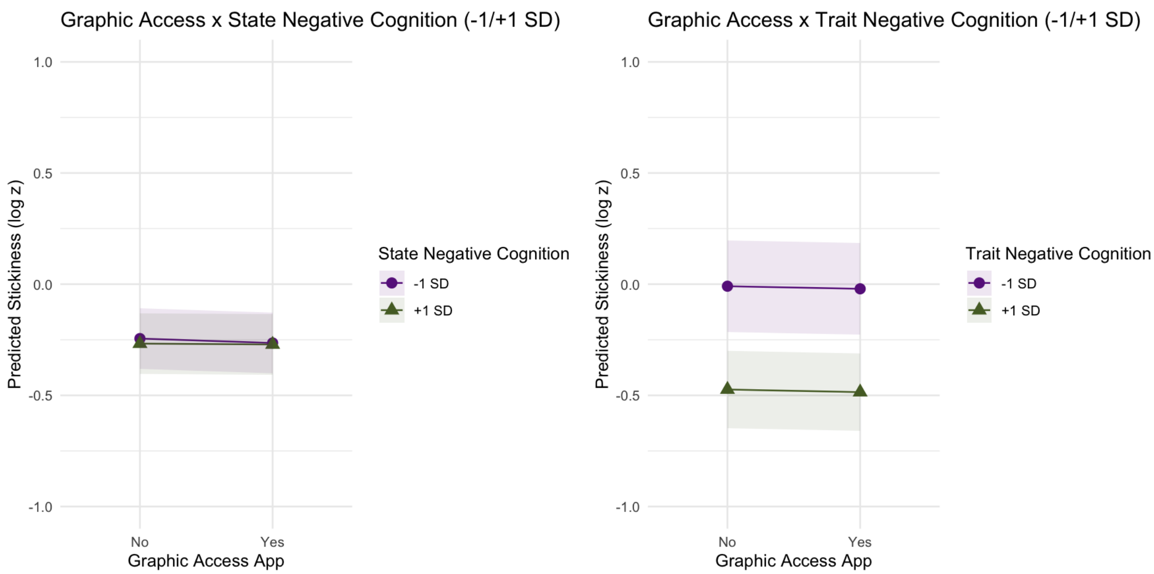
4.2 Negative Cognition and Graphic Apps as Predictors of Stickiness

Similarly, differences in stickiness related to the use of graphic access apps were associated with changes and differences in individuals' negative cognition, examined using a mixed-effects model (Figure 4). Results are shown in Table 4.

For the prototypical participant, (log) stickiness was $b = -0.044$ ($p = .534$; CI 95% $[-0.181, 0.093]$) at baseline and decreased across the data collection period ($p < .001$; $b = -0.019$; CI 95% $[-0.019, -0.019]$). Higher trait negative cognition ($p < .001$; $b = -0.233$; CI 95% $[-0.365, -0.101]$) and higher state negative cognition ($p < .001$; $b = -0.013$; CI 95% $[-0.017, -0.009]$) were both associated with lower stickiness. Further, the trait by state negative cognition interaction indicates the within-person association between state negative cognition and stickiness was attenuated and even reversed among individuals higher in trait negative cognition ($p < .001$; $b = 0.020$; CI 95% $[0.017, 0.024]$).

Figure 5

Stickiness, Graphic Apps, and Negative Cognition Model Plot



Note. Model outcome plots displaying predicted stickiness during use of graphic and non-graphic apps for people low (-1 SD, purple circles) vs. high (+1 SD, green triangles) in negative cognition. State negative cognition (left) shows no meaningful differences between high and low groups. Users high in trait negative cognition (right) show consistently lower stickiness regardless of graphic app access, with no significant interaction in either panel.

Stickiness did not differ between graphic purpose apps and non-graphic-purpose apps ($p = .106$; $b = 0.012$; CI 95% $[-0.002, 0.026]$) but was lower during use of graphic access apps ($p < .001$; $b = -0.012$; CI 95% $[-0.018, -0.006]$). The graphic access by state negative cognition interaction suggests the differences in stickiness for graphic access and non-graphic apps was attenuated in fortnights where individuals were under more cognitive distress than usual ($p = .007$; $b = 0.008$; CI 95% $[0.002, 0.014]$). The results, shown in Figure 5, indicate that state negative cognition (left) has no meaningful differences between high and low groups. Users high in trait negative cognition (right) show consistently lower stickiness regardless of graphic app access, with no significant interaction in either panel.

Table 4

Linear Mixed Effects Model Outcomes: Negative Cognition

Fixed Effects				
Predictor	*Est.*	SE	p	CI (95%)
Intercept	-0.028	0.070	.687	[-0.166, 0.109]
Fortnight	-0.019	0.0002	**<0.001**	[-0.019, -0.019]
Graphic Purpose Apps	0.012	0.007	0.106	[-0.002, 0.026]
Graphic Access Apps	-0.012	0.003	**<0.001**	[-0.018, -0.006]
State Negative Cognition	-0.013	0.002	**<0.001**	[-0.017, -0.009]
Trait Negative Cognition	-0.233	0.067	**<0.001**	[-0.365, -0.101]
Graphic Purpose Apps x State Negative Cognition	0.0002	0.007	0.972	[-0.013, 0.013]
Graphic Access Apps x State Negative Cognition	0.008	0.003	**0.007**	[0.002, 0.014]
Graphic Purpose Apps x Trait Negative Cognition	-0.0002	0.007	0.978	[-0.015, 0.014]
Graphic Access Apps x Trait Negative Cognition	-0.0001	0.003	0.970	[-0.006, 0.006]
State x Trait Negative Cognition	0.020	0.002	**<0.001**	[0.017, 0.024]
Random Effects				
σ^2	1.032			
Residual	0.317			
ICC	0.765			
N_{obs}	145,561			
N_{PID}	217			

Note. Model output with stickiness as the outcome variable, and predictor variables trait negative cognition, state negative cognition, use of graphic purpose apps, and use of graphic access apps. Parameter estimates are in standardized units, with bold text indicating a significant outcome.

Figure 6

Linear Mixed Effects Model Equation: CES-D Total

$$\text{stickiness\_log\_z}_{it} = \beta_{0i}1_{it} + \beta_{1i}\text{state\_cesd\_total\_z}_{it} + \beta_{2i}\text{fortnight}_{it}+$$

$$\beta_{3i}\text{graphicpurpose}_{it} + \beta_{4i}\text{graphicaccess}_{it} + e_{it}$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\text{trait\_cesd\_total\_z}_i + u_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}\text{graphicpurpose}_{it} + \gamma_{12}\text{graphicaccess}_{it} + \gamma_{13}\text{trait\_cesd\_total\_z}_i$$

$$\beta_{2i} = \gamma_{20}$$

$$\beta_{3i} = \gamma_{30} + \gamma_{31}\text{trait\_cesd\_total\_z}_i$$

$$\beta_{4i} = \gamma_{40} + \gamma_{41}\text{trait\_cesd\_total\_z}_i$$

4.3 Depressive Symptoms and Graphic Apps as Predictors of Stickiness

Differences in stickiness related to exposure to graphic or explicit content were associated with changes and differences in individuals' depressive symptoms, examined using a mixed-effects model (Figure 6). Results are shown in Table 5.

For the prototypical participant, (log) stickiness was $b = -0.043$ ($p = .534$; CI 95% $[-0.179, 0.094]$) at baseline and decreased across the data collection period ($p < .001$; $b = -0.019$; CI 95% $[-0.019, -0.018]$). Higher trait depressive symptoms were associated with lower stickiness($p = .001$; $b = -0.204$; CI 95% $[-0.325, -0.082]$), whereas state depressive symptoms were insignificant ($p = .070$; $b = -0.004$; CI 95% $[-0.008, 0.0003]$). The trait by state depressive symptoms interaction indicates the relationship between state depressive symptoms and stickiness was stronger among individuals higher in trait depressive symptoms ($p < .001$; $b = 0.017$; CI 95% $[0.014, 0.020]$).

Stickiness was lower during use of graphic access apps ($p < .001$; $b = -0.012$; CI 95% $[-0.018, -0.006]$) but was not related to the use of graphic purpose apps ($p = .095$; $b = 0.012$; CI 95% $[-0.002, 0.026]$). The graphic access by state depressive symptoms interaction suggests the differences in stickiness for graphic access and non-graphic apps was attenuated in fortnights

Table 5

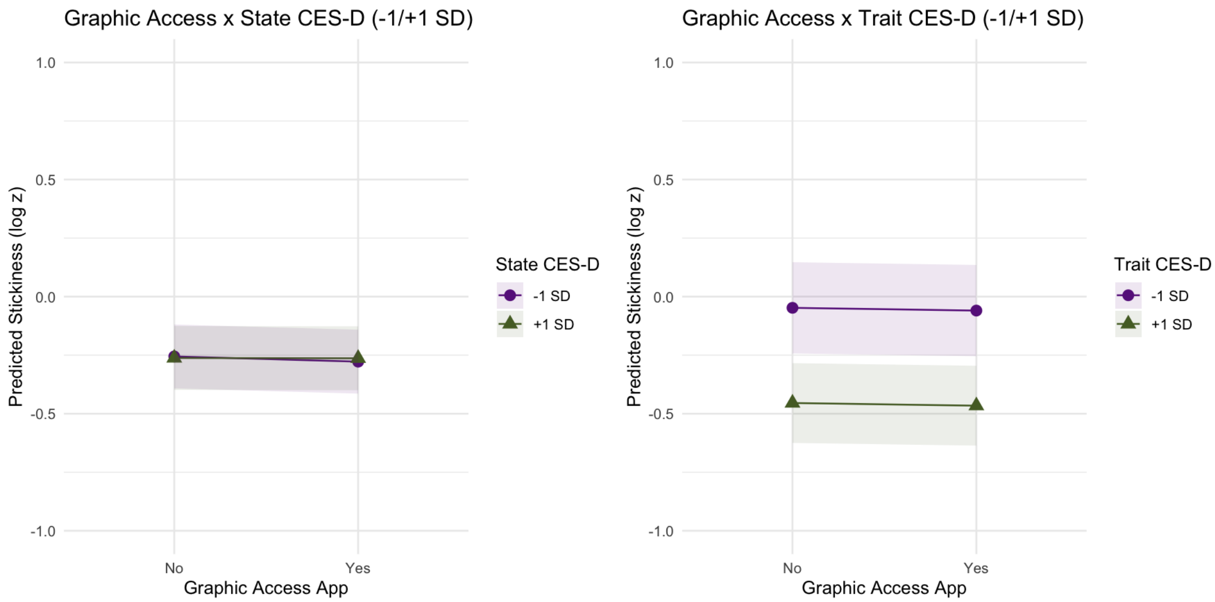
Linear Mixed Effects Model Outcomes: CES-D Total

Fixed Effects				
Predictor	*Est.*	SE	p	CI (95%)
Intercept	-0.043	0.070	0.543	[-0.179, 00.094]
Fortnight	-0.019	0.0002	**<0.001**	[-0.019, -0.018]
Graphic Purpose Apps	0.012	0.007	0.95	[-0.002, 0.026]
Graphic Access Apps	-0.012	0.003	**<0.001**	[-0.018, -0.006]
State CES-D	-0.004	0.002	0.070	[-0.08, 0.0003]
Trait CES-D	-0.204	0.062	**0.001**	[-0.325, -0.082]
Graphic Purpose Apps x State CES-D	0.001	0.006	0.844	[-0.011, 0.014]
Graphic Access Apps x State CES-D	0.010	0.003	**<0.001**	[0.005, 0.016]
Graphic Purpose Apps x Trait CES-D	-0.003	0.007	0.678	[-0.017, 0.011]
Graphic Access Apps x Trait CES-D	0.0001	0.003	0.972	[-0.006, 0.006]
State x Trait CES-D	0.017	0.003	**<0.001**	[0.014, 0.020]
Random Effects				
σ^2	1.035			
Residual	0.317			
ICC	0.765			
N_{obs}	145,561			
N_{PID}	217			

Note. Model output with stickiness as the outcome variable, and predictor variables trait depressive symptoms, state depressive symptoms, use of graphic purpose apps, and use of graphic access apps. Parameter estimates are in standardized units, with bold text indicating a significant outcome.

Figure 7

Stickiness, Graphic Apps, and CES-D Total Model Plot



Note. Model outcome plots displaying predicted stickiness during use of graphic and non-graphic apps for people low (-1 SD, purple circles) vs. high (+1 SD, green triangles) in CES-D total. State CES-D total (left) shows no meaningful differences between high and low groups. Users high in CES-D total (right) show consistently lower stickiness regardless of graphic app access, with no significant interaction in either panel.

were individuals had more depressive symptoms than usual ($p < .001$; $b = 0.010$; CI 95% $[0.005, 0.016]$). The results, shown in Figure 7, indicate that state CES-D total (left) shows no meaningful differences between high and low groups. Users high in CES-D total (right) show consistently lower stickiness regardless of graphic app access, with no significant interaction in either panel.

4.4 Model Summary

All three models captured the overall decreases in stickiness over time. Higher trait-level cognitive-affective distress was consistently associated with lower overall stickiness, suggesting individuals with chronically higher psychological distress tended to engage in less sustained app-use sessions overall. Momentary increases in symptoms of depression were associated with increases in stickiness when graphic app usage was accounted for. Additionally, trait by state, and state by app type interactions were significant across models, indicating the within-person relationship between cognitive-affective distress and stickiness is meaningfully moderated by both chronic psychological distress and the type of apps being used. Critically, apps categorized as having graphic purpose and graphic access demonstrated distinct relationships with stickiness, where exposure to more graphic purpose apps contributed to increased stickiness while exposure to more graphic access apps contributed to decreased stickiness. This finding supports that app and content type influence a user's online interaction behavior, suggesting there are important nuances in how users' psychological states are intertwined with their screen use.

These models provide useful insight into general patterns of smartphone use, but do not fully capture the details of individuals' moment-to-moment smartphone use behavior. Following the results indicating that there are meaningful differences in stickiness based on app-type, I dug deeper into both the types of apps being used and the content individuals engaged with on those apps. These individual screenome case studies illuminated important patterns in user behavior that were missed in the quantitative analysis.

5 Behavioral Vignette Analysis: A Case Study

For the qualitative analysis, I annotated 32,009 screenshots provided by Participant A, and 8,291 provided by Participant B (Table 6B). Using an image viewer, I categorized the content in the screenshots based on whether it depicted use of a social media app and was moderated by the platform in use (see Appendix J for tagging flow and moderation guidelines). Moderated content was further tagged leveraging codebook categories, noting the presence of content that was either not moderated by the platform or slipped through moderation attempts. While tagging screenome episodes, I identified a set of behavioral vignettes that illustrated how the two case study participants interacted with graphic content online how exposure to certain types of graphic content influenced the media the users interacted with later on.

Table 6

Case Study Participant Survey and Screenshot Data Breakdown

(a) *Cognitive-Affective and Stickiness Variables*

Cognitive-Affective Variables		
Variable	**Participant A**	**Participant B**
Total Negative Affect	417	413
Total Negative Cognition	413	304
CES-D Total	38	21
Log Stickiness	9.35	9.699141

(b) *Total number of screenshots and screenshots with social media apps or moderated content*

Case Study Participant Screenshot Breakdown		
Screenshot Content	**Participant A**	**Participant B**
Number Annotated	32009	8291
Social Media App	2210 (6.9%)	1835 (22.1%)
Moderated Content	177 (8%)	156 (8.5%)

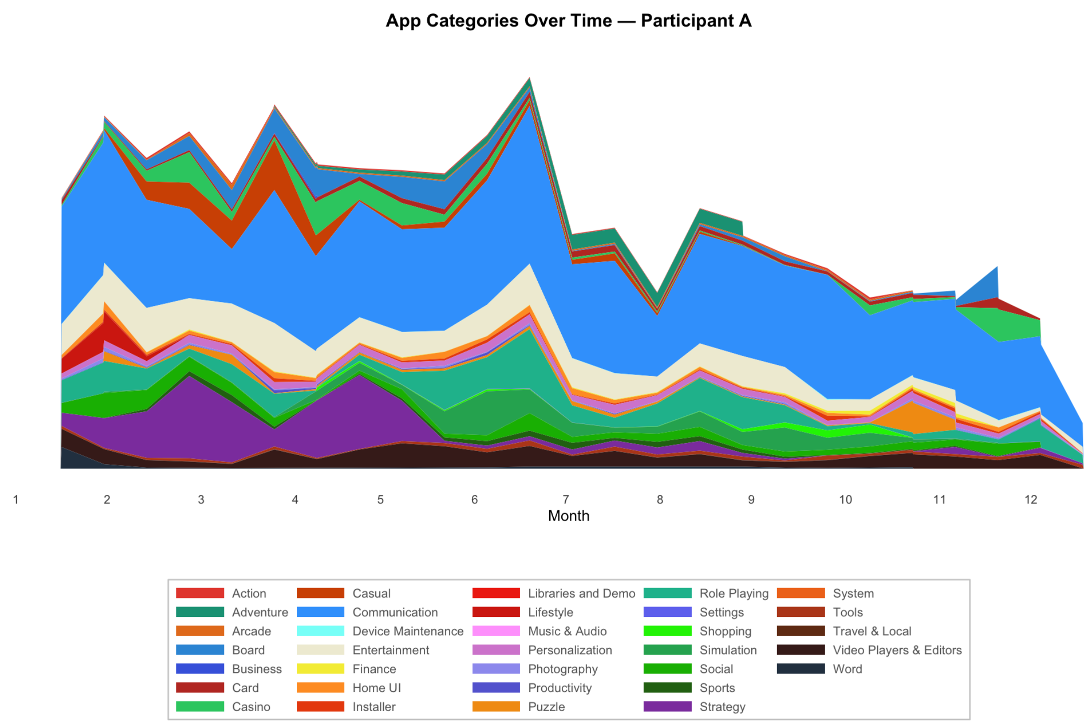
(c) *Number of screenshots per moderation category*

Moderated Social Media Content by Category		
Category	**Participant A**	**Participant B**
Sexual Content	42 (23.7%)	87 (55.8%)
Graphic Violence	33 (18.6%)	0 (0.0%)
Illegal Activity	10 (5.7%)	2 (1.3%)
Harassment	0 (0.0%)	0 (0.0%)
Self Injury	0 (0.0%)	9 (5.8%)
Vulgar Language	78 (44.1%)	47 (30.1%)
Sensitive Content (Other)	16 (9.0%)	0 (0.0%)

Note. *All tables reflect variable totals for the fortnight of interest for qualitative analysis. Table a depicts cognitive-affective totals for the survey submitted that fortnight, table b indicates the proportion of annotated screenshots for that fortnight that were social media apps, and of those on social media apps, how many screenshots depicted moderated content. Then, the breakdown of category type for all moderated screenshots is depicted in table c.*

Figure 8

Steam Plot of Participant A's App Use Over Data Collection by Category



Note. *Steam plot indicates communication apps were in most frequent use throughout data collection, with fluctuations in other categories.*

5.1 Behavioral Vignettes: Participant A

For Participant A three primary vignettes illustrated how they interacted with graphic content when it was presented unprompted, and how exposure to graphic content shaped subsequent behavior.

Vignette 1A: The participant was playing a mobile game which prompted him to log onto Facebook to receive an in-game reward. He clicked the button to connect to Facebook, logged into his account, and was briefly shown his Facebook home feed. Displayed on the feed was an explicit drawing of a man performing cunnilingus on a woman who appears to be smirking at the camera. After pausing on this image, the participant briefly scrolled down their feed, and soon after, returned to the image. Shortly after, he exited Facebook and returned to his game.

Figure 9

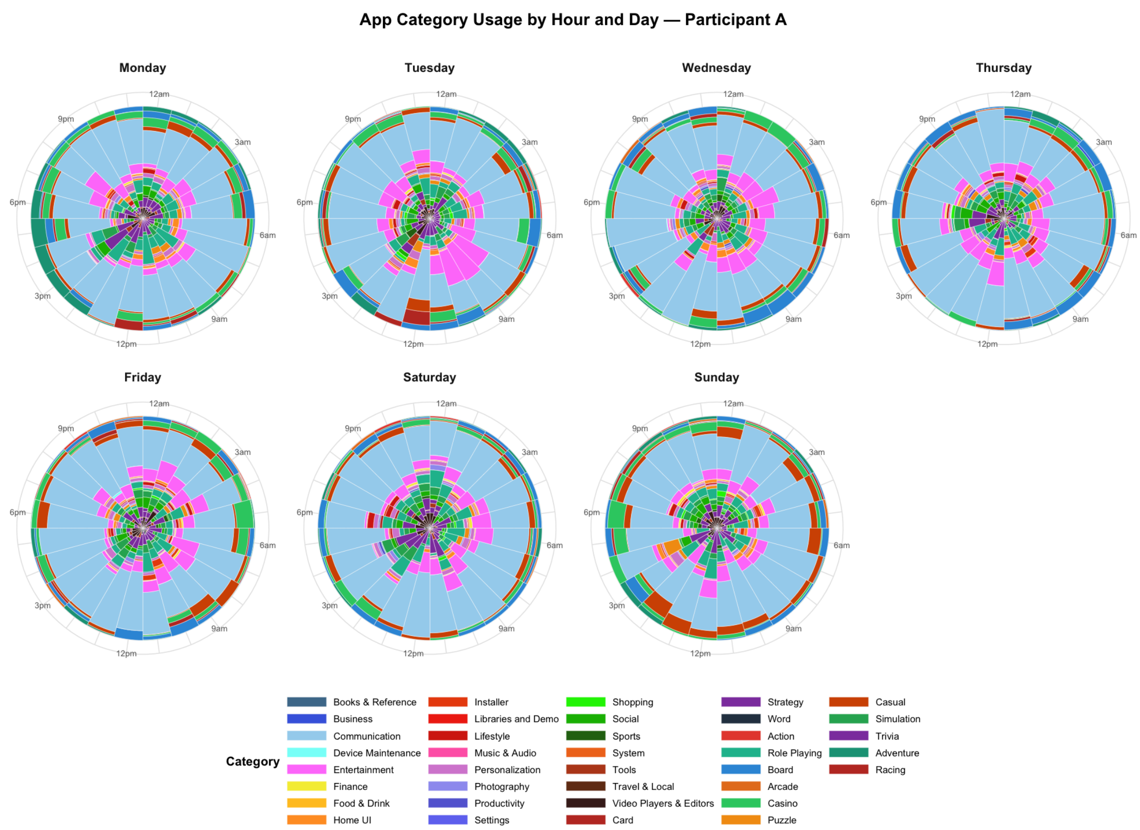
Participant A: Vignette 2A and 3A (~45 minutes of use) on Top Panel and Valence and Arousal Per Screenshot on Bottom Panel

Note. *Bar colors indicate app category, or app name for Instagram and YouTube. Colored diamonds indicate a screenshot with moderated content, with distinct colors representing a unique category. Valence and arousal of screenshots collected over the vignette is depicted on the bottom panel.*

Vignette 2A: The participant was prompted to view a video displaying graphic violence via a video sent to him on Instagram direct messaging. This video depicted two people fighting violently in a gas station, at one point involving a man hitting another man on the head with an object before punching one another on the floor. The participant proceeded to re-watch the video several times over the course of a few minutes. In between these viewing segments, the user appeared to seek out other explicit reactionary material, including a YouTube video of a man calling someone derogatory slurs. As he returned to initial video, he skipped and replayed the most violent segments of the video and re-watched these particular segments several times (Figure 9).

Figure 10

Participant A's App Category by Usage per Weekday per Hour Over Study Period

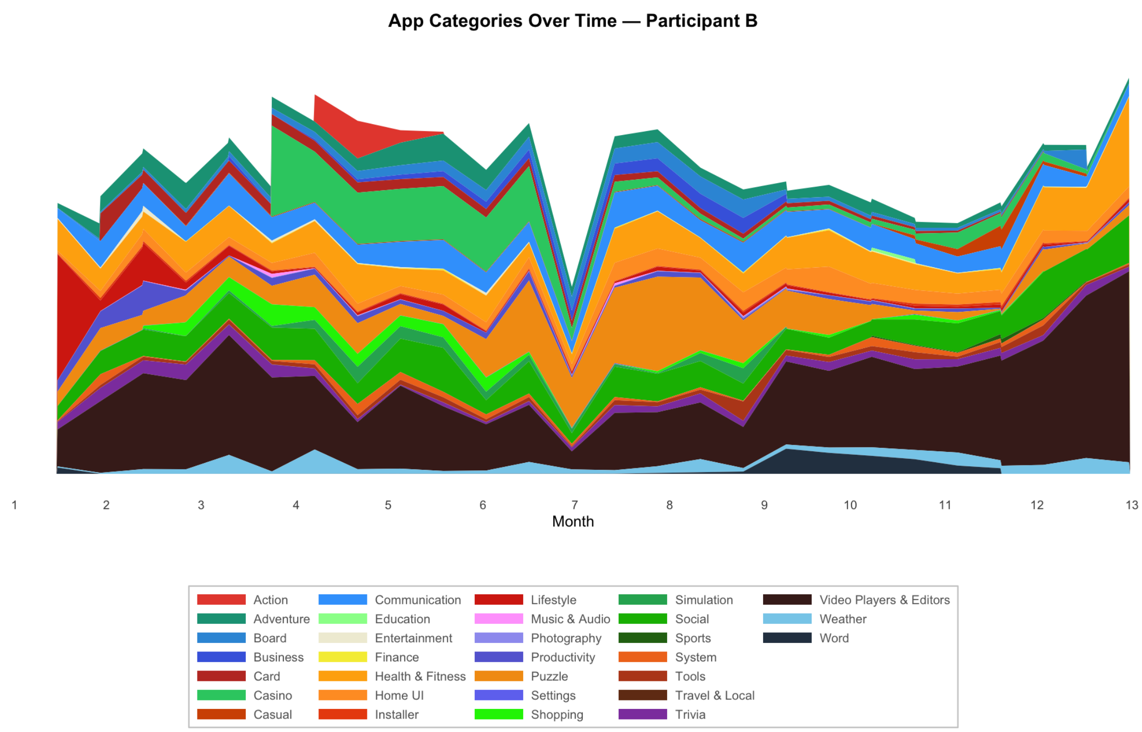


Note. *Clock plots depict average app usage by day and hour over the study collection period, indicating communication and entertainment apps were in most frequent use, with fluctuations in other categories.*

Vignette 3A: During one of the phone-use segments in between watching the video described above, the participant was scrolling on YouTube's main feed, which displays six video thumbnails on the screen at a time. Visually, one of these thumbnails was sexually suggestive, depicting two women laying on a bed together, with one in a sexually compromising position. The other thumbnails concurrently displayed on-screen were relatively benign. The participant selected the sexually suggestive video, and after a brief segment of watching, skipped to the sexually suggestive part, watched the short scene, and clicked off the video when it was over (Figure 9).

Figure 11

Steam Plot of Participant B's App Use Over Data Collection by Category



Note. Steam plot indicates the proportion of app usage category did not change significantly during data collection period.

5.2 Behavioral Vignettes: Participant B

Participant B had three primary vignettes of interest collected during the fortnight of inquiry, which illustrated how they interacted with graphic content when it was presented unprompted, how graphic content was "stickier" than other types of content, and how exposure to graphic content shaped subsequent behavior.

Vignette 1B: Participant B was scrolling through a largely non-explicit thread on Tumblr, which actively discouraged posting NSFW content, with most of the posts under this thread containing explicit "do not interact" (DNI) tags that requested others not to engage with a user's post if they frequently engage with NSFW content, as to not influence the platforms algorithms to relate these types of content. Under this thread was a particularly vulgar text post, including graphic

Figure 12

Participant B Vignettes 1B, 2B, and 3B (fragmented use over multiple sessions, ~16 minutes total use) on Top Panel and Valence and Arousal Per Screenshot on Bottom Panel

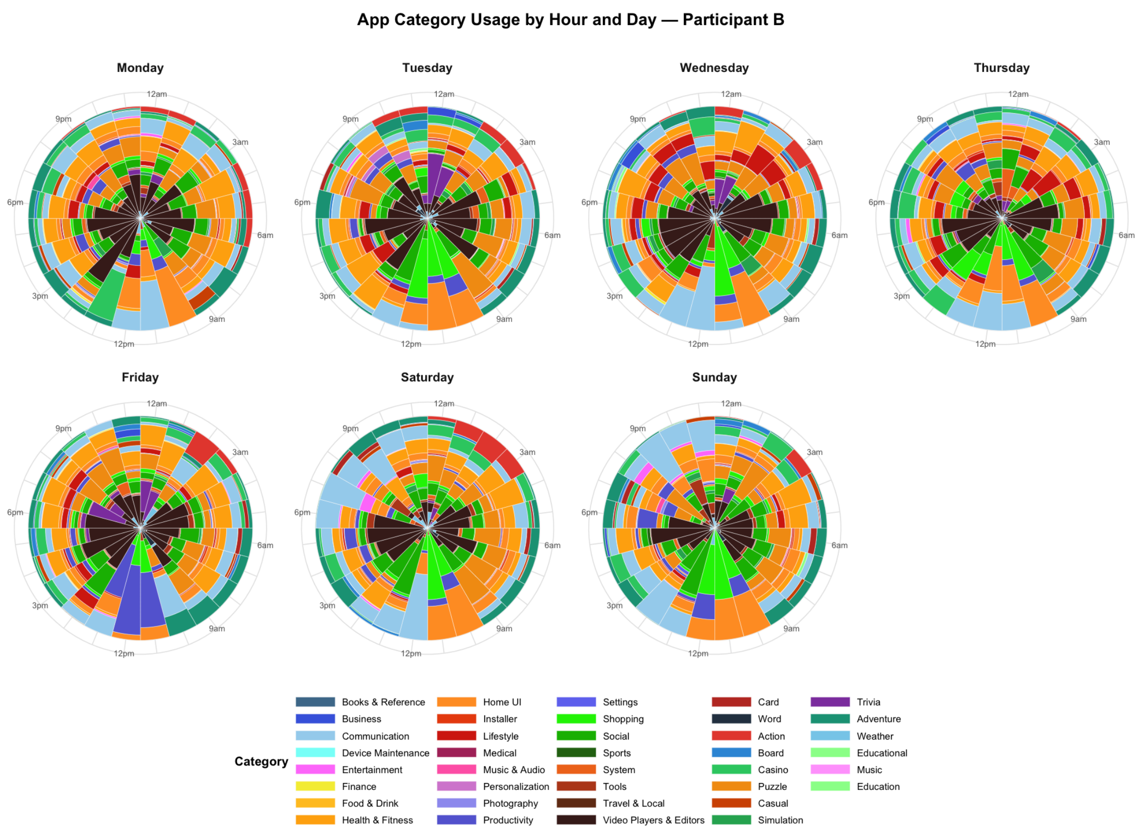
Note. *Bar colors indicate app category, or app name for Tumblr and Reddit. Colored diamonds indicate a screenshot with moderated content, with distinct colors representing a unique category. Valence and arousal of screenshots collected over the vignette is depicted on the bottom panel.*

threatening language about strangulation, vulgar language used as insults, and more. Participant B was particularly captivated by this post, indicated by increased viewing time for this post relative to content displayed around the same time. Interestingly, she was presented with another text post of similar length soon after viewing this one, where the latter did not appear to depict graphic or explicit content. Here, she scrolled past the post, choosing not to engage with it (Figure 12).

Vignette 2B: The participant was scrolling through the main feed of Tumblr when they were suddenly and unexpectedly shown sexually explicit content depicting a naked female body, with nipples, pubis, and vulva censored by emojis. Prior to viewing this content, she was seemingly passively scrolling their feed, indicated by shorter viewing times and lack of interactive engagement

Figure 13

Participant B's App Category by Usage per Weekday per Hour Over Study Period



Note. *Clock plots depict average app usage by day and hour over the study collection period, indicating health and fitness; shopping; travel and local; and communication apps were were in most frequent use throughout the week, with a spike in productivity app usage on Fridays.*

on content pieces. However, she displayed higher viewing time on, and even scrolled back to this image after scrolling away. Shortly after, she came across a fan-fiction post discussing self-harm (Figure 12).

Vignette 3B: Upon being prompted the graphic and explicit material described above, the next interaction with social media displayed drastic differences in content type. The participant opened Reddit and sought out subreddits (particular threads or pages) that displayed intense violent sexual content – primarily depicting BDSM sexual practices, and images or videos of people engaging in sexual intercourse (Figure 12).

5.3 Summary of Behavioral Vignette Patterns

The case study participants exhibited contrasting prototypical interactive behavior on social media platforms: Participant A displayed passive-seeking behavior with the content on their screen, where he opened an app, scrolled through the algorithmically curated feed, and interacted with content that piqued his attention. In contrast, Participant B displayed active-seeking behavior, where they opened an app and queries a search tool for their desired content, community, or thread type before scrolling through the feed. Despite differences in content-seeking behavior, both participants experienced unexpected algorithmic exposure to graphic content, which influenced them to intentionally seek out additional graphic content. Both vignettes illustrated that, for these users, graphic explicit content tended to be ”stickier” and more captivating than non-graphic content of equivalent mediums – particularly for sexual content and graphic violence.

Vignettes 3A and 1B illustrated how these users preferentially engaged with graphic explicit content over non-graphic content, with Participant A opting to view a video with a sexual thumbnail, and Participant B reading a text-post with violence language but skipping a subsequent non-graphic post of similar length. In both instances, it appeared the elicit nature of graphic content captivated the user, causing them to display decreased interest in other types of content. These vignettes in particular contribute to understanding how phenomena like ”click-bait” occur, where posts lure users into engagement by depicting enticing content.

Vignettes 1A and 2B involve unexpected algorithmic exposure to sexual content, which was distracting and led to a loss of initial goal for both participants. Interestingly, in both cases the participants exhibited higher viewing times for the sexual content, and returned back to the content briefly after scrolling away from it. Here, it would appear that not only was graphic sexual content ”stickier”, but it lured these users back in even after they had decided to scroll away.

Finally, vignettes 2A and 3B highlighted how exposure to graphic content can influence behavior across platforms beyond immediate responses to exposure. For participant A, viewing a video depicting graphic violence was followed by seeking out additional graphic explicit content depicting sexual acts, vulgar hateful language, and other graphic violence. Here, it appears the high arousal of a video depicting graphic violence caused him to seek out additional reactionary or explicit content to mimic the emotive response. For participant B, exposure to algorithmically prompted sexual imagery and content related to self-injury was followed with the user seeking violent sexual material on Reddit. In both instances, exposure to graphic explicit content influenced the participants' subsequent actions, with the former participant seeking out additional graphic content across categories and the latter seeking out somewhat related graphic content.

Overall, the qualitative analysis indicated exposure to high arousal graphic content led these participants to seek out additional content to mimic the evocative nature of these materials. Throughout analysis, these patterns held for additional behavioral segments not described above.

Additionally, some behavior appeared to be shaped by content moderation policies on platforms: When Participant B was on Tumblr, they largely sought out non-graphic, non-explicit content; when they were on Reddit, they largely sought out intense, extreme sexual content. This behavior may be indicative of how this user separated their identity per platform, or perhaps how content moderation guidelines shape user interactions on applications. Tumblr has stricter content moderation policies for sexual content than Reddit, prohibiting visual depictions of sexual acts and displays of real human genitals or female-presenting nipples, where Reddit allows such content, potentially providing a simple framework for understanding a stark difference in content seeking per platform.

6 Discussion

6.1 Purpose

The ubiquitous digitization of everyday life has allowed operationally opaque algorithms to shape online interactions and behaviors, particularly type and content of media to which users are exposed. Algorithmic social media platforms are intentionally engineered to optimize for engagement (NVIDIA, n.d; Narayanan, 2023), exploiting psychologically vulnerable users through recommendation algorithms that prioritize engagement over well-being by promoting harmful or explicit content to boost retention (Eklund et al., 2021; Koob & Volkow 2016). Problematic and compulsive phone use highlight the severity of these consequences, and can exacerbate pre-existing states of cognitive-affective distress (Garland et al., 2019). Additionally, harmful content that is disallowed on platforms often makes its way through the moderation process, but addressing moderation practices is difficult due to the ineffectiveness of automated moderation (Milosevic et al., 2022; Bertoldini, 2025), and the psychological harm that comes with human moderation (Spence et al., 2023). While there is an increasing amount of literature aiming to address the harms of algorithmic practices and content moderation, it's evident that little is known about repeated algorithmic exposure to graphic content and how this shapes well-being, nor the behavioral patterns that precede and follow such exposure in naturalistic contexts. This research addresses this gap through a longitudinal mixed-methods design that integrates the mapping of a quantitative analysis of fortnightly surveys onto a qualitative analysis of naturalistic smartphone screenshots.

6.2 Summary of Results

The mixed-effects models underscored both dispositional and momentary negative cognitive-affective states are associated with reduced media stickiness. Across all three models, trait-level effects acted as a more robust predictor, suggesting that those who are dispositionally vulnerable to negative affect, negative cognition, and depressive symptoms are less likely to remain engaged with media over time. When the use of graphic apps was accounted for, momentary increases in depressive symptoms were associated with increases in stickiness, and interactions between app type and state psychological distress were meaningfully moderated by chronic psychological distress *and* the type of apps being used. However, pure statistical analysis failed to capture the nuance of individual moment-to-moment smartphone use behavior, which was illuminated through the qualitative case studies.

Despite the differences in typical media seeking behavior for both case study participants (passive vs. active), both experienced instances where they were (1) algorithmically exposed to graphic explicit content, and (2) sought out graphic content on their own. For both, unexpected exposure to graphic content caused such content to be particularly sticky, indicated by higher viewing times relative to other content, and the user returning to this content after scrolling away. This held for graphic content compared to equivalent mediums of non-graphic content (text, image, video), highlighting that although graphic content takes many forms, it remains stickier relative to non-graphic content of similar format. Additionally, exposure to graphic content appeared to influence how users interacted across platforms; namely, after exposure, users sought out additional types of graphic content across categories. As described in the vignette findings, exposure to content depicting graphic violence led Participant A to seek content depicting sexual content, vulgar language, additional graphic violence, and other sensitive content. For Participant B, algorithmic exposure to nudity and self-injury related content on Tumblr led them to seek out graphically violent sexual material on Reddit, followed by engagement with a text post depicting violent and threatening language. These behavioral vignettes indicate how unprompted exposure to graphic explicit material prompted these two users to seek out additional types of graphic content shortly after.

In order to determine whether there was a relationship between cognitive affective states and exposure to graphic content, I centered the analysis on: (1) which apps allowed participants to access graphic content and how frequent was exposure, (2) whether graphic and explicit content was stickier than other types of content, and (3) whether graphic content exposure was associated with changes in cognitive-affective states.

App classification through app categorization surveys partially indicated which apps allowed study participants' to intentionally or unintentionally access graphic content (RQ 1) by labeling apps as "graphic purpose" and "graphic access" categories, but further distinction is required to determine whether graphic content was intentionally or unintentionally sought out by the user. This is because two apps that allow access to graphic content might differ in the requirements to show such content: for example, an app that requires an explicit user query to display graphic content compared to one that potentially displays unprompted graphic content through algorithmic prompting. Additionally, the case studies highlighted examples of both unintentional and intentional seeking of graphic content within a single app, indicating further nuance regarding how a single app shapes exposure.

When and how often participants were exposed to graphic content (RQ 1.1) was illuminated through case study analysis: quantitative screenshot coding highlighted the proportion of total screenshots depicting moderated content relative to all others, and qualitative observational analysis clarified whether content was displayed unexpectedly through algorithmic prompting or intentionally sought out by the user.

While the case studies indicated graphic and explicit content was stickier relative to other types of content (RQ 1.2), this finding is only applicable to these case study participants and cannot be generalized due to small sample size. Additionally, while survey data for both participants indicates they were under cognitive-affective distress during the case study, there is not enough information to determine whether graphic content exposure is related to this distress (RQ 2), nor whether an existing relationship is dependent on the type of graphic content viewed (RQ 2.1).

6.3 Limitations and Future Directions

Constraints in answering these research questions were driven by the study design and lack of comprehensive qualitative analysis. The survey data that fueled the quantitative analysis was self-reported and collected every two weeks, causing survey responses to reflect a participants' perceived psychological state in retrospect, encompassing the entire two-week period. Naturally, this method of collection may fail to capture fluctuations in day-to-day or moment-to-moment changes in psychological experiences. In addition, the difference in temporal scale between the surveys and screenshots taken every five seconds made it difficult to investigate a relationship between cognitive-affective state and content consumption. Specifically, the behavioral vignettes from the case studies occurred over less than an hour of total screen use for each participant, and it is not clear whether these sessions significantly influenced the participants' survey results. Future research may consider more frequent survey collection to narrow this gap. Additionally, due to lower inter-rater agreement across the app categorization surveys, other app categories could not be meaningfully leveraged in this analysis. Due to an emphasis on the role social media platforms play in algorithmic exposure to graphic content, future analysis should focus on these apps in particular as a predictor variable for stickiness.

Case studies were conducted for only two participants due to the time-consuming process of manually annotating each screenshot. In addition, these participants were similar in demographic background, screen time, and levels of reported depression symptoms. Future research should expand the case study by including participants of varying demographic backgrounds, levels of indicated depressive symptoms, screen-time, and engagement with apps that provide access to graphic content. Analysis would also benefit from an increased time-period for observing content behavior to draw conclusions about deviations from typical behavior and changes over time leveraging the longitudinal data.

As an observational study, a significant limitation was in understanding the intentionality of the user. The analysis of the behavioral vignettes was subject to interpretive bias and lacked participant input into the drivers of their interactions with particular content types. Future research may consider additional raters in order to account for this more comprehensively. Additionally, causal relationships cannot be interpreted from this analysis, and future research should examine the directionality of the effects from the findings and explore whether graphic content engagement reflects active emotion regulation or incidental exposure. Thus, while the case studies appear to support the notion that graphic content is stickier than non-graphic content, it isn't certain whether this is due to a user's psychological state of mind, the content of the material itself, or both.

Lastly, data collection occurred during a specific sociocultural and temporal context at the height of the COVID-19 pandemic. Future research should include a more contemporary time period to reflect changes in social media platform policies and algorithms, as well as potential changes in participant psychological states due to the distressing nature of the pandemic.

6.4 Implications for Practice

Much of the current literature is focused on summary statistics: what content users engage with the most, how long they spend on an application, and what generates higher viewing time across populations. The findings presented in this thesis highlight an underexamined aspect of user behavior online – ironically, the actual behavior itself. By constructing a digital phenotype that maps psychological states onto moment-to-moment phone behavior in a naturalistic longitudinal sample, this analysis illustrates the dynamic, person-dependent relationship between digital media interaction and cognitive-affective states of users, and how this relationship is shaped through the design of applications and platforms. The case study vignette analysis in particular helps to illuminate nuanced behaviors that are otherwise unseen in mere quantitative analysis, perhaps suggesting an unmet need for meticulous, longitudinal, cross-platform analysis for individual users.

From a practical standpoint, these findings illustrate a compelling picture for media platform design and legal framework curation, where understanding what keeps emotionally vulnerable individuals engaged with online content is important for avoiding psychological distress and related adverse consequences on human well-being. The indicated stickiness of graphic and explicit content in particular raises important questions about who algorithmic systems are reaching, given that exposure to harmful content can exacerbate and reinforce pre-existing psychological distress (Garland et al., 2019). Additionally, the trait by state interaction in the models further solidifies the importance of individual differences in understanding how algorithmic content exposure shapes user experience individuals high in trait negative cognitive-affective states appear to be disproportionately affected by momentary states when it comes to media engagement, suggesting that algorithmically delivered content may have more significant effects on these users. This in particular highlights a need for affect-sensitive algorithmic design that is responsive to engagement metrics and the cognitive-affective states of users in real time, and a legal framework limiting the algorithmic centering of engagement that distributes emotionally evocative content to users, especially those who appear to be in vulnerable psychological states.

The literature surrounding the addictive features of online algorithms, problematic phone use, and exposure to harmful content is growing in light of rapidly evolving and prevalent access to technology. Of more modern interest are strategies that can be implemented to suppress these consequences: researchers advise interventions ranging from cognitive-behavioral therapy (Bong et al., 2021) to smartphone usage tracking application (Rahmillah et al., 2023) to mitigate problematic use of smartphones. Others have recommended exercise (Liu et al., 2022) or altering screen UI by setting screen display to grayscale (Olsen et al., 2022). However, despite intervention efforts a fundamental conundrum takes hold: the features of smartphone applications that make them addictive are ubiquitous, and with the prevalence of the digital economy, decreasing or opting-out of use inevitability raises barriers. Without access to smartphones, individuals place themselves in an infrastructural exile away from the default operational and communication channels of society. Advised interventions for decreasing smartphone use is entirely user-dependent, placing the burden of resistance on users and requiring them to continually battle against systems engineered specifically to maintain their attention, leaving them in a stubborn feedback loop of distress. Thus, perhaps a more effective intervention strategy is one that focuses efforts on proposing revisions to the systems at the center of the fallout.

6.5 Policy Recommendation

Considering the consequences of graphic content exposure online that occur due to engagement-driven recommendation algorithms and faulty content moderation systems, I hope to leverage this thesis to propose a set of revisions to content moderation policies toward a proactive, user-centered design distinct from the existing reactive removal framework. These revisions are focused on automated moderation, user flagging, user tagging and filtering, and robust content moderation transparency from platforms.

Platforms should be required to invest in improving the sensitivity and specificity of automated content moderation models, reducing the volume of disturbing content that requires human review. Considering the current necessity of human review and model trainers, when human intervention is necessary platforms should be mandated to implement exposure restrictions on the duration and volume of content an individual moderator reviews in a given period. Additionally, mandated psychological support like trauma-informed counseling needs to be readily available as an employment requirement. In light of recent lawsuits against major platforms, they should be prohibited from cutting costs by outsourcing moderation labor to third-party contractors in jurisdictions lacking sufficient labor protections, as this system exploits vulnerable workers who are underpaid and unsupported. Further, platforms should be required to publish information regarding their labor practices for both automated and human-led moderation as part of their transparency disclosures to create accountability for the ghost work that has been historically ignored.

In addition, I propose technical recommendations for improving content moderation policies. Moderation infrastructure should be extended to generate descriptive metadata regarding removed material to increase regulatory oversight, which promotes accountability and may highlight patterns in the effectiveness of automated content moderation for distinct types of content. Drawing on the content tagging features of Reddit and Tumblr, platforms should allow users to tag and filter posts by content category, allowing the collaborative contextualization of potentially harming or disturbing materials. In addition, automated moderation algorithms should be trained to tag sensitive content to assist with carrying the burden of accountability. Filtering tools can allow users to block categories of sensitive content – for example, a user struggling with body image may want to block all content related to diets or eating disorders – allowing users to opt in or out of exposure that could be detrimental to their psychological well-being.

Further, platforms should integrate scaffolded content moderation policies. For categories of content that is allowed but has documented psychological risk for vulnerable populations, such as content related to eating disorders or self harm, platforms should limit exposure by age-group and harmful content interaction. Similar to TikTok's existing framework, such content should be completely blocked for users under 18 and should be ineligible to be displayed through algorithmic prompting. Additionally, by integrating a content tagging system, platforms can collect data on a user's engagement with harmful content, potentially downweighing similar content when a user appears to be engaging with this content disproportionately.

Implementing these recommendations would require sustained investment from platforms, meaningful regulatory legal frameworks, and continued empirical research to evaluate the effectiveness of these interventions on user well-being. The history of social media regulation indicates progress will be incremental and contested both at the legal and platform level, but the difficulty of pursuing these changes does not diminish its urgency and importance. The evidence presented indicates that users predisposed to cognitive-affective distress are among those who become trapped in an intentionally designed reinforcement loop with harmful, unmoderated content. A human-centered approach to platform policy is a necessary reorientation away from the existing framework reliant on engagement, exposure, and exploitation, toward centering the psychological well-being of the very users that fuel these platform economies. My hope with these suggestions is that they will be evaluated, tested, and continuously critiqued and revised as advocacy, research, and legislative efforts move toward a human-centered framework for developing and regulating technological systems.

6.6 Conclusion

This research endeavor aimed to address a meaningful gap in the present literature by examining how algorithmic exposure to unmoderated graphic and explicit content interacted with psychological states and patterns in user behavior. Through a longitudinal mixed-methods approach, this research offers a naturalistic account of how digital media engagement is shaped by the design of embedded platform algorithms, even without user awareness. Across analyses, consistent evidence emerged that dispositional and momentary cognitive-affective states are meaningfully associated with media engagement. Graphic content access was a significant moderator of these relationships, suggesting emotionally evocative content draws distressed users back into engagement. The qualitative behavioral analysis illustrated behaviors preceding and following algorithmic exposure and sought access to graphic content across platforms, highlighting the stickiness and influence of graphic explicit content.

This research establishes that algorithmic platforms are active conduits of content, serving as catalysts for cognitive-affective experiences, with disproportionate consequences on those under existing distress. Here, I hope to inform human-centered platform design sensitive to well-being, involving meaningful regulatory frameworks and a commitment by platforms to center users over existing engagement metrics governing platform economies. The users most harmed by these systems are those least equipped to resist their addictive design; while engagement remains the primary metric for algorithmic decision-making, these consequences will persist. More than anything, these findings indicate a necessary revision to existing legal frameworks and regulatory oversight to revise and reshape platform design and accountability in order to ensure the well-being and flourishing of all users.

7 References

L. (n.d.). *Lost in Translation: The Global South and the Flaws of Content Moderation | Al Jazeera Media Institute.* Retrieved 2026-05-18, from https://institute.aljazeera.net/en/ajr/article/3419

4chan. (2026). *Rules - 4chan.* Retrieved 2026-04-19, from https://4chan.org/rules

Abuhamdah, S. M. A., & Naser, A. Y. (2023, November). Smart phone addiction and its mental health risks among university students in Jordan: a cross-sectional study. *BMC psychiatry*, *23*(1), 812. doi: 10.1186/s12888-023-05322-6

Americas, T. L. R. H. APACrefauthors (2023, March). Screen violence: a real threat to mental health in children and adolescents. *Lancet Regional Health - Americas*, *19*, 100473. Retrieved 2026-04-21, from https://pmc.ncbi.nlm.nih.gov/articles/PMC10025407/ doi: 10.1016/j.lana.2023.100473

Amirthalingam, J., & Khera, A. (n.d.). Understanding Social Media Addiction: A Deep Dive. *Cureus*, *16*(10), e72499. Retrieved 2026-04-21, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11594359/ doi: 10.7759/cureus.72499

Anderson, L. (2012, June). *Snuff: Murder and torture on the internet, and the people who watch it.* Retrieved 2026-05-23, from https://www.theverge.com/2012/6/13/3076557/snuff-murder-torture-internet-people-who-watch-it

Aspen, E., & Singh. (n.d.). *The flaws in the content moderation system: The Middle East case study.* Retrieved 2026-05-23, from https://mei.edu/publication/flaws-content-moderation-system-middle-east-case-study/

Attwood, F. (2014, November). Immersion: 'extreme' texts, animated bodies and the media. *Media, Culture & Society*, *36*(8), 1186–1195. Retrieved 2026-05-23, from https://doi.org/10.1177/0163443714544858 doi: 10.1177/0163443714544858

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015, October). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. Retrieved 2026-04-22, from https://doi.org/10.18637/jss.v067.i01 doi: 10.18637/jss.v067.i01

Battisti, E., Alfiero, S., & Leonidou, E. (2022, November). Remote working and digital transformation during the COVID-19 pandemic: Economic–financial impacts and psychological drivers for employees. *Journal of Business Research*, *150*, 38–50. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC9186428/ doi: 10.1016/j.jbusres.2022.06.010

Benvenuti, M. F., Bernardy, J. L., de Toledo, T. N., Santiago, J., Siqueira, J. d. O., & Silveira, P. S. P. (2024, September). The role of schedules of reinforcement in the experimental analysis of social behavior. *Perspectives on Behavior Science*, *47*(4), 717–738. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11582125/ doi: 10.1007/s40614-024-00421-5

The blind spots in AI moderation: Research reveals systematic failures. (n.d.). Retrieved 2026-05-23, from https://www.hertie-school.org/en/news/detail/content/simon-munzert-blind-spots-in-ai-moderation

Bong, S. H., Won, G. H., & Choi, T. Y. (2021, February). Effects of Cognitive-Behavioral Therapy Based Music Therapy in Korean Adolescents with Smartphone and Internet Addiction. *Psychiatry Investigation*, *18*(2), 110–117. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC7960754/ doi: 10.30773/pi.2020.0155

Bonta, R. (2025, November). *Protecting Our Kids from Social Media Addiction Act (SB 976).* Retrieved 2026-04-19, from https://oag.ca.gov/sb976

Booth, R., & editor, R. B. U. t. (2024, December). More than 140 Kenya Facebook moderators diagnosed with severe PTSD. *The Guardian*. Retrieved 2026-05-08, from https://www.theguardian.com/media/2024/dec/18/kenya-facebook-moderators-sue-after-diagnoses-of-severe-ptsd

Brinberg, M., Ram, N., Wang, J., Sundar, S. S., Cummings, J. J., Yeykelis, L., & Reeves, B. (2023, July). Screenertia: Understanding "Stickiness" of Media Through Temporal Changes in Screen Use. *Communication Research*, *50*(5), 535–560. Retrieved 2026-04-19, from https://doi.org/10.1177/00936502211062778 doi: 10.1177/00936502211062778

Buolamwini, J., & Gebru, T. (2018, January). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency* (pp. 77–91). PMLR. Retrieved 2026-04-19, from https://proceedings.mlr.press/v81/buolamwini18a.html

California SB396 | 2025-2026 | Regular Session. (n.d.). Retrieved 2026-04-19, from https://legiscan.com/CA/text/SB396/id/3123861

Cerit, M., Lee, A. Y., Hancock, J., Miner, A., Cho, M.-J., Muise, D., ... Reeves, B. (2025, February). Person-Specific Analyses of Smartphone Use and Mental Health: Intensive Longitudinal Study. *JMIR Formative Research*, *9*(1), e59875. Retrieved 2026-04-19, from https://formative.jmir.org/2025/1/e59875 doi: 10.2196/59875

Complaint for injunctive and other relief. (2011, April). *Biotechnology Law Report*, *30*(2), 273–279. Retrieved 2026-04-19, from https://journals.sagepub.com/doi/full/10.1089/blr.2011.9967 doi: 10.1089/blr.2011.9967

Cortés-Albornoz, M. C., Ramírez-Guerrero, S., García-Guáqueta, D. P., Vélez-Van-Meerbeke, A., & Talero-Gutiérrez, C. (2023, September). Effects of remote learning during COVID-19 lockdown on children's learning abilities and school performance: A systematic review. *International Journal of Educational Development*, *101*, 102835. doi: 10.1016/j.ijedudev.2023.102835

Crowhurst, S., & Hosseinzadeh, H. (2024, June). Risk Factors of Smartphone Addiction: A Systematic Review of Longitudinal Studies. *Public Health Challenges*, *3*(2), e202. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC12039634/ doi: 10.1002/puh2.202

de Alarcón, R., de la Iglesia, J. I., Casado, N. M., & Montejo, A. L. (2019, January). Online Porn Addiction: What We Know and What We Don't—A Systematic Review. *Journal of Clinical Medicine*, *8*(1), 91. Retrieved 2026-04-19, from https://www.mdpi.com/2077-0383/8/1/91 doi: 10.3390/jcm8010091

De', R., Pandey, N., & Pal, A. (2020, December). Impact of digital surge during Covid-19 pandemic:

A viewpoint on research and practice. *International Journal of Information Management*, *55*, 102171. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC7280123/ doi: 10.1016/j.ijinfomgt.2020.102171

Ding, K., Shen, Y., Liu, Q., & Li, H. (2023, December). The Effects of Digital Addiction on Brain Function and Structure of Children and Adolescents: A Scoping Review. *Healthcare*, *12*(1), 15. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC10779052/ doi: 10.3390/healthcare12010015

Dutton. (2026, May). *Snuff and Stuff: A History of Internet Shock Content.* Retrieved 2026-06-01, from https://horrificscribblings.com/snuff-and-stuff-a-history-of-internet-shock-content/

Elhai, J. D., Dvorak, R. D., Levine, J. C., & Hall, B. J. (2017, January). Problematic smartphone use: A conceptual overview and systematic review of relations with anxiety and depression psychopathology. *Journal of Affective Disorders*, *207*, 251–259. doi: 10.1016/j.jad.2016.08.030

Epskamp, S., Deserno, M., & Bringmann, L. (n.d.). *mlVAR: Multi-Level Vector Autoregression.* Retrieved 2026-04-22, from https://sachaepskamp.r-universe.dev/mlVAR

Erdős, C., Kelemen, O., Pócs, D., Paulik, E., Papp, A., Horváth, E., . . . Széll, K. (2025, January). Pornography-Watching Disorder and Its Risk Factors Among Young Adults: Cross-Sectional Survey. *Journal of Medical Internet Research*, *27*(1), e49860. Retrieved 2026-04-19, from https://www.jmir.org/2025/1/e49860 doi: 10.2196/49860

Foster-McCray, S. (2024). When Algorithms See Us: An Analysis of Biased Corporate Social Media Algorithm Programming and the Adverse Effects These Social Media Algorithms Create When They Recommend Harmful Content to Unwitting Users Comments. *Southern Journal of Policy and Justice*, *18*, 1–28. Retrieved 2026-05-08, from https://heinonline.org/HOL/P?h=hein.journals/srebwsude18&i=6

Garland, E. L., Froeliger, B., & Howard, M. O. (2014, January). Mindfulness training targets neurocognitive mechanisms of addiction at the attention-appraisal-emotion interface. *Frontiers* *in Psychiatry*, *4*, 173. doi: 10.3389/fpsyt.2013.00173

Georgia SB351 | 2023-2024 | Regular Session. (n.d.). Retrieved 2026-04-19, from https://legiscan.com/GA/text/SB351/id/2973892

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press.

Goldman, E. (2025, November). *Colorado's Mandatory Social Media "Warning Labels" Are Unconstitutional-NetChoice v. Weiser.* Retrieved 2026-04-19, from https://blog.ericgoldman.org/archives/2025/11/colorados-mandatory-social-media-warning-labels-are-unconstitutional-netchoice-v-weiser.htm

Governor Lamont and Attorney General Tong Announce Legislation To Combat Youth Social Media Addiction. (n.d.). Retrieved 2026-04-19, from https://portal.ct.gov/governor/news/press-releases/2026/01-2026/governor-lamont-attorney-general-tong-announce-legislation-to-combat-youth-social-media-addiction

Grant, J. E., Lust, K., & Chamberlain, S. R. (n.d.). Problematic smartphone use associated with greater alcohol consumption, mental health issues, poorer academic performance, and impulsivity. *Journal of Behavioral Addictions*, *8*(2), 335–342. Retrieved 2026-05-18, from https://pmc.ncbi.nlm.nih.gov/articles/PMC6609450/ doi: 10.1556/2006.8.2019.32

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016, November). Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges. *Perspectives on psychological science : a journal of the Association for Psychological Science*, *11*(6), 838–854. Retrieved 2026-03-23, from https://pmc.ncbi.nlm.nih.gov/articles/PMC5572675/ doi: 10.1177/1745691616650285

Hardy, J., & Stewart, C. (n.d.). *Gore and violent extremism: How extremist groups exploit 'gore' sites to view and share terrorist material.* Retrieved 2026-05-23, from https://www.isdglobal.org/digital-dispatch/gore-and-violent-extremism

-how-extremist-groups-exploit-gore-sites-to-view-and-share-terrorist-material/

Hartmann, D., Oueslati, A., Staufer, D., Pohlmann, L., Munzert, S., & Heuer, H. (2025, April). Lost in Moderation: How Commercial Content Moderation APIs Over- and Under-Moderate Group-Targeted Hate Speech and Linguistic Variations. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–26). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-05-23, from https://dl.acm.org/doi/10.1145/3706598.3713998 doi: 10.1145/3706598.3713998

Instagram. (2020). *Introducing Instagram Reels.* Retrieved 2026-04-19, from https://about.instagram.com/blog/announcements/introducing-instagram-reels-announcement

James, L. (2025, October). *Attorney General James Announces that Social Media Companies Must Start Reporting Their Content Moderation Policies to Combat Online Hate.* Retrieved 2026-04-19, from https://ag.ny.gov/press-release/2025/attorney-general-james-announces-social-media-companies-must-start-reporting

James, R. J., O'Malley, C., & Tunney, R. J. (2019, January). Gambling on Smartphones: A Study of a Potentially Addictive Behaviour in a Naturalistic Setting. *European Addiction Research*, *25*(1), 30–40. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC6482978/ doi: 10.1159/000495663

Jerab, D. (2025, January). *The Influence of Emotional Triggers and Social Sharing Behaviors on the Virality of Marketing Campaigns across Different Digital Platforms* [SSRN Scholarly Paper]. Rochester, NY: Social Science Research Network. Retrieved 2026-06-01, from https://papers.ssrn.com/abstract=5092293 doi: 10.2139/ssrn.5092293

Johar, A. (2026, March). *TikTok, Meta compromised safety in algorithm race — report.* Retrieved 2026-05-23, from https://www.mobileworldlive.com/meta/tiktok-meta-compromised-safety-in-algorithm-race-report/

Johnson. (n.d.). SENATE BILL 2097.

Keeler, J., Conde Ludtke, L., Yang, Q., Raschke Rameh, V., Ward, R., Treasure, J., & Carter, B. (2026, March). Associations of Problematic Smartphone Use and Smartphone Screen Time With Eating Disorder Psychopathology in Non-Clinical Samples: A Systematic Review. *JMIR Mental Health*, *13*, e88572. Retrieved 2026-05-18, from https://pmc.ncbi.nlm.nih.gov/articles/PMC12980065/ doi: 10.2196/88572

Kelly, C. A., & Sharot, T. (2025, January). Web-browsing patterns reflect and shape mood and mental health. *Nature Human Behaviour*, *9*(1), 133–146. Retrieved 2026-05-08, from https://www.nature.com/articles/s41562-024-02065-6 doi: 10.1038/s41562-024-02065-6

Kelly, H. (n.d.). *Why you may see violent or weird posts on Instagram, Facebook and TikTok - The Washington Post.* Retrieved 2026-06-01, from https://www.washingtonpost.com/technology/2025/03/13/social-media-algorithm-strange/

Kim, I., Boffa, J., Cho, M., Conroy, D. E., Kline, N., Haber, N., ... Ram, N. (2025, October). *Stanford Screenomics: An Open-source Platform for Unobtrusive Multimodal Digital Trace Data Collection from Android Smartphones.* medRxiv. Retrieved 2026-03-23, from https://www.medrxiv.org/content/10.1101/2025.06.24.25329707v2 (ISSN: 3067-2007 Pages: 2025.06.24.25329707) doi: 10.1101/2025.06.24.25329707

Kim, I., Boffa, J., Cho, M., Conroy, D. E., Kline, N., Haber, N., ... Ram, N. (2026, March). An open-source platform for multimodal digital trace data collection from smartphones. *Nature Health*, 1–12. Retrieved 2026-03-23, from https://www.nature.com/articles/s44360-026-00072-7 doi: 10.1038/s44360-026-00072-7

Kim, S. A. (2017). Social Media Algorithms: Why You See What You See Technology Explainers. *Georgetown Law Technology Review*, *2*(1), 147–154. Retrieved 2026-04-19, from https://heinonline.org/HOL/P?h=hein.journals/gtltr2&i=145

Kiyak, C., Cetinkaya, D., McAlaney, J., Hodge, S., & Ali, R. (2024, September). Interrupting Dissociation of Players through Real-Time Digital Tasks during Online Gambling. *International Journal of Human–Computer Interaction*, *40*(18), 5247–5258. Retrieved

2026-05-08, from https://doi.org/10.1080/10447318.2023.2233127 (_eprint: https://doi.org/10.1080/10447318.2023.2233127) doi: 10.1080/10447318.2023.2233127

Koller, M. (2016, December). robustlmm: An R Package for Robust Estimation of Linear Mixed-Effects Models. *Journal of Statistical Software*, *75*, 1–24. Retrieved 2026-04-22, from https://doi.org/10.18637/jss.v075.i06 doi: 10.18637/jss.v075.i06

Koob, G. F., & Volkow, N. D. (2016, August). Neurobiology of addiction: a neurocircuitry analysis. *The Lancet. Psychiatry*, *3*(8), 760–773. doi: 10.1016/S2215-0366(16)00104-8

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017, December). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. Retrieved 2026-04-22, from https://doi.org/10.18637/jss.v082.i13 doi: 10.18637/jss.v082.i13

Lee, M., Lee, A. P. M., & Press, A. (2026, March). *Jury finds Meta's platforms are harmful to children in 1st wave of social media addiction lawsuits.* Retrieved 2026-06-01, from https://www.pbs.org/newshour/nation/jury-finds-metas-platforms-are-harmful-to-children-in-1st-wave-of-social-media-addiction-lawsuits (Section: Nation)

Linkedin. (2020, June). *LinkedIn and Microsoft help 25 million job seekers get back to work.* Retrieved 2026-04-19, from https://news.linkedin.com/2020/june/linkedin-and-microsoft-help-25-million-job-seekers-get-back-to-w

Liu, H., Soh, K. G., Samsudin, S., Rattanakoses, W., & Qi, F. (2022, October). Effects of exercise and psychological interventions on smartphone addiction among university students: A systematic review. *Frontiers in Psychology*, *13*, 1021285. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC9580397/ doi: 10.3389/fpsyg.2022.1021285

Lost in Content Moderation. (n.d.). Retrieved 2026-05-23, from https://www.weizenbaum-institut.de/news/detail/lost-in-content-moderation/

Louisiana Laws - Louisiana State Legislature. (n.d.). Retrieved 2026-04-19, from https://

legis.la.gov/legis/LawPrint.aspx?d=1337818

McMath, P. (2023, June). *Secure Online Child Interaction and Age Limitation Act.*

Meta. (n.d.). *Community Standards | Transparency Center.* Retrieved 2026-04-19, from https://transparency.meta.com/policies/community-standards/

Metzler, H., & Garcia, D. (2024, September). Social Drivers and Algorithmic Mechanisms on Digital Media. *Perspectives on Psychological Science*, *19*(5), 735–748. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11373151/ doi: 10.1177/17456916231185057

Milli, S., Carroll, M., Wang, Y., Pandey, S., Zhao, S., & Dragan, A. (n.d.). *Engagement, User Satisfaction, and the Amplification of Divisive Content on Social Media.* Retrieved 2026-05-23, from http://knightcolumbia.org/content/engagement-user-satisfaction-and-the-amplification-of-divisive-content-on-social-media

Milosevic, T., Van Royen, K., & Davis, B. (2022). Artificial Intelligence to Address Cyberbullying, Harassment and Abuse: New Directions in the Midst of Complexity. *International Journal of Bullying Prevention*, *4*(1), 1–5. Retrieved 2026-05-18, from https://pmc.ncbi.nlm.nih.gov/articles/PMC8872854/ doi: 10.1007/s42380-022-00117-x

Minguillon, J., Lopez-Gordo, M. A., Renedo-Criado, D. A., Sanchez-Carrion, M. J., & Pelayo, F. (2017, October). Blue lighting accelerates post-stress relaxation: Results of a preliminary study. *PLoS ONE*, *12*(10), e0186399. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC5648169/ doi: 10.1371/journal.pone.0186399

Muñoz-Sastre, D., Rodrigo-Martín, L., & Rodrigo-Martín, I. (2021, November). The Role of Twitter in the WHO's Fight against the Infodemic. *International Journal of Environmental Research and Public Health*, *18*(22), 11990. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC8621779/ doi: 10.3390/ijerph182211990

Narayanan, A. (2023). Understanding Social Media Recommendation Algorithms. Retrieved 2026-04-21, from https://doi.org/10.7916/khdk-m460 doi: 10.7916/khdk-m460

Nawaz, F. A., Riaz, M. M. A., Banday, N. u. a., Singh, A., Arshad, Z., Derby, H., & Sultan, M. A.

(2024, February). Social media use among adolescents with eating disorders: a double-edged sword. *Frontiers in Psychiatry*, *15*, 1300182. Retrieved 2026-04-21, from https://pmc.ncbi.nlm.nih.gov/articles/PMC10884122/ doi: 10.3389/fpsyt.2024.1300182

Ndayambaje, E., & Okereke, P. U. (2025, May). The Psychopathology of Problematic Smartphone Use (PSU): A Narrative Review of Burden, Mediating Factors, and Prevention. *Health Science Reports*, *8*(5), e70843. doi: 10.1002/hsr2.70843

Nouvian, T. (2025, January). *Families sue TikTok in France over teen suicides they say are linked to harmful content | AP News.* Retrieved 2026-04-19, from https://apnews.com/article/tiktok-france-trial-suicide-lawsuit-fa8f979c3121a3c5712d52a300c9005f

NVIDIA. (n.d.). *What is a Recommendation System?* Retrieved 2026-04-21, from https://www.nvidia.com/en-us/glossary/recommendation-system/

NY State Senate Bill 2023-S7694A. (2023). Retrieved 2026-04-19, from https://www.nysenate.gov/legislation/bills/2023/S7694/amendment/A

Olson, J. A., Sandra, D. A., Chmoulevitch, D., Raz, A., & Veissière, S. P. L. (2022, May). A Nudge-Based Intervention to Reduce Problematic Smartphone Use: Randomised Controlled Trial. *International Journal of Mental Health and Addiction*, 1–23. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC9112639/ doi: 10.1007/s11469-022-00826-w

Parker, S., Odio, S., & Mosseri, A. (2012, January). *Infinite Scrolling* (No. US20120011430A1). Retrieved 2026-04-19, from https://patents.google.com/patent/US20120011430A1/en

Paterna, A., Alcaraz-Ibáñez, M., Aguilar-Parra, J. M., Salavera, C., Demetrovics, Z., & Griffiths, M. D. (2024, April). Problematic smartphone use and academic achievement: A systematic review and meta-analysis. *Journal of Behavioral Addictions*, *13*(2), 313–326. Retrieved 2026-05-18, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11220804/ doi: 10.1556/2006.2024.00014

Paxton, K. (n.d.). *Texas HB20 | 2025 | 89th Legislature 2nd Special Session.* Retrieved 2026-04-19, from https://legiscan.com/TX/bill/HB20/2025/X2

Pedersen, T. L. (n.d.). *The Composer of Plots.* Retrieved 2026-05-08, from https://patchwork.data-imaginist.com/

Pellegrino, A., Stasi, A., & Bhatiasevi, V. (2022, November). Research trends in social media addiction and problematic social media use: A bibliometric analysis. *Frontiers in Psychiatry*, *13*, 1017506. Retrieved 2026-04-21, from https://pmc.ncbi.nlm.nih.gov/articles/PMC9707397/ doi: 10.3389/fpsyt.2022.1017506

Perez, S. (2020, November). *Facebook's Snapchat-like 'Vanish Mode' feature arrives on Messenger and Instagram.* Retrieved 2026-04-19, from https://techcrunch.com/2020/11/12/facebooks-snapchat-like-vanish-mode-feature-arrives-on-messenger-and-instagram/

Posit | Data Science Platform for Enterprise Teams. (n.d.). Retrieved 2026-04-22, from https://posit.co/

Potenza, M. N. (2013, August). Neurobiology of Gambling Behaviors. *Current opinion in neurobiology*, *23*(4), 660–667. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC3803105/ doi: 10.1016/j.conb.2013.03.004

Rahmillah, F. I., Tariq, A., King, M., & Oviedo-Trespalacios, O. (2023, August). Evaluating the Effectiveness of Apps Designed to Reduce Mobile Phone Use and Prevent Maladaptive Mobile Phone Use: Multimethod Study. *Journal of Medical Internet Research*, *25*, e42541. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC10498313/ doi: 10.2196/42541

Reddit. (n.d.). *Reddit Rules.* Retrieved 2026-04-19, from https://redditinc.com/policies/reddit-rules

Reeves, B., Robinson, T., & Ram, N. (2020, January). Time for the Human Screenome Project. *Nature*, *577*(7790), 314–317. Retrieved 2026-04-19, from https://www.nature.com/articles/d41586-020-00032-5 (Bandiera_abtest: a Cg_type: Comment Subject_term: Media) doi: 10.1038/d41586-020-00032-5

Rejer, I., Jankowski, J., Dreger, J., & Lorenz, K. (2024, January). Viewer Engagement in Response to Mixed and Uniform Emotional Content in Marketing Videos—An Electroencephalographic Study. *Sensors (Basel, Switzerland)*, *24*(2), 517. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC10818430/ doi: 10.3390/s24020517

Research, E. (2024, September). *The True Crime Consumer Report by Edison Research and audiochuck.* Retrieved 2026-06-01, from https://www.edisonresearch.com/the-true-crime-consumer-report-by-edison-research-and-audiochuck/

Revelle, W. (2026, March). *psych: Procedures for Psychological, Psychometric, and Personality Research.* Retrieved 2026-04-22, from https://cran.r-project.org/web/packages/psych/index.html

Roberts, S. (2019). *Behind the Screen: Content Moderation in the Shadows of Social Media.* doi: https://doi.org/10.2307/j.ctvhrcz0v

Ross, E. (2024, October). *What is Causing Our Epidemic of Loneliness and How Can We Fix It? | Harvard Graduate School of Education.* Retrieved 2026-04-19, from https://www.gse.harvard.edu/ideas/usable-knowledge/24/10/what-causing-our-epidemic-loneliness-and-how-can-we-fix-it

Satici, S. A., Gocet Tekin, E., Deniz, M. E., & Satici, B. (2023). Doomscrolling Scale: its Association with Personality Traits, Psychological Distress, Social Media Use, and Wellbeing. *Applied Research in Quality of Life*, *18*(2), 833–847. doi: 10.1007/s11482-022-10110-7

SB396 Bill Information. (n.d.). Retrieved 2026-04-19, from https://arkleg.state.ar.us/Bills/Detail

Schissler, M. (2025, July). Beyond Hate Speech and Misinformation: Facebook and the Rohingya Genocide in Myanmar. *Journal of Genocide Research*, *27*(3), 445–470. Retrieved 2026-04-19, from https://doi.org/10.1080/14623528.2024.2375122 (_eprint: https://doi.org/10.1080/14623528.2024.2375122) doi: 10.1080/14623528.2024.2375122

Section 230: An Overview [legislation]. (n.d.). Retrieved 2026-04-19, from https://www.congress.gov/crs-product/R46751

Senate Bill 7072 (2021) - The Florida Senate. (2021). Retrieved 2026-04-19, from https://www.flsenate.gov/Session/Bill/2021/7072/

Sen. Bryan, R. H. D.-N. (1998, October). *S.2326 - 105th Congress (1997-1998): Children's Online Privacy Protection Act of 1998* [legislation]. Retrieved 2026-04-19, from https://www.congress.gov/bill/105th-congress/senate-bill/2326 (Archive Location: 1998-07-17)

SF 1807 Introduction - 94th Legislature (2025 - 2026). (n.d.). Retrieved 2026-04-19, from https://www.revisor.mn.gov/bills/94/2025/0/SF/1807/versions/0/

Shock site | Semantic Scholar. (n.d.). Retrieved 2026-05-23, from https://www.semanticscholar.org/topic/Shock-site/2104608

Snapchat. (2026). *Snap Inc. Launches Spotlight, a New Entertainment Platform for User Generated Content within Snapchat.* Retrieved 2026-04-19, from https://investor.snap.com/news/news-details/2020/Snap-Inc.-Launches-Spotlight-a-New-Entertainment-Platform-for-User-Generated-Content-within-Snapchat/default.aspx

Social Media Victims Law Center - Social Media Litigation Lawyers. (2026). Retrieved 2026-06-01, from https://socialmediavictims.org/

Spence, R., Bifulco, A., Bradbury, P., Martellozzo, E., & DeMarco, J. (2023, September). The psychological impacts of content moderation on content moderators: A qualitative study. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *17*(4). Retrieved 2026-05-08, from https://cyberpsychology.eu/article/view/33166 doi: 10.5817/CP2023-4-8

Stevens, T., Brevers, D., Chambers, C. D., Lavric, A., McLaren, I. P. L., Mertens, M., ... Verbruggen, F. (2015, March). How does response inhibition influence decision making when gambling? *Journal of Experimental Psychology. Applied*, *21*(1), 15–36. doi: 10.1037/xap0000039

Taskin, S., Yildirim Kurtulus, H., Satici, S. A., & Deniz, M. E. (2024, April). Doomscrolling and mental well-being in social media users: A serial mediation through mindfulness and secondary traumatic stress. *Journal of Community Psychology*, *52*(3), 512–524. doi: 10.1002/jcop.23111

Tateno, M., Teo, A. R., Ukai, W., Kanazawa, J., Katsuki, R., Kubo, H., & Kato, T. A. (2019, July). Internet Addiction, Smartphone Addiction, and Hikikomori Trait in Japanese Young Adult: Social Isolation and Social Network. *Frontiers in Psychiatry*, *10*. Retrieved 2026-04-22, from https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2019.00455/full doi: 10.3389/fpsyt.2019.00455

Thigpen, N. N., Keil, A., & Freund, A. M. (2018, February). Responding to Emotional Scenes: Effects of Response Outcome and Picture Repetition on Reaction Times and the Late Positive Potential. *Cognition & emotion*, *32*(1), 24–36. Retrieved 2026-04-21, from https://pmc.ncbi.nlm.nih.gov/articles/PMC6174536/ doi: 10.1080/02699931.2016.1266305

TikTok. (2020). *Keeping our community safe during the COVID-19 outbreak.* Retrieved 2026-04-19, from https://newsroom.tiktok.com/keeping-our-community-safe-during-the-coronavirus-outbreak?lang=en-IN

TikTok. (2026). *Policies & engagement | TikTok.* Retrieved 2026-04-19, from https://www.tiktok.com/safety/en/policies-and-engagement/overview

TikTok Lawsuit - Teen Addiction & Mental Health Harm. (n.d.). Retrieved 2026-04-19, from https://socialmediavictims.org/tiktok-lawsuit/

Tumblr. (2026). *Tumblr.* Retrieved 2026-04-19, from https://www.tumblr.com/policy/en/user-guidelines

Ungless, E. L., Markl, N., & Ross, B. (2024a, July). *Experiences of Censorship on TikTok Across Marginalised Identities.* arXiv. Retrieved 2026-04-19, from http://arxiv.org/abs/2407.14164 (arXiv:2407.14164 [cs]) doi: 10.48550/arXiv.2407.14164

Ungless, E. L., Markl, N., & Ross, B. (2024b, July). *Experiences of Censorship on TikTok Across Marginalised Identities.* arXiv. Retrieved 2026-04-19, from http://arxiv.org/abs/2407.14164 (arXiv:2407.14164 [cs]) doi: 10.48550/arXiv.2407.14164

Wakeman, G. (2024, November). *5 Bloody Spectacles at Ancient Rome's Colos-*

seum. Retrieved 2026-06-01, from https://www.history.com/articles/colosseum-gladiators-ancient-rome-spectacles (Section: Feature)

Wang, J., & Wang, S. (2025, May). The Emotional Reinforcement Mechanism of and Phased Intervention Strategies for Social Media Addiction. *Behavioral Sciences*, *15*(5), 665. Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC12108933/ doi: 10.3390/bs15050665

Wheelwright, T. (2026, January). *Cell Phone Usage Stats 2026: Americans Check Their Phones 186 Times a Day.* Retrieved 2026-05-23, from https://www.reviews.org/mobile/cell-phone-addiction/

Wickham, H. (n.d.). *Create Elegant Data Visualisations Using the Grammar of Graphics.* Retrieved 2026-05-08, from https://ggplot2.tidyverse.org/

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., . . . Yutani, H. (2019, November). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. Retrieved 2026-04-22, from https://joss.theoj.org/papers/10.21105/joss.01686 doi: 10.21105/joss.01686

X. (2026). *The X Rules: Safety, privacy, authenticity, and more.* Retrieved 2026-04-19, from https://help.x.com/en/rules-and-policies/x-rules

Xie, J., Xu, X., Zhang, Y., Tan, Y., Wu, D., Shi, M., & Huang, H. (2023, December). The effect of short-form video addiction on undergraduates' academic procrastination: a moderated mediation model. *Frontiers in Psychology*, *14*, 1298361. Retrieved 2026-04-22, from https://pmc.ncbi.nlm.nih.gov/articles/PMC10756502/ doi: 10.3389/fpsyg.2023.1298361

Xu, R. (2021, September). *Spectatorship and the Consumption of Dying at Public Executions.* Retrieved 2026-06-01, from https://cmsmc.org/opeds/spectatorship-and-the-consumption-of-dying

Xue, H., Gong, X., & Stevens, H. (2022, June). COVID-19 Vaccine Fact-Checking Posts on Facebook: Observational Study. *Journal of Medical Internet Research*, *24*(6), e38423.

Retrieved 2026-04-19, from https://pmc.ncbi.nlm.nih.gov/articles/PMC9217154/ doi: 10.2196/38423

Yan, T., Su, C., Xue, W., Hu, Y., & Zhou, H. (2024, June). Mobile phone short video use negatively impacts attention functions: an EEG study. *Frontiers in Human Neuroscience*, *18*, 1383913. Retrieved 2026-04-22, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11236742/ doi: 10.3389/fnhum.2024.1383913

Yao, Q., Gao, Y., Liu, C., Li, X., Jin, W., & Wang, Q. (2025, May). The impact of childhood trauma on short video addiction: psychological and morphological correlates. *Scientific Reports*, *15*, 18999. Retrieved 2026-05-09, from https://pmc.ncbi.nlm.nih.gov/articles/PMC12125358/ doi: 10.1038/s41598-025-04020-5

Yousef, A. M. F., Alshamy, A., Tlili, A., & Metwally, A. H. S. (2025, March). Demystifying the New Dilemma of Brain Rot in the Digital Era: A Review. *Brain Sciences*, *15*(3), 283. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC11939997/ doi: 10.3390/brainsci15030283

YouTube. (n.d.). *Building YouTube Shorts, a new way to watch & create on YouTube.* Retrieved 2026-04-19, from https://blog.youtube/news-and-events/building-youtube-shorts/

YouTube. (2026). *YouTube's Community Guidelines - YouTube Help.* Retrieved 2026-04-19, from https://support.google.com/youtube/answer/9288567?hl=en

Zhang, X., Wu, Y., & Liu, S. (2019, September). Exploring short-form video application addiction: Socio-technical and attachment perspectives. *Telematics and Informatics*, *42*, 101243. Retrieved 2026-05-09, from https://www.sciencedirect.com/science/article/pii/S0736585319303302 doi: 10.1016/j.tele.2019.101243

Zhu, W., Zhang, Y., Lan, Y., & Song, X. (2025, August). Smartphone dependence and its influence on physical and mental health. *Frontiers in Psychiatry*, *16*, 1281841. Retrieved 2026-05-08, from https://pmc.ncbi.nlm.nih.gov/articles/PMC12401922/ doi: 10.3389/fpsyt.2025.1281841

§ 59.1-577.1. (Effective January 1, 2026) Social media platforms; responsibilities and prohibitions

related to minors. (n.d.). Retrieved 2026-04-19, from https://law.lis.virginia.gov/vacode/title59.1/chapter53/section59.1-577.1/

Appendix A

Survey Instruments

Table A1

CES-D Survey Measures, where italicized variables were reverse scored

CES-D Survey Measures					
Variable Name	**Survey Question (0-3 scale)**	**Variable Name**	**Survey Question (0-3 scale)**	**Name of Selection Option**	**Value**
cesd1	I was bothered by things that don't usually bother me.	*cesd12_R*	I was happy.	Rarely or none of the time (less than 1 day)	0
cesd2	I did not feel like eating; my appetite was poor.	cesd13	I talked less than usual.	Some or little of the time (1-2 days)	1
cesd3	I feel that I could not shake off the blues even with help from my family or friends.	cesd14	I felt lonely.	Occasionally or moderate amount of time (3-4 days)	2
cesd4_R	I felt I was just as good as other people .	cesd15	People were unfriendly.	Most or all of the time (5-7 days)	3
cesd5	I had trouble keeping my mind on what I was doing.	*cesd16_R*	I enjoyed life.		
cesd6	I felt depressed.	cesd17	I had crying spells.		
cesd7	I felt that everything I did was an effort.	cesd18	I felt sad.		
cesd8_R	I felt hopeful about the future.	cesd19	I felt that people dislike me.		
cesd9	I thought my life had been a failure	cesd20	I could not get "going".		
cesd10	I felt fearful.	cesd_total	N/A - sum of all questions		
cesd11	My sleep was restless.				

Table A2

Cognitive-Affective Survey Measures; Negative Cognition (left) and Negative Affect (right)

Negative Cognition Survey Measures	
Variable Name	**Survey Question (0-100 scale)**
rumination	Looking back over the last two weeks, I felt I was thinking about the same things over and over again
attention	Looking back over the last two weeks, how difficult did you find the following things? - Keeping my attention on the tasks I was doing
concentration	Looking back over the last two weeks, how difficult did you find the following things? - Concentrtating when doing a task
restless	Looking back over the last two weeks, I felt restless or fidgety
impatient	Looking back over the last two weeks, I felt impatient
spur_of_moment	Looking back over the last two weeks, I felt I made "spur of the moment" decisions

Negative Affect Survey Measures	
Variable Name	**Survey Question (0-100 scale)**
depressed	Looking back over the last two weeks, I felt DEPRESSED
lonely	Looking back over the last two weeks, I felt LONELY
sad	Looking back over the last two weeks, I felt SAD
anxious	Looking back over the last two weeks, I felt ANXIOUS
worried	Looking back over the last two weeks, I felt WORRIED

Appendix B

Variable Descriptives

Table B1

Negative Affect, Negative Cognition, and CES-D Total Variable Descriptives

Variable	N	Mean	SD	Median	IQR	Min	Max	Skewness	Kurtosis
CES-D Total	3655	13.03	12.68	9	16	0	60	1.25	0.96
CES-D Total (Centered and Scaled)	3655	0	1	-0.32	1.261	-1.03	3.70	1.25	0.96
Trait CES-D Total (Centered and Scaled)	3655	0	1	-0.30	1.373	-1.16	3.94	1.10	0.68
State CES-D Total (Centered and Scaled)	3655	0	1	-0.07	0.820	-5.83	7.98	0.64	5.15
Negative Affect Total	3655	218.83	179.37	200	321	0	743	0.38	-1.08
Negative Affect Total (Centered and Scaled)	3655	0	1	-0.10	1.790	-1.22	2.92	0.38	-1.08
Trait Negative Affect Total (Centered and Scaled)	3655	0	1	-0.16	1.796	-1.36	2.38	0.29	-1.15
State Negative Affect Total (Centered and Scaled)	3655	0	1	-0.04	0.927	-5.81	5.52	0.13	2.95
Negative Cognition Total	3655	206.58	166.53	191	297	0	763	0.36	-1.04
Negative Cognition Total (Centered and Scaled)	3655	0	1	-0.09	1.783	-1.24	3.34	0.36	-1.04
Trait Negative Cognition Total (Centered and Scaled)	3655	0	1	-0.01	1.829	-1.36	2.46	0.25	-1.20
State Negative Cognition Total (Centered and Scaled)	3655	0	1	-0.06	0.918	-5.47	5.84	0.29	3.28

Table B2

Stickiness Variable Descriptives

Variable	N	Mean	SD	Median	IQR	Min	Max	Skewness	Kurtosis
Stickiness	3655	47443.46	109342.68	21537.56	37998.25	0	2979512.09	12.28	237.69
Log of Stickiness	3655	9.77	1.79	9.98	1.582	0	14.91	-1.75	4.85
Log of Stickiness (Centered and Scaled)	3655	0	1	0.12	0.886	-5.47	2.88	-1.75	4.85
Trait Log of Stickiness (Centered and Scaled)	3655	0	1	0.01	1.176	-6.60	4.71	-0.96	4.87

Appendix C

Variable Correlations

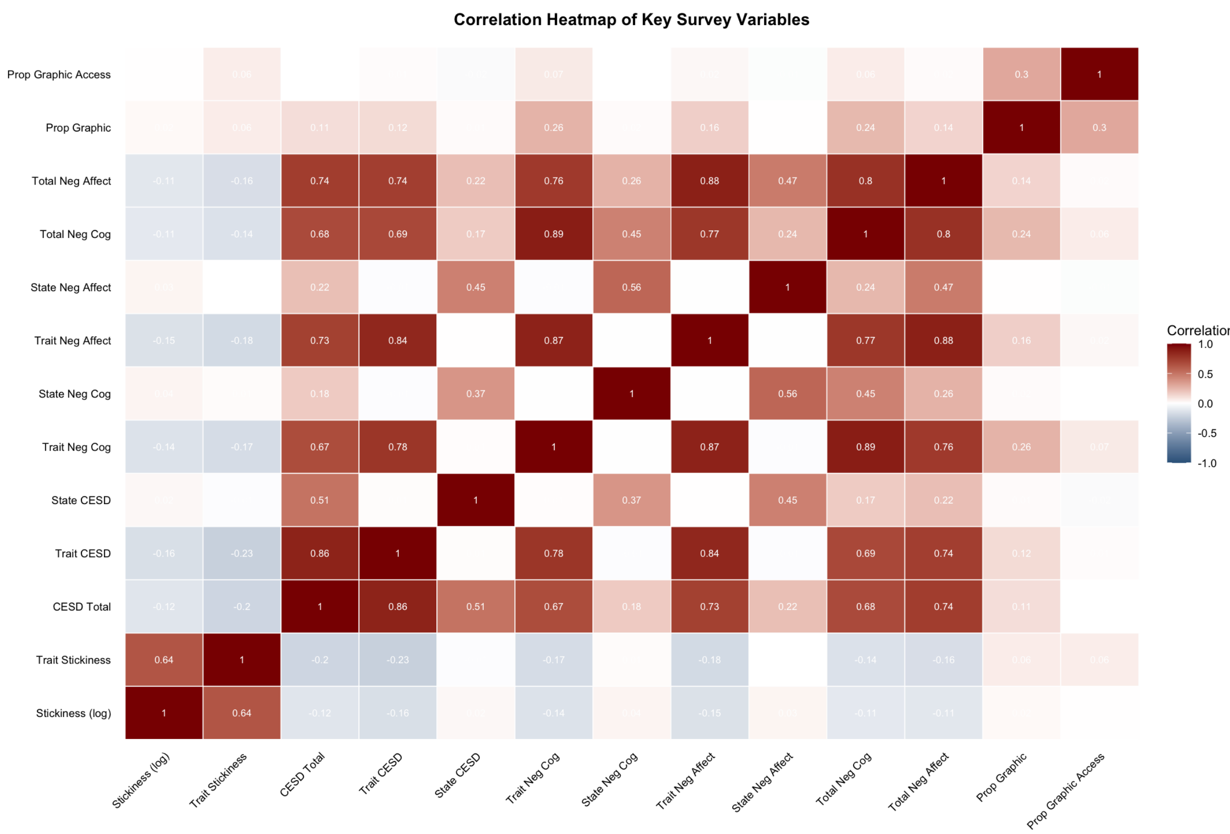
Table C1

Variable Correlations

Measure	1	2	3	4	5	6	7	8	9	10	11	12
1. Fortnight	–	-0.16	0.10	-0.10	-0.06	-0.08	-0.13	-0.05	-0.19	-0.14	-0.07	-0.19
2. Log Stickiness (Centered and Scaled)	-0.16	–	0.58	-0.11	-0.15	0.05	-0.11	-0.14	0.04	-0.10	-0.13	0.04
3. Trait Log Stickiness Total (Centered and Scaled)	-0.16	0.58	–	-0.22	-0.25	0	-0.19	-0.21	0	-0.17	-0.19	0
4. CES-D Total (Centered and Scaled)	-0.10	-0.11	-0.22	–	0.88	0.48	0.76	0.75	0.20	0.71	0.70	0.17
5. Trait CES-D Total (Centered and Scaled)	-0.06	-0.15	-0.25	0.88	–	0	0.76	0.85	-0.01	0.72	0.80	-0.02
6. State CES-D Total (Centered and Scaled)	-0.08	0.05	0	0.48	0	–	0.20	0	0.44	0.16	0	0.38
7. Negative Affect Total (Centered and Scaled)	-0.13	-0.11	-0.19	0.76	0.76	0.20	–	0.89	0.45	0.82	0.78	0.25
8. Trait Negative Affect Total (Centered and Scaled)	-0.05	-0.14	-0.21	0.75	0.85	0	0.89	–	-0.01	0.80	0.88	-0.01
9. State Negative Affect Total (Centered and Scaled)	-0.19	0.04	0	0.20	-0.01	0.44	0.45	-0.01	–	0.23	-0.01	0.56
10. Negative Cognition Total (Centered and Scaled)	-0.14	-0.10	-0.17	0.71	0.72	0.16	0.82	0.80	0.23	–	0.90	0.42
11. Trait Negative Cognition Total (Centered and Scaled)	-0.07	-0.13	-0.19	0.70	0.80	0	0.78	0.88	-0.01	0.90	–	-0.01
12. State Negative Cognition Total (Centered and Scaled)	-0.19	0.04	0	0.17	-0.02	0.38	0.25	-0.01	0.56	0.42	-0.01	–

Figure C1

Variable Correlation Heatmap



Note. *Strongest correlations depicted with opaque colors, with positive correlative relationships in red and negative ones in blue.*

Appendix D

Variable Intraclass Correlation Coefficients

Table D1

Variable Intraclass Correlation Coefficients

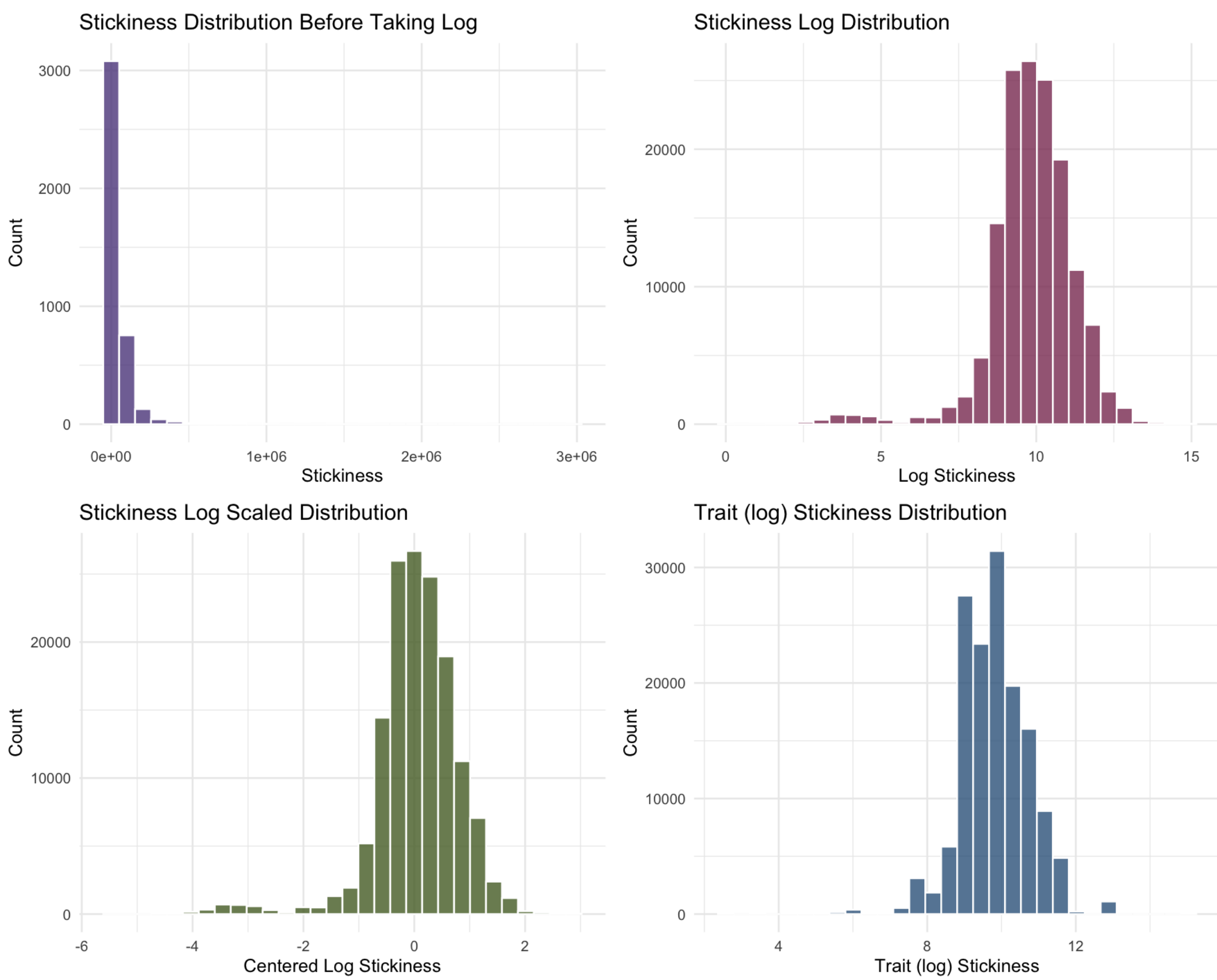
Variable	ICC
Log of Stickiness	0.40
Negative Affect Total	0.79
Negative Cognition Total	0.81
CES-D Total	0.79

Appendix E

Variable Distribution Plots

Figure E1

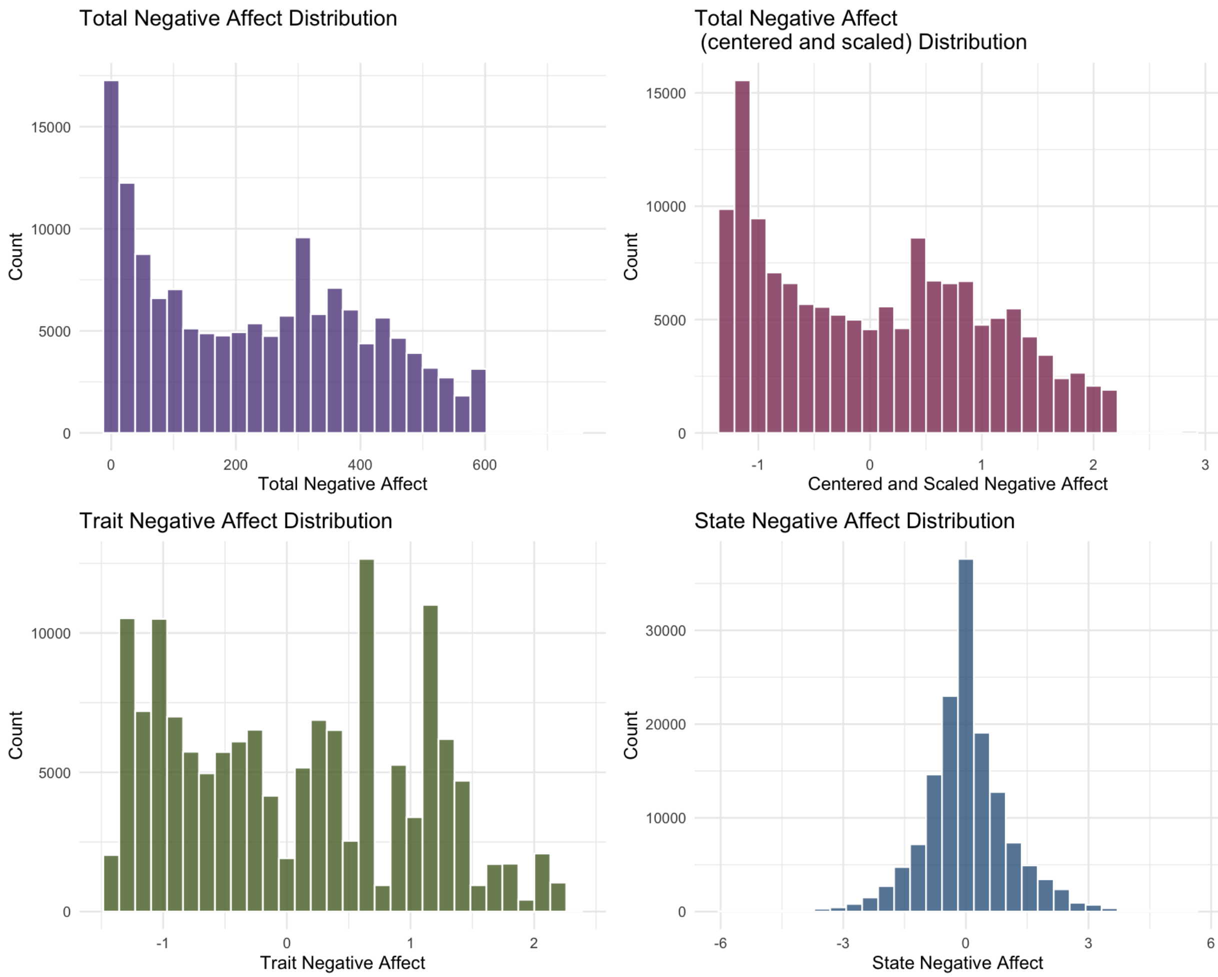
Stickiness Variable Distribution



Note. Distribution histograms of total and trait stickiness variable on various scales, with count of observations on y-axis.

Figure E2

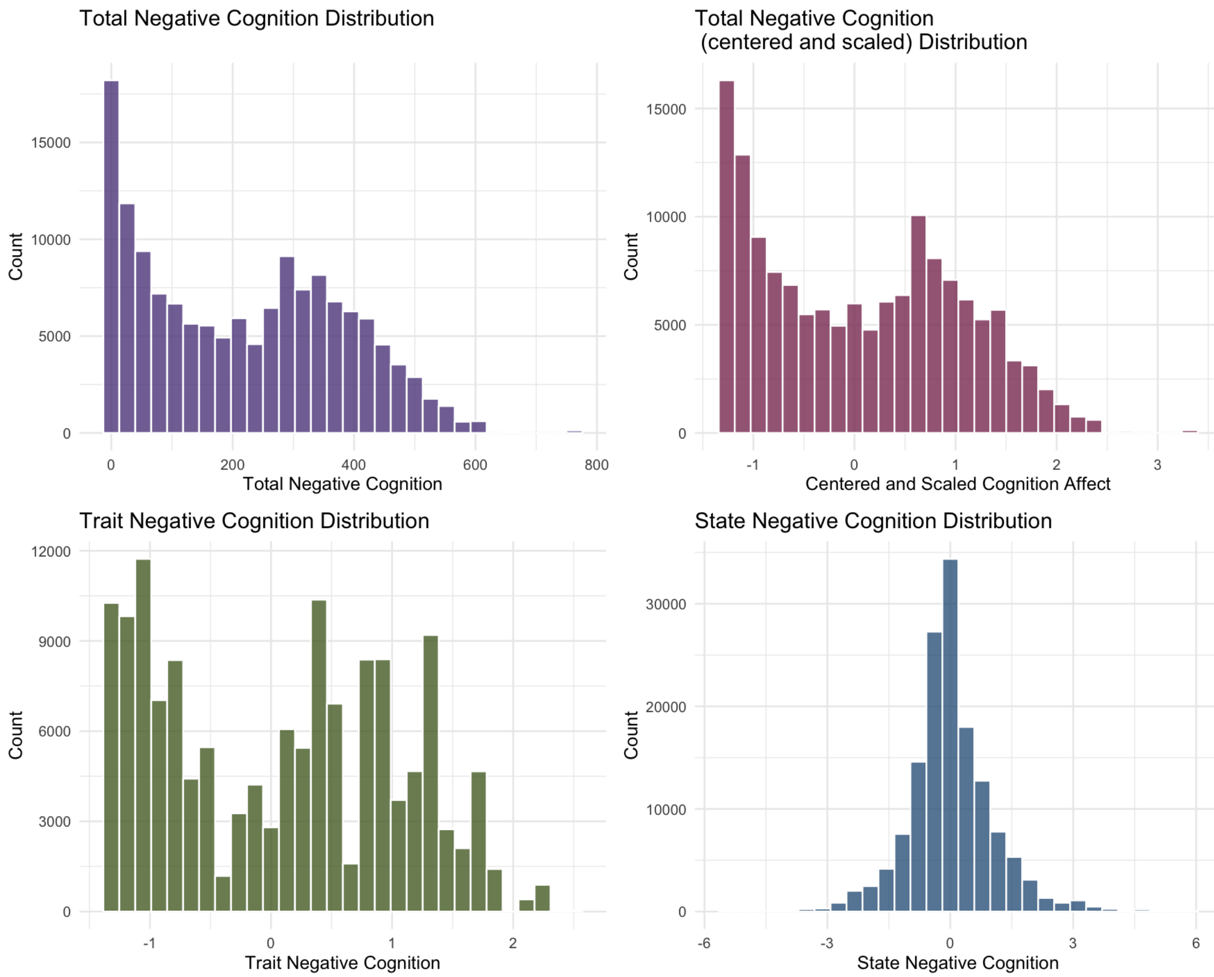
Negative Affect Variable Distribution



Note. *Distribution histograms of total, trait, and state negative affect variable on various scales, with count of observations on y-axis.*

Figure E3

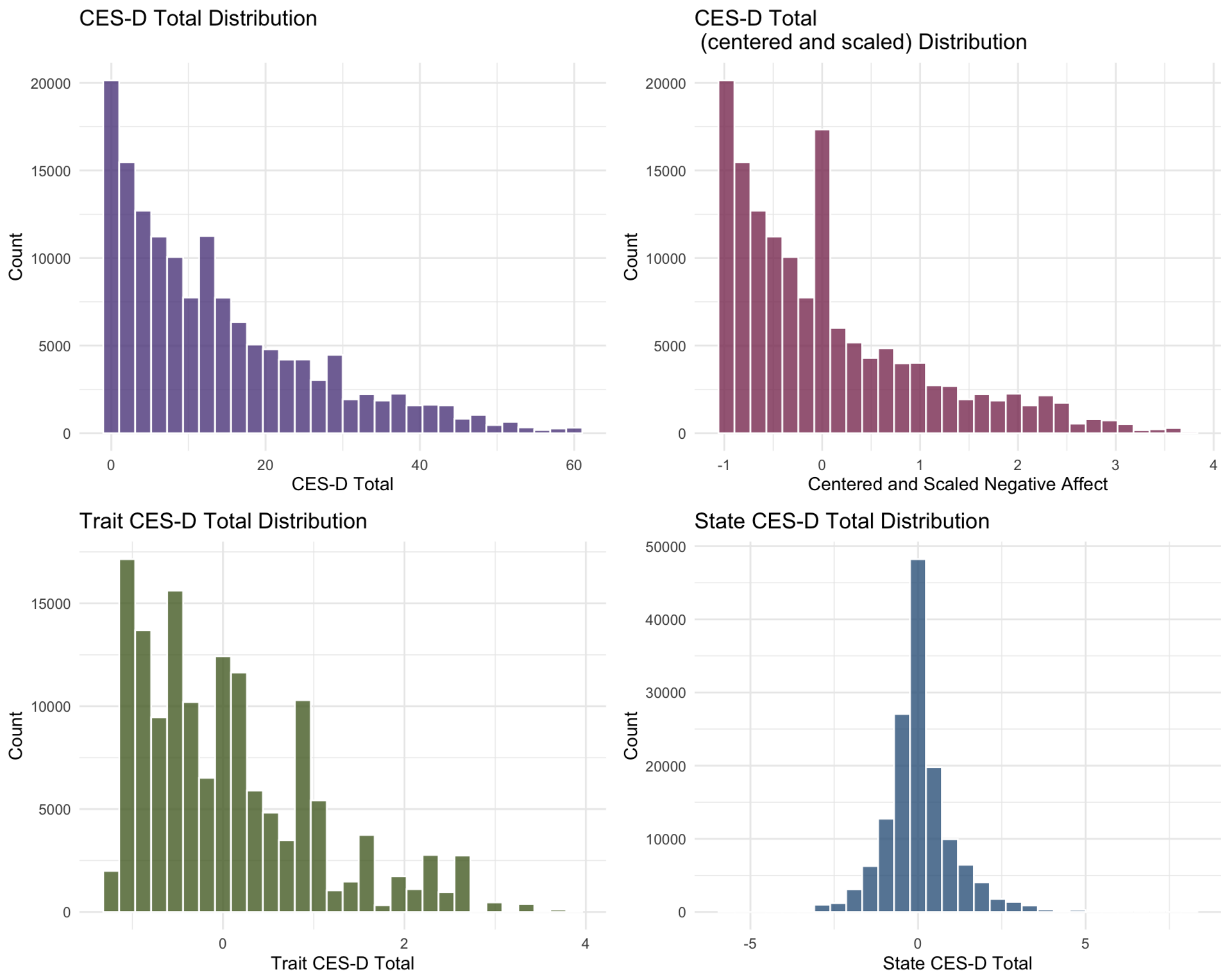
Negative Cognition Variable Distribution



Note. Distribution histograms of total, trait, and state negative cognition variable on various scales, with count of observations on y-axis.

Figure E4

CES-D Total Variable Distribution



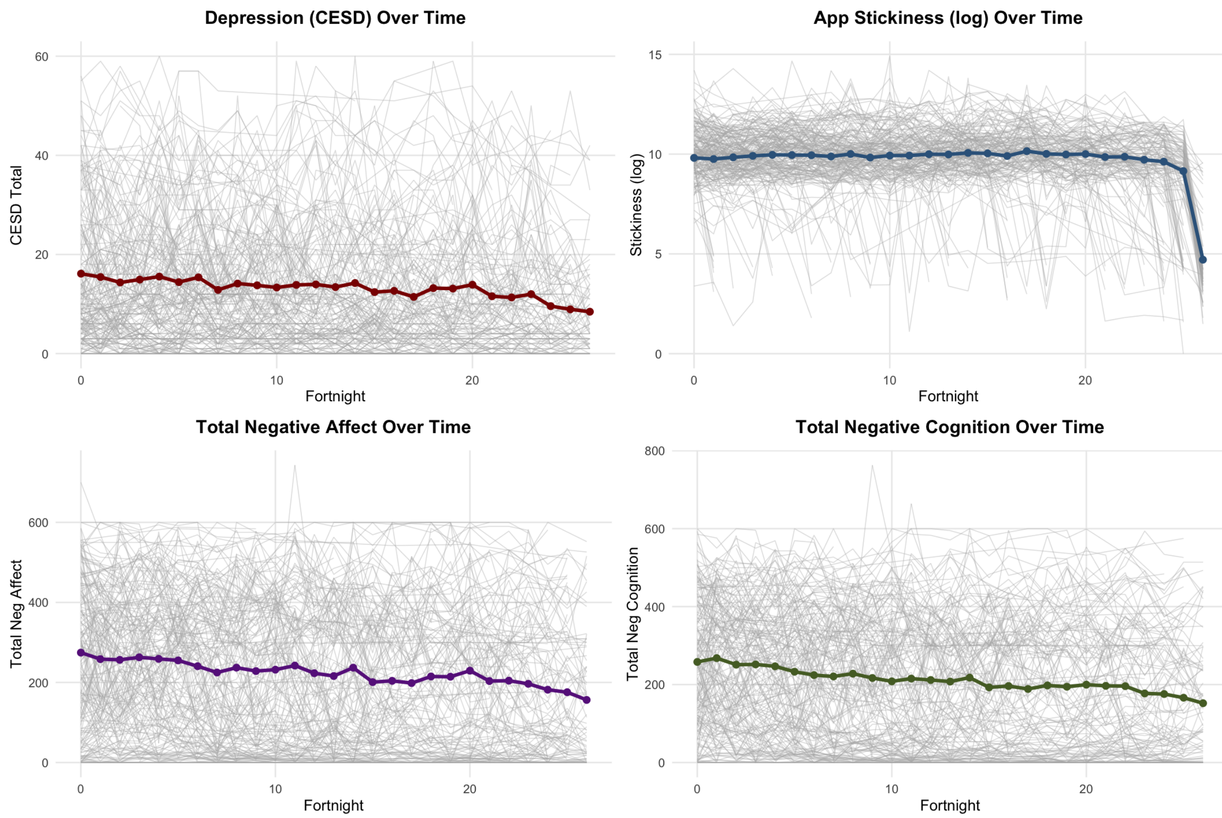
Note. *Distribution histograms of total, trait, and state CES-D variable on various scales, with count of observations on y-axis.*

Appendix F

Variable Intraclass Correlation Coefficients

Figure F1

Spaghetti Plot for all 217 Participants



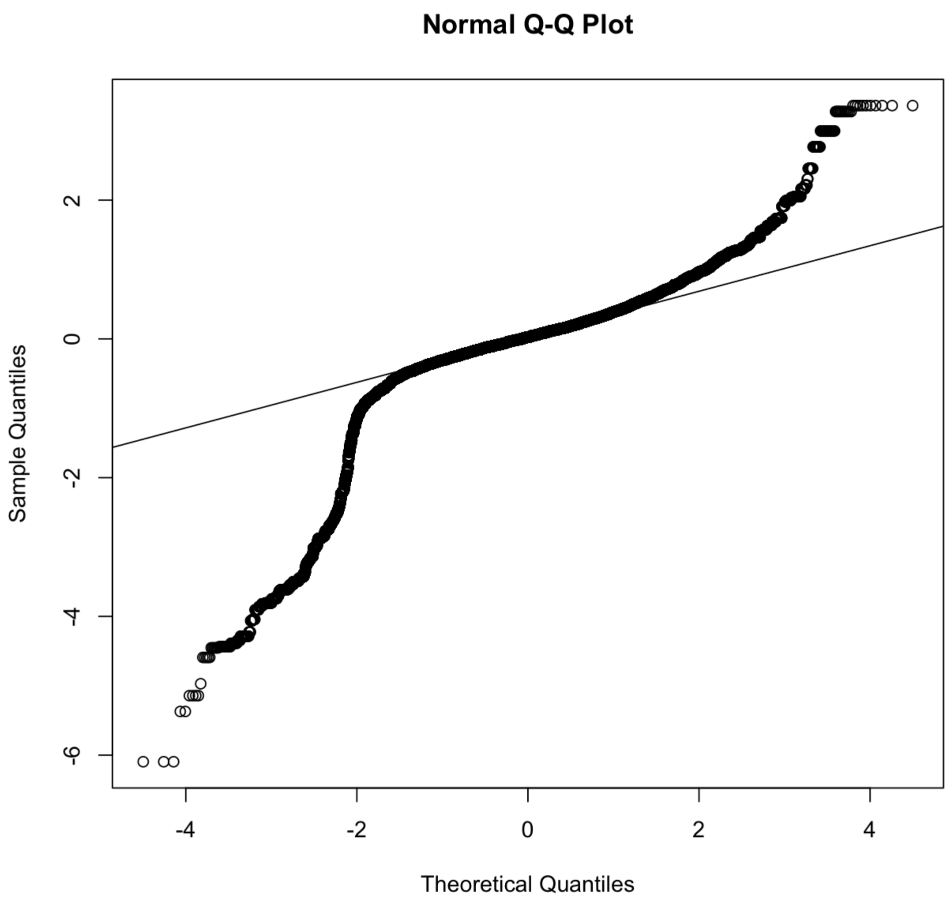
Note. *Grey lines are represent participants, where the colored line is the group mean. There is significant variation in individual cognitive-affective experiences over the study period within and between individuals, and generally the mean negative cognitive-affective states trended downward with time.*

Appendix G

Quantile-Quantile Plot for Stickiness (log z) Distribution

Figure G1

Quantile-Quantile Plot for Stickiness (log z) Distribution



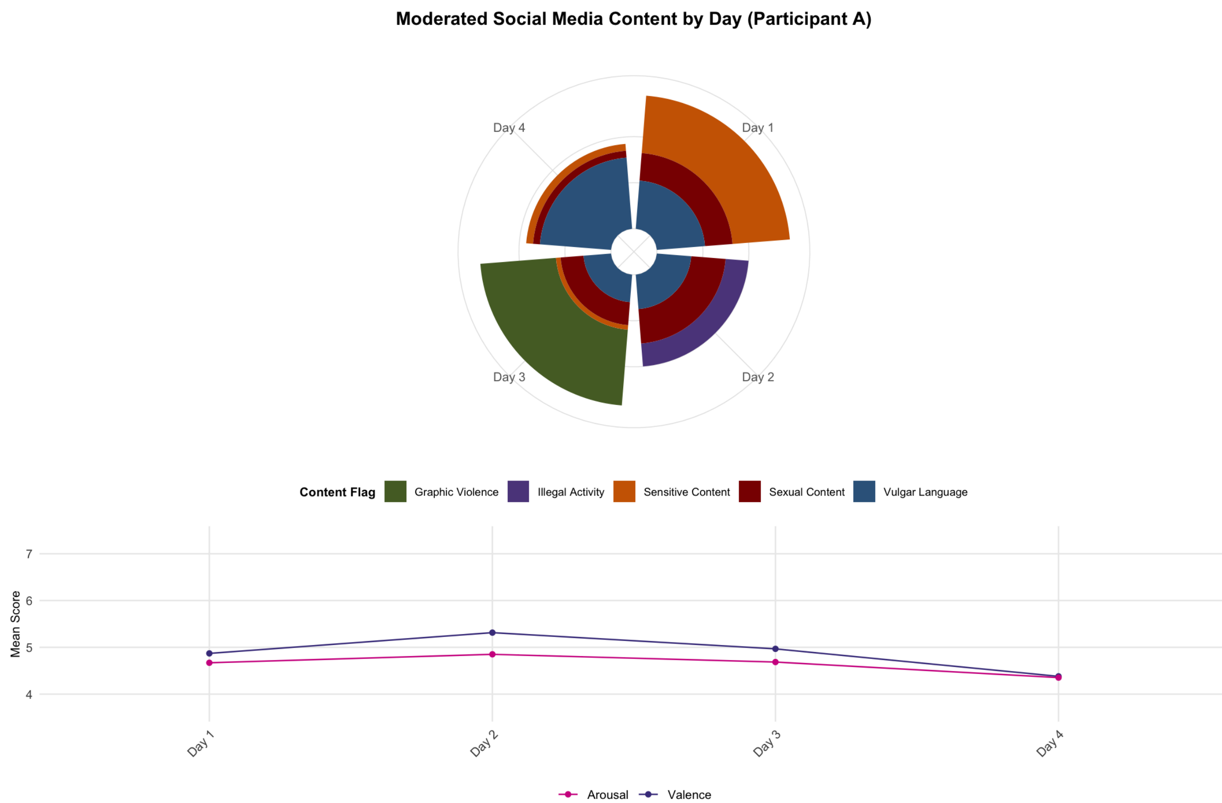
Note. The Quantile-Quantile Plot assesses whether (log z) stickiness follows a normal distribution. The deviations from the normal distribution reference line indicate many extreme highs and lows in the variable even after log transforming.

Appendix H

Participant A Proportion of Moderated Content Categories

Figure H1

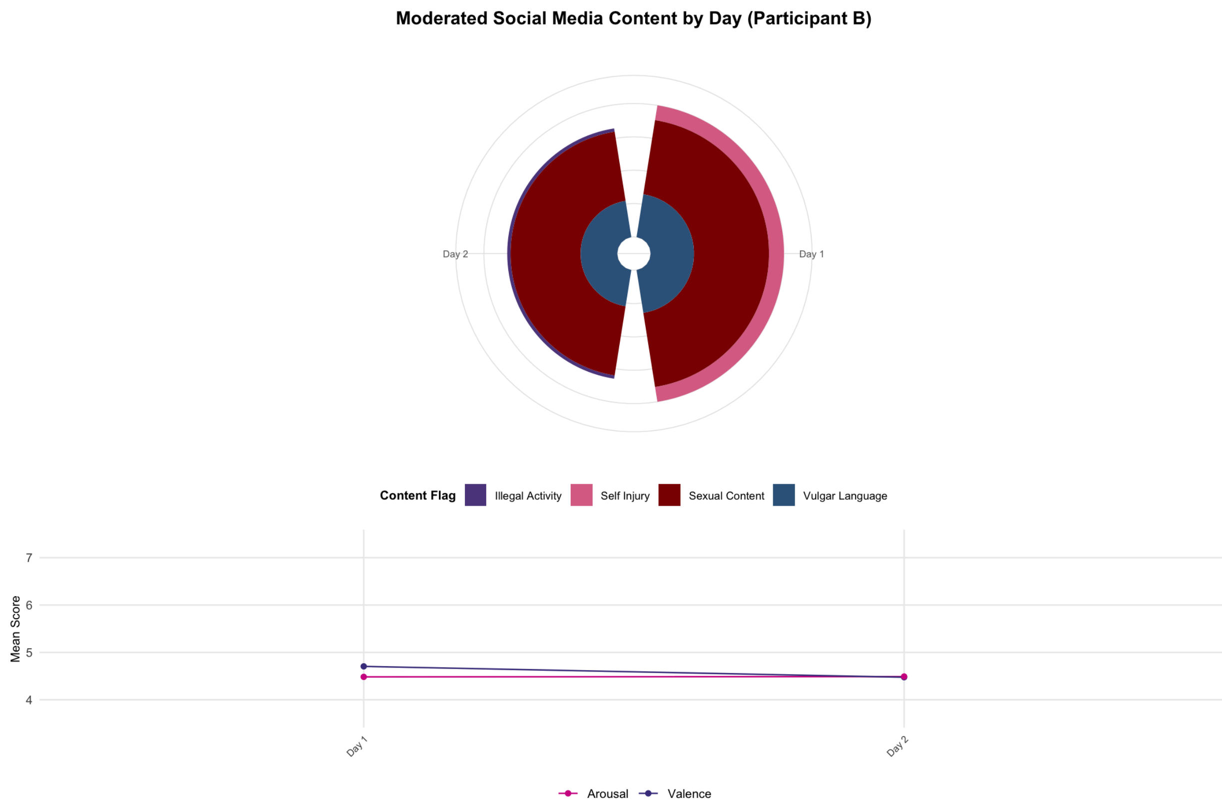
Proportion of Moderated Content Categories Over Behavioral Vignette Collection Period



Note. *Proportion of total moderated screenshots per category collected over behavioral vignette coding, with a unique color representing a distinct category. Mean valence and arousal of screenshots depicted in bottom panel. Participant A had significantly more sensitive content during the first day of tagging, and more graphic violence the third day of tagging.*

Figure H2

Proportion of Moderated Content Categories Over Behavioral Vignette Collection Period



Note. Proportion of total moderated screenshots per category collected over behavioral vignette coding, with a unique color representing a distinct category. Mean valence and arousal of screenshots depicted in bottom panel. Participant B had significantly more sexual content during both days of tagging.

Appendix I

App Coding Survey

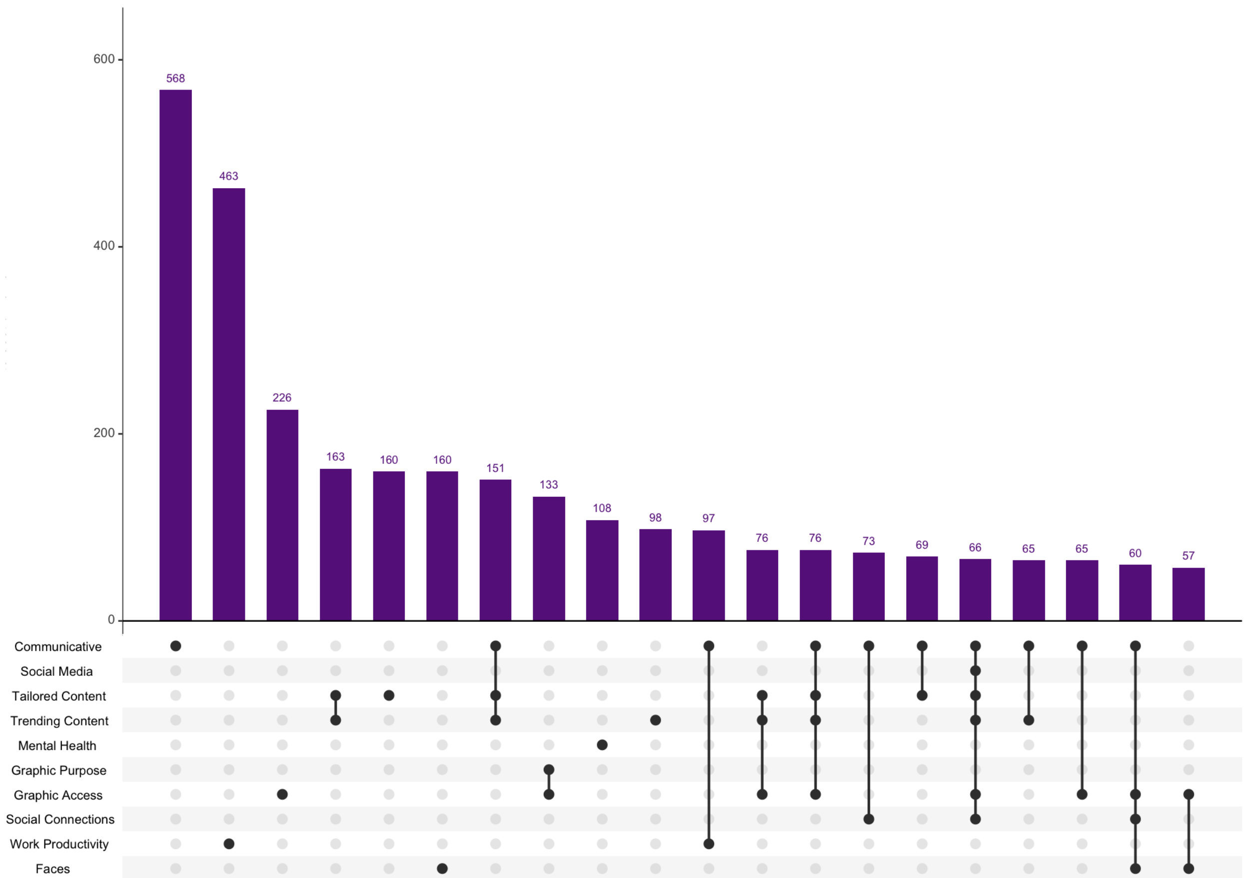
Table I1

Qualtrics App Rating Survey Questions and Descriptions

Question	Description
Is this a Communicative app?	Communicative apps refer to apps where users can communicate and exchange information with other users/entities Users may provide and/or receive information through the app. Communicative apps may or may not have large social networks. Examples include email apps, messaging apps, news media apps, or browsing apps like Safari, Google Chrome, Firefox, etc.
Is this a Social Media app?	Social media refers to digital platforms and technologies that facilitate the creation, sharing, and exchange of information, ideas, interests, and other forms of expression via virtual communities and networks. Social media platforms enable users to interact with content and other users through posts, comments, likes, shares, and direct messaging. Social media apps include apps that allow users to interact with content, even if users themselves cannot create their own content. Examples include Facebook, Twitter, Instagram, LinkedIn, and TikTok.
Does this app regularly present content tailored to a user's past behavior and interactions?	Many apps personalize users' experience using algorithms that deliver new content based on a users' past behavior and preferences. Examples of apps that use such algorithms to present content include TikTok, Spotify and YouTube.
Does the app regularly prioritize showing a user popular or trending content, even when it doesn't align with the user's usual interests?	Many apps personalize users' experience using algorithms that deliver new content based on popular topics and trends. For example X/Twitter's "Trending" section highlights topics or hashtags that are currently popular across the platform.
Is the purpose of this app to help users cultivate mental health and wellness?	Help-seeking behavior is defined as "searching for or requesting help from others via formal or informal mechanisms, such as through mental health services." We are interested in examining this help-seeking behavior in digital platforms. To do this, I will look at mobile applications for both formal help-seeking behavior (mental health apps) and informal help-seeking behavior (wellness apps). Mental Health mobile applications are mobile applications that help users assess, support, regulate, and treat their mental health. Wellness mobile applications are a broader category and include any mobile applications that aim to help people improve their well-being by promoting healthy habits and lifestyles.
Is the purpose of this app to display graphic or explicit content to users?	An application built for the purpose of displaying graphic or explicit content. This may include apps with a user rating of 18+, dating apps, pornography apps, violent video game apps, and more. These apps facilitate user's access to and viewing of graphic or explicit content. Apps that are NOT built for the purpose of displaying graphic or explicit content include apps that are specifically built for an age-restricted audience of 12 and under (e.g. YouTube kids) or specifically non-graphic content (e.g. Duolingo, Cash App, Sephora).
Does this app provide users access to graphic or explicit content?	An application that may provide a user with graphic or explicit content, consensually or non-consensually. Examples may include social media apps, internet explorers, messaging apps, and more.
Is this app designed to facilitate social interactions or connections between users?	(e.g., interpersonal relationships, community interaction, social engagement)
Is this app designed to support work-related tasks or professional productivity?	(e.g., task management, productivity tools, professional networking)
If the user encounters faces on this application, are they likely to be non-familiar or familiar faces?	N/A

Figure I1

UpSet Plot for Number of Apps Coded Across All Categories



Note. *UpSet plot shows intersections across multiple categories, similar to a Venn diagram. The majority of apps coded fell into "Communicative" and "Work Productivity" categories.*

Appendix J

Content Moderation Category Codebook

Table J1

Depiction of Screenshot Tagging Flow Conducted via Image Viewer

content type	primary categorization	primary cat type	secondary categorization	secondary cat type	tertiary categorization	tertiary cat type	description
social app	social media app name (text)	text					IE instagram, facebook, tiktok, youtube, x, reddit, 4chan
content type	1. video	categorical					Description
	2. image						
	3. text						
	4. other		descriptive tag	text			
moderated content	1. Yes	boolean					Description
	2. No						
sexual content	1. Yes	boolean	1. nudity	categorical	1. Bare Nudity	categorical	Description
					2. Semi Nudity		
					3. Eduational Nudity		
					4. Realistic Depiction		
					5. Unrealistic Depiction		
			2. education discussion		Discussion Topic	text	
			3. other (sexually suggestive)		descriptive tag	text	
			4. No				
	2. No						
graphic violence	1. humans	categorical					Description
	2. animals						
	3. other		descriptive tag	text			
illegal activity	1. Yes	boolean	1. promotes or encourages illegal activity	categorical			Description
			2. depicts illegal activity				
			3. is illegal activity		descriptive tag	text	
	2. No						
age-restricted activity		boolean	1. gambling	categorical			Description
			2. legal drug, tobacco, alcohol use				
	1. Yes		3. legal use of firearms		descriptive tag	text	
	2. No						
harassment and bullying	1. Yes	boolean	1. hateful speech	categorical			Description
			2. incites violence				
			3. attacks protected groups		descriptive tag	text	
	2. No						
self injury	1. Yes	boolean	1. self-mutilation or suicide	categorical	1. actual self-injury	categorical	Description
			2. eating disorders		2. depicts actual self injury		
			3. No		3. promotoes or glorifies		
					4. educational discussion		
	2. No						
vulgar language	1. Yes	boolean	1. swear words	categorical			Description
			2. other vulgar expression(s)		descriptive tag	text	
	2. No						
sensitive content	1. Yes	boolean					Description
	2. No		descriptive tag	text			

Table J2

Content Moderation Guidelines by Platform

Category	Nudity	Sexually suggestive content	Graphic violence	Illegal activity	Age-restricted activity	Hate speech and bullying	Self-injury	Vulgar language	Sensitive content (other)
Description	Images or depictions of individuals wearing little to no clothing.	Content that implies, hints at, or alludes to sexual themes or activity, but does not display sexual activity.	The depiction of violent acts in a way that is detailed, explicit, and often realistic. E.g. intense displays of blood, gore, mutilation, and other forms of physical harm.	Activities that are not legal.	Activities that are only legal past a particular age.	Abusive or threatening speech or writing that expresses prejudice on the basis of ethnicity, religion, sexual orientation, or similar grounds.	Depictions of one intentionally causing harm to one's body, including self mutilation, eating disorders, and more.	Words and phrases considered offensive, rude, or inappropriate in most social contexts. Includes swear words, insults, and sexually explicit or crude expressions.	Does the platform offer a sensitive content/view anyway warning? Content that is not listed in other categories but may be censored or considered sensitive.
Example	A depiction of people engaging in sexual intercourse.	An emoji combination that hints at sexual themes but does not directly say it is referring to sex.	A depiction of two people engaging in a physical fight that results in bloodshed.	Depictions of the use of illegal drugs.	Depictions of gambling; use of alcohol, legal drugs, or firearms.	Content containing slurs.	A depiction of someone intentionally cutting into their skin.	Profanity such as "fuck" and "shit".	A person accidentally injuring themselves where the content does not depict the resulting bodily harm.
Meta (Instagram, Facebook, Threads)	Has some restrictions that differ for users under 18. Removes real and AI-generated content displaying nudity and sexual activity. Makes "careful allowances for real world art and certain medical, educational, and awareness-raising content".	Does not allow sexually suggestive content or fetish content. Exceptions for content that is educational, art work, depicts gender affirming surgery, and others, for users 18 and up with a content warning. Exceptions for users 18+ for text based discussions of sexual content.	Does not allow videos of people, living or deceased, in non-medical contexts; Sadistic remarks. Most other graphic content allowed for users 18+ with warning label.	May remove content if requested based on local laws if against community standards. Does not allow high-risk drugs (drugs that have a high potential for misuse, addiction, or are associated with serious health risks) outside discussions of recovery.	Limits content to users 18+ and 21+ depending on activity depicted. Bans the selling of any of these items.	Works with tier based structure: T1: Universal protections for all users; T2: Additional protections for all Minors, Private Adults and Limited Scope Public Figures; T3: Additional protections for Private Minors, Private Adults, and Minor Involuntary Public Figures; T4: Additional protections for Private Minors only	Does not allow content that promotes or shows graphic depictions. Some content restricted to 18+, some also provides resources, some also provide warning screen indicating sensitive nature.	Only removed if used to incite hate toward a person or group of people.	Yes, offers sensitive content warning.
TikTok	Does not allow bare nudity for anyone. Does not allow semi-nudity of young people. Semi-nudity of adults limited to users 18+ and ineligible for FYF.	Does not allow sexually explicit language by anyone. Does not allow content by young people that intends to be sexually suggestive. Content 18+ and ineligible for the FYF if it shows intimate kissing, sexualized framing, or sexualized behavior by adults, or if it shows sex products.	Does not allow gory, gruesome, disturbing, or extremely violent content. Content is 18+, and ineligible for the FYF if it shows human or animal blood, extreme physical fighting, or graphic footage of events that would otherwise violate rules but are in the public interest to view. Content is also ineligible for the FYF if it shows fictional graphic violence or potentially distressing or mildly graphic material. Bans animal abuse, cruelty, neglect, trade, or other forms of animal exploitation.	Bans any violent threats, promotion of violence, incitement to violence, or promotion of criminal activities that may harm people, animals, or property; Human trafficking and smuggling; Sexual and physical abuse.	Bans the facilitation or marketing of gambling or gambling-like activities; Content 18+ and FYF ineligible if glamorizes gambling or gambling-like activities; Bans the trade of alcohol, tobacco products, and or drugs; showing, possessing, or using drugs; Bans showing young people possessing or using alcohol, tobacco products, or drugs; Content 18+ and FYF ineligible if discusses drugs or other regulated substances; Bans the trade or marketing of firearms or explosive weapons, or content showing or promoting them if they are not used in a safe or appropriate setting; Bans selling/trading regulated goods/services.	Bans hate speech, hateful behavior, or promotion of hateful ideologies — implicit or explicit. Content ineligible for FYF if indirectly demeaning protected groups in the context of discussing social issues. Does not allow violent or hateful orgs/people, or inciting violence/hate onto others. Bans harassing, degrading, or bullying statements or behavior.	Bans showing, promoting, or sharing plans for suicide or self-harm. Bans showing or promoting disordered eating and dangerous weight loss behaviors, or facilitating the trade or marketing of weight loss or muscle gain products. Content 18+ and FYF ineligible if it shows or promotes potentially harmful weight management, or markets weight loss or muscle gain products (+ in some regions, content that shows cosmetic surgery and doesn't include risk warnings).	Does not specify.	Yes, offers sensitive content warning.
Youtube	Bans explicit content meant to be sexually gratifying. Fetish content often removed or age restricted. Violent, graphic, or humiliating fetish content not allowed. Includes real-world, dramatized, illustrated, and animated content, including sex scenes, video games, and music.	Appears to adhere to same guidelines as above (not specified).	Does not allow violent or gory content intended to shock or disgust viewers, or content encouraging others to commit violent acts. Includes animal abuse.	Bans content intended to praise, promote, or aid violent extremist or criminal organizations. Bans selling of illegal or regulated goods/services. Bans hard drug use or creation, poison sale or creation, and instructions for academic cheating. Bans additional illegal activity.	Age restricted content (does not specify what age) Content that promotes a cannabis dispensary; reviews brands of nicotine e-liquid; facilitates access to, promotes, or depicts online gambling, including content from approved sites.	Bans hate speech against protected groups. Bans content that targets someone with prolonged insults or slurs based on their physical traits or protected group status. Bans harmful behaviors, like threats or doxxing. Stricter approach on content that targets minors. May include content that includes harassment if the primary purpose is educational, documentary, scientific, or artistic in nature (including comedy performances).	Bans promoting suicide, self-harm, or eating disorders, that is intended to shock or disgust, or that poses a considerable risk to viewers. Allows discussing these topics in a non harmful way.	Use of explicit language may result in age restriction, content removal, or strike (probation).	Yes, offers sensitive content warning.
X	Allowed as long as all parties consent, content is properly labeled, and "not prominently displayed".	Same as above (does not specify for suggestive implicit content).	Some allowed. Allowed if properly labeled, not prominently displayed and is not excessively gory or depicting sexual violence. Explicitly threatening, inciting, glorifying, or expressing desire for violence is not allowed.	Not allowed to use platform to further or encourage illegal activities. Includes selling, buying, or facilitating transactions in illegal goods or services, as well as certain types of regulated goods or services.	Does not specify.	Not allowed. May not share abusive content, engage in the targeted harassment of someone, or incite other people to do so. Prohibits abuse or attacks on identity.	Bans promoting or encouraging self-harm, suicide, or eating disorders. Allows discussing topics in a non harmful or educational way.	Typically allowed as long as it is not used in the context of hate speech.	Yes, offers sensitive content warning.
Reddit	Allowed and dependent on community. Only banned if depicting minors.	Allowed and dependent on community.	Allowed and dependent on community.	"Avoid posting illegal content or soliciting or facilitating illegal or prohibited transactions."	Allowed and dependent on community.	Bans communities and users that incite violence or that promote hate based on identity or vulnerability. Specifically bans doxxing and impersonation.	Allowed and dependent on community.	Allowed and dependent on community.	Yes, offers sensitive content warning and users can tag for NSFW.
4Chan	Allowed and dependent on board.	Allowed and dependent on board.	Allowed and dependent on board.	"Don't post, share, request, or link to anything that breaks U.S. or local laws."	Platform is strictly 18+. Allowed and dependent on board.	Don't share or ask for personal information (doxxing), and don't organize or take part in group spam or raids. Starting fights between boards is also against the rules.	Allowed and dependent on board.	Allowed and dependent on board.	Yes — users must specify whether posts are NSFW or risk ban.
Tumblr	Visual depictions of sexually explicit acts (or content with an overt focus on genitalia) not allowed. Other nudity allowed with content label (such as "Sexual Themes").	Allowed.	Does not allow violent or gory content intended to shock or disgust viewers. Does not allow content that encourages or incites violence, or glorifies acts of violence or the perpetrators.	Bans conducting illegal behavior, like fraud or phishing; Specifically bans using the platform for election interference and facilitation of human trafficking.	Does not specify.	"Don't engage in targeted abuse, bullying, or harassment. Don't engage in the unwanted sexualization or sexual harassment of others." Violence and hate speech banned.	Bans promoting or encouraging self-harm, suicide, or eating disorders. Allows discussing topics in a non harmful or educational way.	Allowed.	Yes, offers sensitive content warning and users can add a content label.

Appendix K

R Code

Note that “ “ is used in place of " " throughout the code due to latex rendering issues. The actual R code uses " " for labels.

 # Load in Packages

```
require(pacman)
pacman::p_load(haven, gt, gtable, gtsummary, viridis,
   corrplot, psych, entropy, nFactors, effects,
   interactions, lme4, lmerTest, pompom, mlVAR, patchwork,
   ggeffects, lubridate, tidyverse)
```

 # Data Import

```
#(excluded for participant privacy)
```

 # Data Cleaning

```
#(excluded for conciseness)
```

 # Data Filtering

```
# filter for participants with screenshot and survey data
stick_perpid <- verified_participants %>%
  filter(sushi==1 & eventlog==1 & screenshot_stats==1) %>%
  select(email) %>%
  left_join(stickiness_pidfn, by=``email``)
length(unique(stick_perpid$email))
stick_pid_unique <- stick_perpid %>%
  distinct(email)


# filter survey data to only include above participants
```

```
SUSHI_PL <- verified_participants %>%
  filter(sushi==1 & eventlog==1 & screenshot_stats==1) %>%
  select(pid) %>%
  left_join(SUSHI_PL, by=``pid``)
SUSHI_FN <- verified_participants %>%
  filter(sushi==1 & eventlog==1 & screenshot_stats==1) %>%
  select(pid) %>%
  left_join(SUSHI_FN, by=``pid``)

# remove the ``onboarding`` survey source
SUSHI_FN <- SUSHI_FN %>%
  filter(datasource != ``onboarding``)

# select for variables of interest
SUSHI_FN <- SUSHI_FN %>%
  select(pid, w5, email, fortnight, depressed, lonely, sad,
      anxious, worried, rumination, stressed, happy, calm,
     satisfied, relieved, rumination, attention,
     concentration, restless, impatient, spur_of_moment,
     cesd_total)

# select for variables of interest
SUSHI_PL <- SUSHI_PL %>%
  select(pid, w5, email, screenshotcount, survey_total,
     survey_filtered, age_mode, gender_mode, white_mode,
     nativeam_mode, asian_mode, black_mode, pacificis_mode,
      other_mode, prefer_not_answer_mode, hispanic_mode,
```

```
    education_mode, income_mode, political_affiliation_
    mode, )

# describe the datasets
describe(SUSHI_FN)
str(SUSHI_FN)
length(unique(SUSHI_FN$pid))
describe(SUSHI_PL)
str(SUSHI_PL)
length(unique(SUSHI_PL$email))
```

 # Get Unique App Ratings

```
apps_unique_ratings <- apps %>%
  distinct() %>%
  arrange(app)
# use max and min filtering to keep ``yes`` ratings
apps_unique_ratings <- apps_unique_ratings %>%
  ungroup() %>%
  mutate(ga_num = ifelse(graphicaccess == ``Yes``, 1, 0),
    gp_num = ifelse(graphicpurpose == ``Yes``, 1, 0))
apps_unique_ratings_f <- apps_unique_ratings %>%
  group_by(app) %>%
  filter(gp_num == max(gp_num)) %>%
  filter(ga_num == max(ga_num))
# merge onto df
apps_merged_stickiness <- stickiness_pidfnapp %>%
  right_join(apps_unique_ratings_f, by = ``app``, keep=F)
```

 # Stickiness Variable

```
# look at summary statistics
describe(stick_perpid)
summary(stick_perpid$stickiness)

# log transform stickiness due to distribution skew
stick_perpid <- stick_perpid %>%
  filter(!is.na(fortnight),
         !is.na(stickiness)) %>%
  mutate(stickiness_log = log1p(as.numeric(stickiness)))
describe(stick_perpid)
summary(stick_perpid$stickiness_log)

# add to SUSHI_FN for later use in creating state and trait
    variables
SUSHI_FN <- left_join(SUSHI_FN, stick_perpid, by = c(``
   email``, ``fortnight``))
```

 # Create Trait and State Variables

```
# calculate summary cognitive-affective scales
scales <- SUSHI_FN %>%
  group_by(pid, fortnight) %>%
  summarize(total_negcog = sum(c(rumination, attention,
     concentration, restless, impatient, spur_of_moment),
     na.rm=TRUE), total_negaff = sum(c(depressed, lonely,
     sad, anxious, worried, stressed), na.rm=TRUE), total_
     stickiness = sum(stickiness_log), na.rm=TRUE)
```

```
# calculate trait variables
trait_scales <- SUSHI_FN %>%
  select(pid, fortnight, cesd_total) %>%
  right_join(scales, by =c(``pid``, ``fortnight``)) %>%
  group_by(pid) %>%
  summarize(trait_negcog = mean(total_negcog, na.rm=TRUE),
     trait_negaff = mean(total_negaff, na.rm=TRUE), trait_
     cesdtotal = mean(cesd_total, na.rm=TRUE), trait_
     stickiness = mean(total_stickiness, na.rm=TRUE))

# merge trait_scales to scales
scales <- SUSHI_FN %>%
  select(pid, fortnight, cesd_total) %>%
  right_join(scales, by =c(``pid``, ``fortnight``)) %>%
  left_join(trait_scales, by=``pid``)

# merge trait variables and calculate state variables
scales <- scales %>%
  mutate(state_negcog = total_negcog - trait_negcog, state_
     negaff = total_negaff - trait_negaff, state_cesdtotal
     = cesd_total - trait_cesdtotal)

# look at structure and stats
str(scales)
describe(scales)
```

```
# calculate IQR for every numeric column
sapply(scales[sapply(scales, is.numeric)], IQR, na.rm =
   TRUE)

# merge in stickiness data
scales <- SUSHI_PL %>%
  select(pid, email) %>%
  right_join(stick_perpid, by=``email``) %>%
  left_join(scales, by=c(``pid``, ``fortnight``))

# remove rows that have any NA values in any column
scales <- scales %>%
  drop_na()
```

 # Compute Correlations

```
cor_vars <- scales %>%
  select(fortnight, stickiness_log, total_negcog, trait_
     negcog, state_negcog, total_negaff, trait_negaff,
     state_negaff, cesd_total, trait_cesdtotal, state_
     cesdtotal)
# correlate all variables, exclude pid, email, redundant
   stickiness vars, screenshot count, survey total, and
   cesd binary
round(cor(cor_vars, use=``complete.obs``), 2)
```

 # Center and Scale Variables

```
# scale and center all variables (z score)
sc_vars <- scales %>%
```

```
  mutate(across(c(stickiness_log, trait_stickiness, cesd_
     total, trait_cesdtotal, state_cesdtotal,trait_negcog,
     state_negcog, trait_negaff, state_negaff, total_negcog
     , total_negaff), ~ as.numeric(scale(.)), .names = ``{.
     col}_z``))
describe(sc_vars)

# calculate IQR for every numeric column
sapply(sc_vars[sapply(sc_vars, is.numeric)], IQR, na.rm =
   TRUE)

# add scales data frame to stickiness and app
sc_vars <- left_join(sc_vars, apps_merged_stickiness, by =
   c(``email``, ``fortnight``))
describe(sc_vars)

# correlations of sc variables
sc_vars_num <- sc_vars %>% select(where(is.numeric))
round(cor(sc_vars_num, use = ``complete.obs``), 2)
```

Plot Variable Distributions

```
# example code for cesd shown below, done for all four
   variables
var_colors <- c(
  ``cesd_total`` = ``mediumpurple4``,
  ``cesd_total_z`` = ``hotpink4``,
  ``trait_cesdtotal_z`` = ``darkolivegreen``,
```

```
  ``state_cesdtotal_z`` = ``steelblue4``
)

p1 <- ggplot(sc_vars, aes(x = cesd_total, fill = ``cesd_
   total``)) +
  geom_histogram(color = ``white``, alpha = 0.8) +
  scale_fill_manual(values = var_colors) +
  theme_minimal(base_size = 11) +
  theme(legend.position = ``none``) +
  labs(title = ``CES-D Total Distribution``, x = ``CES-D
     Total``, y = ``Count``)

p2 <- ggplot(sc_vars, aes(x = cesd_total_z, fill = ``cesd_
   total_z``)) +
  geom_histogram(color = ``white``, alpha = 0.8) +
  scale_fill_manual(values = var_colors) +
  theme_minimal(base_size = 11) +
  theme(legend.position = ``none``) +
  labs(title = ``CES-D Total \n (centered and scaled)
     Distribution``,
       x = ``Centered and Scaled Negative Affect``, y = ``
         Count``)

p3 <- ggplot(sc_vars, aes(x = trait_cesdtotal_z, fill = ``
   trait_cesdtotal_z``)) +
  geom_histogram(color = ``white``, alpha = 0.8) +
  scale_fill_manual(values = var_colors) +
```

```
  theme_minimal(base_size = 11) +
  theme(legend.position = ``none``) +
  labs(title = ``Trait CES-D Total Distribution``, x = ``
     Trait CES-D Total``, y = ``Count``)


p4 <- ggplot(sc_vars, aes(x = state_cesdtotal_z, fill = ``
   state_cesdtotal_z``)) +
  geom_histogram(color = ``white``, alpha = 0.8) +
  scale_fill_manual(values = var_colors) +
  theme_minimal(base_size = 11) +
  theme(legend.position = ``none``) +
  labs(title = ``State CES-D Total Distribution``, x = ``
     State CES-D Total``, y = ``Count``)

(p1 | p2) / (p3 | p4)
```

 # Analyze Participant Demographics

```
table(SUSHI_PL$gender_mode)
mean(SUSHI_PL$age_mode, na.rm=TRUE)
table(SUSHI_PL$white_mode)
table(SUSHI_PL$black_mode)
table(SUSHI_PL$asian_mode)
table(SUSHI_PL$hispanic_mode)
table(SUSHI_PL$nativeam_mode)
table(SUSHI_PL$other_mode)
table(SUSHI_PL$prefer_not_answer_mode)
```

```
table(SUSHI_PL$education_mode)
table(SUSHI_PL$income_mode)
table(SUSHI_PL$political_affiliation_mode)

# 45 people selected they were hispanic
# 138 selected they were white
SUSHI_PL %>%
  summarise(n = sum(white_mode == 1 & hispanic_mode == 1,
     na.rm = TRUE))
# 25 selected white and hispanic
SUSHI_PL %>%
  summarise(n = sum(is.na(white_mode) & hispanic_mode == 1,
      na.rm = TRUE))
# 20 selected hispanic and left white blank
SUSHI_PL %>%
  summarise(n = sum(white_mode == 1 & hispanic_mode == 2,
     na.rm = TRUE))
# 113 selected white and nonhispanic

# look for multiracial participants
SUSHI_PL %>%
  filter(rowSums(select(., white_mode, black_mode, asian_
     mode, nativeam_mode, pacificis_mode), na.rm = TRUE) >
     1) %>%
  select(pid, white_mode, black_mode, asian_mode, nativeam_
     mode, pacificis_mode, hispanic_mode)
```

 # Calculate App Usage Summaries

```
# summarize graphic app usage per participant per fortnight
graphic_summary <- sc_vars %>%
  group_by(pid, fortnight) %>%
  summarize(prop_graphic = mean(graphicpurpose == ``Yes``,
     na.rm = TRUE), count_graphic = sum(graphicpurpose ==
     ``Yes``, na.rm = TRUE))

# merge back into sc_vars
sc_vars <- sc_vars %>%
  left_join(graphic_summary, by = c(``pid``, ``fortnight``)
     )

# scale graphic purpose app use
sc_vars$prop_graphic_z <- scale(sc_vars$prop_graphic)
graphic_access_summary <- sc_vars %>%
  group_by(pid, fortnight) %>%
  summarize(prop_graphic_access = mean(graphicaccess == ``
     Yes``, na.rm = TRUE), count_graphic_access = sum(
     graphicaccess == ``Yes``, na.rm = TRUE))

# merge back into sc_vars
sc_vars <- sc_vars %>%
  left_join(graphic_access_summary, by = c(``pid``, ``
     fortnight``))

# scale graphic purpose app use
sc_vars$prop_graphic_access_z <- scale(sc_vars$prop_graphic
```

```
   _access)
```

 # Calculate Intraclass Correlation Coefficients

```
# stickiness total
stickiness0_fit <- lmer(stickiness_log_z ~ 1 + (1 | pid),
   data = sc_vars)
summary(stickiness0_fit)
sticky_components <- as.data.frame(VarCorr(stickiness0_fit)
   )
sticky_btwn <- sticky_components$vcov[1]  # random
   intercept variance
sticky_residual <- sticky_components$vcov[2] # residual
   variance
sticky_ICC <- sticky_btwn / (sticky_btwn + sticky_residual)
sticky_ICC

# ces-d total
cesd0_fit <- lmer(cesd_total ~ 1 + (1 | pid), data = sc_
   vars)
summary(cesd0_fit)

# negative affect total
negaff0_fit <- lmer(total_negaff ~ 1 + (1 | pid), data = sc
   _vars)
summary(negaff0_fit)

# negative cognition total
```

```
negcog0_fit <- lmer(total_negcog ~ 1 + (1 | pid), data = sc
   _vars)
summary(negcog0_fit)
```

Stickiness as a Function of Negative Affect and Fortnight

```
# STICKINESS AS FUNCTION OF NEGATIVE AFFECT AND FORTNIGHT
sticky1_fit_negaff <- lmer(formula = stickiness_log_z ~ 1 +
    fortnight + trait_negaff_z + state_negaff_z + trait_
   negaff_z*state_negaff_z +
(1|pid), data=sc_vars,na.action=na.exclude)
summary(sticky1_fit_negaff)

sc_vars <- sc_vars %>%
  mutate(pred_sticky_negaff = predict(sticky1_fit_negaff))

confint(sticky1_fit_negaff)
```

Check Model Skew

```
# CHECK MODEL SKEW
# check original vs log transformed
ggplot(sc_vars, aes(x = stickiness_log_z)) + geom_histogram
   (bins = 50)
# check residuals
qqnorm(resid(sticky1_fit_negaff))
qqline(resid(sticky1_fit_negaff))
```

Sensitivity Analysis

```
sticky_robust <- rlmer(stickiness_log_z ~ 1 + fortnight +
```

```
    trait_negaff_z + state_negaff_z + trait_negaff_z * state
    _negaff_z + (1 | pid), data = sc_vars)
summary(sticky_robust)
```

 # Stickiness as a Function of Negative Cognition and Fortnight

```
sticky1_fit_negcog <- lmer(formula = stickiness_log_z ~ 1 +
     fortnight + trait_negcog_z +
state_negcog_z + trait_negcog_z*state_negcog_z +
(1|pid), data=sc_vars, na.action=na.exclude)
summary(sticky1_fit_negcog)


sc_vars <- sc_vars %>%
  mutate(pred_sticky_negcog = predict(sticky1_fit_negcog))
confint(sticky1_fit_negcog)
```

 # Stickiness as a Function of CES-D Total and Fortnight

```
# STICKINESS AS FUNCTION OF CES-D TOTAL AND FORTNIGHT
sticky1_fit_cesd <- lmer(formula = stickiness_log_z ~ 1 +
   fortnight + trait_cesdtotal + state_cesdtotal + trait_
   cesdtotal*state_cesdtotal + (1|pid), data=sc_vars, na.
   action=na.exclude)
summary(sticky1_fit_cesd)


sc_vars <- sc_vars %>%
  mutate(pred_sticky_cesd = predict(sticky1_fit_cesd))
confint(sticky1_fit_cesd)
```

 # Stickiness as a Function of Fortnight, Negative Affect, Graphic Purpose Apps, and Graphic Access Apps

```
sticky_negaff_gracont <- lmer(formula = stickiness_log_z ~
   1 + fortnight + graphicpurpose + graphicaccess + state_
   negaff_z + state_negaff_z*graphicpurpose + state_negaff_
   z*graphicaccess + trait_negaff_z + trait_negaff_z*
   graphicpurpose + trait_negaff_z*graphicaccess + trait_
   negaff_z*state_negaff_z + (1|pid), data=sc_vars, na.
   action=na.exclude)

summary(sticky_negaff_gracont)
confint(sticky_negaff_gracont)

sc_vars <- sc_vars %>%
  mutate(pred_sticky_negaff_gracont = predict(sticky_negaff
     _gracont))

# PLOT OUTCOMES
# graphic access on x axis, state CES-D as -1/+1 SD lines
state_negaff_gracont <- ggpredict(sticky_negaff_gracont,
                terms = c(``graphicaccess``, ``state_negaff
                   _z [-1, 1]``)) %>%
  as.data.frame() %>%
  ggplot(aes(x = x, y = predicted, color = group, fill =
     group, shape = group)) +
  geom_line(aes(group = group)) +
  geom_point(size = 3) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high, group
     = group), alpha = 0.10, color = NA) +
```

```
  labs(x = ``Graphic Access App``,
       y = ``Predicted Stickiness (log z)``,
       color = ``State Negative Affect``,
       fill = ``State Negative Affect``,
       shape = ``State Negative Affect``,
       title = ``Graphic Access x State Negative Affect (-1
          /+1 SD)``) +
  scale_color_manual(values = c(``darkorchid4``, ``
     darkolivegreen``),
                    labels = c(``-1 SD``, ``+1 SD``)) +
  scale_fill_manual(values = c(``darkorchid4``, ``
     darkolivegreen``),
                   labels = c(``-1 SD``, ``+1 SD``)) +
  scale_shape_manual(values = c(16, 17), # creates filled
     circle and triangle
                    labels = c(``-1 SD``, ``+1 SD``)) +
  coord_cartesian(ylim = c(-1, 1)) +
  theme_minimal()

# graphic access on x axis, trait CES-D as -1/+1 SD lines
trait_negaff_gracont <- ggpredict(sticky_negaff_gracont,
                terms = c(``graphicaccess``, ``trait_negaff
                   _z [-1, 1]``)) %>%
  as.data.frame() %>%
  ggplot(aes(x = x, y = predicted, color = group, fill =
     group, shape = group)) +
  geom_line(aes(group = group)) +
```

```
  geom_point(size = 3) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high, group
     = group), alpha = 0.10, color = NA) +
  labs(x = ``Graphic Access App``,
       y = ``Predicted Stickiness (log z)``,
       color = ``Trait Negative Affect``,
       fill = ``Trait Negative Affect``,
       shape = ``Trait Negative Affect``,
       title = ``Graphic Access x Trait Negative Affect (-1
         /+1 SD)``) +
  scale_color_manual(values = c(``darkorchid4``, ``
     darkolivegreen``),
                     labels = c(``-1 SD``, ``+1 SD``)) +
  scale_fill_manual(values = c(``darkorchid4``, ``
     darkolivegreen``),
                    labels = c(``-1 SD``, ``+1 SD``)) +
  scale_shape_manual(values = c(16, 17),
                     labels = c(``-1 SD``, ``+1 SD``)) +
  coord_cartesian(ylim = c(-1, 1)) +
  theme_minimal()

state_negaff_gracont + trait_negaff_gracont
```

 # Stickiness as a Function of Fortnight, Negative Cognition, Graphic Purpose Apps, and Graphic Access Apps

```
sticky_negcog_gracont <- lmer(formula = stickiness_log_z ~
   1 + fortnight + graphicpurpose + graphicaccess + state_
```

```
    negcog_z + state_negcog_z*graphicpurpose + state_negcog_
    z*graphicaccess + trait_negcog_z + trait_negcog_z*
    graphicpurpose + trait_negcog_z*graphicaccess + trait_
    negcog_z*state_negcog_z + (1|pid), data=sc_vars, na.
    action=na.exclude)
summary(sticky_negcog_gracont)
confint(sticky_negcog_gracont)

sc_vars <- sc_vars %>%
  mutate(pred_sticky_negcog_gracont = predict(sticky_negcog
     _gracont))

# PLOT OUTCOMES
# graphic access on x axis, state CES-D as -1/+1 SD lines
state_negcog_gracont <- ggpredict(sticky_negcog, gracont,
        terms = c(``graphicaccess``, ``state_negcog_z [-1,
           1]``)) %>%
  as.data.frame() %>%
  ggplot(aes(x =x, y=predicted, color=group, fill=group,
     shape=group)) +
  geom_line(aes(group=group)) +
  geom_point(size=3) +
  geom_ribbon(aes(ymin=conf.low, y.max=conf.high, group=
     group), alpha=0.10, color=NA) +
  labs(x=``Graphic Access App``,
        y = ``Predicted Stickiness (log z)``,
        color = ``State Negative Cognition``,
```

```
        fill = ``State Negative Cognition``,
        shape = ``State Negative Cognition``,
        title = ``Graphic Access Apps x State Negative
           Cognition (-1/+1 SD)``) +
  scale_color_manual(values=c(``darkorchid4``, ``
     darkolivegreen``),
                 labels = c(``-1 SD``, ``+1 SD``)) +
  scale_fill_manual(values=c(``darkorchid4``, ``
     darkolivegreen``),
                 labels = c(``-1 SD``, ``+1 SD``)) +
  scale_shape_manual(values = c(16,17),
                 labels = c(``-1 SD``, ``+1 SD``)) +
  coord_cartesian(ylim=c(-1, 1)) +
  theme_minimal()

#graphic access on x axis, trait CES-D as -1/+1 SD lines
trait_negcog_gracont <- ggpredict(sticky_negcog, gracont,
        terms = c(``graphicaccess``, ``trait_negcog_z [-1,
           1]``)) %>%
  as.data.frame() %>%
  ggplot(aes(x =x, y=predicted, color=group, fill=group,
     shape=group)) +
  geom_line(aes(group=group)) +
  geom_point(size=3) +
  geom_ribbon(aes(ymin=conf.low, y.max=conf.high, group=
     group), alpha=0.10, color=NA) +
  labs(x=``Graphic Access App``,
```

```
        y = ``Predicted Stickiness (log z)``,
        color = ``Trait Negative Cognition``,
        fill = ``Trait Negative Cognition``,
        shape = ``Trait Negative Cognition``,
        title = ``Graphic Access Apps x Trait Negative
          Cognition (-1/+1 SD)``) +
  scale_color_manual(values=c(``darkorchid4``, ``
    darkolivegreen``),
              labels = c(``-1 SD``, ``+1 SD``)) +
  scale_fill_manual(values=c(``darkorchid4``, ``
    darkolivegreen``),
              labels = c(``-1 SD``, ``+1 SD``)) +
  scale_shape_manual(values = c(16,17),
              labels = c(``-1 SD``, ``+1 SD``)) +
  coord_cartesian(ylim=c(-1, 1)) +
  theme_minimal()

state_negcog_gracont + trait_negcog_gracont
```

Stickiness as a Function of Fortnight, CES-D Total, Graphic Purpose Apps, and Graphic Access Apps

```
# STICKINESS AS A FUNCTION OF FORTNIGHT, CES-D TOTAL,
   GRAPHIC PURPOSE APPS, AND GRAPHIC ACCESS APPS (EQUATION
   AND PLOTS)
sticky_cesd_gracont <- lmer(formula = stickiness_log_z ~ 1
   + fortnight + graphicpurpose + graphicaccess + state_
   cesdtotal_z + state_cesdtotal_z*graphicpurpose + state_
```

```
    cesdtotal_z*graphicaccess + trait_cesdtotal_z + trait_
    cesdtotal_z*graphicpurpose + trait_cesdtotal_z*
    graphicaccess + trait_cesdtotal_z*state_cesdtotal_z +
    (1|pid), data=sc_vars, na.action=na.exclude)
summary(sticky_cesd_gracont)
confint(sticky_cesd_gracont)

# save for plotting
sc_vars <- sc_vars %>%
  mutate(pred_sticky_cesd_gracont = predict(sticky_cesd_
     gracont))

# graphic access on x axis, state CES-D as -1/+1 SD lines
state_cesd_gracont <- ggpredict(sticky_cesd_gracont,
                   terms = c(``graphicaccess``, ``state_
                      cesdtotal_z [-1, 1]``)) %>%
  as.data.frame() %>%
  ggplot(aes(x = x, y = predicted, color = group, fill =
     group, shape = group)) +
  geom_line(aes(group = group)) +
  geom_point(size = 3) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high, group
     = group), alpha = 0.10, color = NA) +
  labs(x = ``Graphic Access App``,
       y = ``Predicted Stickiness (log z)``,
       color = ``State CES-D``,
       fill = ``State CES-D``,
```

```
        shape = ``State CES-D``,
        title = ``Graphic Access x State CES-D (-1/+1 SD)``)
            +
    scale_color_manual(values = c(``darkorchid4``, ``
      darkolivegreen``),
                        labels = c(``-1 SD``, ``+1 SD``)) +
    scale_fill_manual(values = c(``darkorchid4``, ``
      darkolivegreen``),
                       labels = c(``-1 SD``, ``+1 SD``)) +
    scale_shape_manual(values = c(16, 17), # creates filled
      circle and triangle
                        labels = c(``-1 SD``, ``+1 SD``)) +
    coord_cartesian(ylim = c(-1, 1)) +
    theme_minimal()

  # graphic access on x axis, trait CES-D as -1/+1 SD lines
  trait_cesd_gracont <- ggpredict(sticky_cesd_gracont,
                terms = c(``graphicaccess``, ``trait_
                  cesdtotal_z [-1, 1]``)) %>%
    as.data.frame() %>%
    ggplot(aes(x = x, y = predicted, color = group, fill =
      group, shape = group)) +
    geom_line(aes(group = group)) +
    geom_point(size = 3) +
    geom_ribbon(aes(ymin = conf.low, ymax = conf.high, group
      = group), alpha = 0.10, color = NA) +
    labs(x = ``Graphic Access App``,
```

```
        y = ``Predicted Stickiness (log z)``,
        color = ``Trait CES-D``,
        fill = ``Trait CES-D``,
        shape = ``Trait CES-D``,
        title = ``Graphic Access x Trait CES-D (-1/+1 SD)``)
            +
  scale_color_manual(values = c(``darkorchid4``, ``
    darkolivegreen``),
                     labels = c(``-1 SD``, ``+1 SD``)) +
  scale_fill_manual(values = c(``darkorchid4``, ``
    darkolivegreen``),
                    labels = c(``-1 SD``, ``+1 SD``)) +
  scale_shape_manual(values = c(16, 17),
                     labels = c(``-1 SD``, ``+1 SD``)) +
  coord_cartesian(ylim = c(-1, 1)) +
  theme_minimal()

state_cesd_gracont + trait_cesd_gracont
```

Quantitative Analysis of Coded Screenshots (PID 175)

```
# proportion of social media apps
pid175_annotated %>%
  count(is_social_app) %>%
  mutate(proportion = n / sum(n))
# 29799 false (93%) 2210 true (6.9%)


# proportion of moderated content on social media apps
```

```
pid175_annotated %>%
  filter(is_social_app == ``True``) %>%
  count(is_moderated) %>%
  mutate(proportion = n / sum(n))
# 2033 false (92%) 177 true (8%)

# of this moderated content, what is the content type
pid175_annotated %>%
  filter(is_social_app == ``True``, is_moderated == ``True
     ``) %>%
  summarise(across(c(has_sexual_content, has_graphic_
     violence, has_illegal_activity, has_harassment, has_
     self_injury, has_vulgar_language, has_sensitive_
     content),
                 list(
                      count = ~ sum(. == ``True``, na.rm =
                         TRUE),
                      proportion = ~ mean(. == ``True``, na.
                         rm = TRUE)
                 ))) %>%
  pivot_longer(everything(),
               names_to = c(``category``, ``.value``),
               names_sep = ``_(?=[^_]+$)``)
```

 # Quantitative Analysis of Coded Screenshots (PID 237)

```
# proportion of social media apps
pid237_annotated %>%
```

```
  count(is_social_app) %>%
  mutate(proportion = n / sum(n))
# 6456 false (77.9%) 1835 (22.1%)

# proportion of moderated content on social media apps
pid237_annotated %>%
  filter(is_social_app == ``True``) %>%
  count(is_moderated) %>%
  mutate(proportion = n / sum(n))
# 1679 false (91.5%) 177 true (8.5%)

# of this moderated content, what is the content type
pid237_annotated %>%
  filter(is_social_app == ``True``, is_moderated == ``True
     ``) %>%
  summarise(across(c(has_sexual_content, has_graphic_
     violence, has_illegal_activity, has_harassment, has_
     self_injury, has_vulgar_language, has_sensitive_
     content),
                 list(
                     count = ~ sum(. == ``True``, na.rm =
                        TRUE),
                     proportion = ~ mean(. == ``True``, na.
                        rm = TRUE)
                   ))) %>%
  pivot_longer(everything(),
               names_to = c(``category``, ``.value``),
```

```
                names_sep = ``_(?=[^_]+$)``)
```

 # Vignette Tagging Plot 175 (Content Moderation Categories, Valence, Arousal)

```
# code adapted for both case study participants
# flag colors and labels
flag_colors <- c(
  ``has_sexual_content`` = ``red4``,
  ``has_graphic_violence`` = ``darkolivegreen``,
  ``has_illegal_activity`` = ``mediumpurple4``,
  ``has_harassment`` = ``peachpuff2``,
  ``has_self_injury`` = ``palevioletred``,
  ``has_vulgar_language`` = ``steelblue4``,
  ``has_sensitive_content`` = ``darkorange3``
)

flag_labels <- c(
  ``has_sexual_content`` = ``Sexual Content``,
  ``has_graphic_violence`` = ``Graphic Violence``,
  ``has_illegal_activity`` = ``Illegal Activity``,
  ``has_harassment`` = ``Harassment``,
  ``has_self_injury`` = ``Self Injury``,
  ``has_vulgar_language`` = ``Vulgar Language``,
  ``has_sensitive_content`` = ``Sensitive Content``
)

day_labels <- c(
  ``2020-12-28`` = ``Day 1``,
```

```
  ``2020-12-29`` = ``Day 2``,
  ``2020-12-30`` = ``Day 3``,
  ``2020-12-31`` = ``Day 4``
)

# prep flag data
pid175_plot <- pid175_annotated %>%
  mutate(date = as.character(as.Date(datetime)),
         day_label = factor(day_labels[date], levels = c(``
           Day 1``, ``Day 2``, ``Day 3``, ``Day 4``))) %>%
  filter(is_social_app == ``True``,
         date >= ``2020-12-28``,
         date <= ``2020-12-31``) %>%
  pivot_longer(cols = c(has_sexual_content, has_graphic_
     violence, has_illegal_activity, has_harassment, has_
     self_injury, has_vulgar_language, has_sensitive_
     content),
  names_to = ``category``, values_to = ``flagged``) %>%
  filter(flagged == ``True``)

# count flags per day per category
flag_counts_175 <- pid175_plot %>%
  group_by(day_label, category) %>%
  summarise(count = n(), .groups = ``drop``)

# valence and arousal summary
va_summary <- pid175_annotated %>%
```

```
  mutate(date = as.character(as.Date(datetime)),
         day_label = factor(day_labels[date], levels = c(``
            Day 1``, ``Day 2``, ``Day 3``, ``Day 4``))) %>%
  filter(is_social_app == ``True``,
         date >= ``2020-12-28``,
         date <= ``2020-12-31``) %>%
  group_by(day_label) %>%
  summarise(valence = mean(valence, na.rm = TRUE),
            arousal = mean(arousal, na.rm = TRUE))

# radial bar chart
radial_175 <- ggplot() +
  geom_bar(
    data = flag_counts_175,
    aes(x = day_label, y = count, fill = category),
    stat = ``identity``, position = ``stack``, width = 0.9
  ) +
  coord_polar(theta = ``x``, start = 0) +
  scale_fill_manual(values = flag_colors, labels = flag_
     labels, name = ``Content Flag``) +
  scale_y_continuous(limits = c(-10, NA)) +
  theme_minimal(base_size = 11) +
  theme(
    plot.title = element_text(hjust = 0.5, face = ``bold``,
        margin = margin(b = 20)),
    axis.title = element_blank(),
    axis.text.y = element_blank(),
```

```
    axis.text.x = element_text(size = 9, color = ``grey40
      ``),
    panel.grid.major = element_line(color = ``grey90``,
      linewidth = 0.3),
    panel.grid.minor = element_blank(),
    legend.position = ``bottom``,
    legend.title = element_text(face = ``bold``, size = 9),
    legend.text = element_text(size = 8)
  ) +
  labs(title = ``Moderated Social Media Content by Day (
     Participant A)``)

# valence and arousal line plot
va_175 <- va_summary %>%
  pivot_longer(cols = c(valence, arousal), names_to = ``
     measure``, values_to = ``value``) %>%
  ggplot(aes(x = day_label, y = value, color = measure,
     group = measure)) +
  geom_line() +
  geom_point() +
  scale_color_manual(values = c(``valence`` = ``
     darkslateblue``, ``arousal`` = ``violetred``),
  labels = c(``valence`` = ``Valence``, ``arousal`` = ``
     Arousal``)) +
  labs(x = NULL, y = ``Mean Score``, color = NULL) +
  coord_cartesian(ylim = c(3.6, 7.4)) +
  theme_minimal() +
```

```
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1, size
       = 9),
    axis.title.y = element_text(size = 9),
    legend.position = ``bottom``,
    panel.grid.minor = element_blank()
  )

# combine
radial_175 / va_175 + plot_layout(heights = c(2, 1))
```

 # Vignette Plot - Participant B

```
# example code shown below for Participant B; completed one
    for both case studies

app_colors_237 <- c(
  ``Instagram`` = ``slateblue1``,
  ``YouTube`` = ``firebrick2``,
  ``Twitter`` = ``#1DA1F2``,
  ``Tumblr`` = ``#35465C``,
  ``Reddit`` = ``#FF4500``,
  ``Social`` = ``pink``,
  ``Gaming`` = ``plum``,
  ``Browser`` = ``lightblue``,
  ``System`` = ``#AAAAAA``,
  ``Finance`` = ``lightyellow``,
  ``Files`` = ``mistyrose2``,
```

```
  ``Utility`` = ``mistyrose2``,
  ``NA`` = ``#DDDDDD``
)

flag_colors <- c(
  ``has_sexual_content`` = ``#FF1493``,
  ``has_graphic_violence`` = ``#CB0FFF``,
  ``has_illegal_activity`` = ``#2C3E50``,
  ``has_age_restricted`` = ``white``,
  ``has_harassment`` = ``#FB5607``,
  ``has_self_injury`` = ``yellow``,
  ``has_vulgar_language`` = ``#39FF14``,
  ``has_sensitive_content`` = ``#00F5FF``
)

flag_shapes <- c(
  ``has_sexual_content`` = 18,
  ``has_graphic_violence`` = 18,
  ``has_illegal_activity`` = 18,
  ``has_age_restricted`` = 18,
  ``has_harassment`` = 18,
  ``has_self_injury`` = 18,
  ``has_vulgar_language`` = 18,
  ``has_sensitive_content`` = 18
)

# group apps into categories, leave coded apps separate
```

```
categorize_app_237 <- function(name) {
  case_when(
    name == ``Instagram`` ~ ``Instagram``,
    name == ``YouTube`` ~ ``YouTube``,
    name == ``Twitter`` ~ ``Twitter``,
    name == ``Tumblr - Culture, Art, Chaos`` ~ ``Tumblr``,
    name == ``Boost for reddit`` ~ ``Reddit``,
    name %in% c(``Discord - Talk, Video Chat & Hang Out
       with Friends``,
                ``Messaging``, ``Gmail``,
                ``Twitch: Livestream Multiplayer Games &
                   Esports``) ~ ``Social``,
    name %in% c(``Big Farm: Mobile Harvest - Free Farming
       Game``, ``HQ Trivia``, ``Simon Tatham's␣Puzzles``,␣
       ``Little␣Alchemy␣2``,␣``Pokemon␣GO``,␣``Flow␣Fit:␣
       Sudoku``,␣``Trump's Empire: Idle game``, ``Boost for
        reddit``) ~ ``Gaming``,
    name %in% c(``Google Chrome: Fast & Secure``, ``Google
       ``) ~ ``Browser``,
    name %in% c(``System UI``, ``Home UI``, ``Installer``,
       ``Stanford Screenomics``, ``Clock``, ``Camera``) ~
       ``System``,
    name %in% c(``Swagbucks LIVE``, ``SB Answer - Surveys
       that Pay``,
                ``Luck Is On Your Side``, ``SimpliSafe Home
                    Security App``) ~ ``Finance``,
    name %in% c(``Google Photos``, ``QR code reader``) ~ ``
```

```
        Files``,
      name %in% c(``Weather by WeatherBug: Live Radar Map &
        Forecast``,
                  ``Twilight: Blue light filter for better
                    sleep``) ~ ``Utility``,
      is.na(name) | name == ```` ~ ``NA``, TRUE ~ ``System``
    )
  }

  # prep data with row index
  df_subset_237 <- rbind(
    pid237_merged[1330:1390, ],
    pid237_merged[2040:2130, ],
    pid237_merged[2437:2458, ]
  ) %>%
    mutate(
      app_group = categorize_app_237(name),
      row_index = row_number(),
      time_label = format(datetime, ``%H:%M``)
    )

  # collapse into blocks using row_index
  df_blocks_237 <- df_subset_237 %>%
    mutate(block_id = cumsum(app_group != lag(app_group,
       default = first(app_group)))) %>%
    group_by(block_id, app_group) %>%
    summarise(
```

```
    start = min(row_index),
    end   = max(row_index),
    .groups = ``drop``
  )

# flags using row_index
df_flags_237 <- df_subset_237 %>%
  select(row_index, all_of(flag_cols)) %>%
  tidyr::pivot_longer(cols = all_of(flag_cols), names_to =
     ``flag``, values_to = ``value``) %>%
  filter(value == ``True``)

# x axis breaks and labels
x_breaks <- df_subset_237$row_index[seq(1, nrow(df_subset_
   237), by = 10)]
x_labels <- df_subset_237$time_label[seq(1, nrow(df_subset_
   237), by = 10)]

# panel 1: app bar with flags
p1 <- ggplot() +
  geom_rect(
    data = df_blocks_237,
    aes(xmin = start, xmax = end, ymin = -0.4, ymax = 0.4,
       fill = app_group),
    color = ``white``, linewidth = 0.3, alpha = 0.5
  ) +
  geom_point(
```

```
      data = df_flags_237,
      aes(x = row_index, y = 0, shape = flag, color = flag),
      size = 2.5, stroke = 0.4
    ) +
    scale_fill_manual(
    values = app_colors_237,
    name   = ``App``,
    limits = c(``Tumblr``, ``Reddit``, ``Social``, ``Browser
      ``, ``System``, ``Utility``, ``NA``)
  ) +
    scale_color_manual(values = flag_colors, name = ``Content
        Flag``) +
    scale_shape_manual(values = flag_shapes, name = ``Content
        Flag``) +
    scale_x_continuous(breaks = x_breaks, labels = x_labels)
      +
    coord_cartesian(ylim = c(-0.6, 0.6)) +
    theme_minimal(base_size = 11) +
    theme(
      axis.title.y     = element_blank(),
      axis.text.y      = element_blank(),
      axis.ticks.y     = element_blank(),
      axis.title.x     = element_blank(),
      axis.text.x      = element_blank(),
      panel.grid       = element_blank(),
      legend.position  = ``top``,
      legend.title     = element_text(face = ``bold``, size =
```

```
        9),
      legend.text       = element_text(size = 8),
      plot.title        = element_text(size = 13, face = ``
        bold``),
      plot.background  = element_rect(fill = ``white``, color
        = NA),
      panel.background = element_rect(fill = ``#F8F8F8``,
        color = NA)
    ) +
    labs(title = ``PID 237``) +
  guides(
      color = guide_legend(nrow = 2, byrow = TRUE),
      shape = guide_legend(nrow = 2, byrow = TRUE),
      fill  = guide_legend(nrow = 2, byrow = TRUE)
    )

  # reorder legend
  p1 <- p1 +
    theme(legend.box = ``horizontal``) +
    guides(
      color = guide_legend(nrow = 2, byrow = TRUE, order = 2)
        ,
      shape = guide_legend(nrow = 2, byrow = TRUE, order = 2)
        ,
      fill  = guide_legend(nrow = 2, byrow = TRUE, order = 1)
    )
```

```
# panel 2: valence and arousal
p2 <- ggplot(df_subset_237, aes(x = row_index)) +
  geom_line(aes(y = valence, color = ``Valence``),
     linewidth = 0.7, na.rm = TRUE) +
  geom_line(aes(y = arousal, color = ``Arousal``),
     linewidth = 0.7, na.rm = TRUE) +
  scale_color_manual(values = c(``Valence`` = ``
     darkolivegreen``, ``Arousal`` = ``violetred1``), name
     = ````) +
  scale_x_continuous(breaks = x_breaks, labels = x_labels)
     +
  theme_minimal(base_size = 11) +
  theme(
    axis.text.x      = element_text(angle = 45, hjust = 1,
       size = 9),
    axis.title.x     = element_text(face = ``bold``, size =
        11),
    axis.title.y     = element_text(size = 9),
    panel.grid.major = element_line(color = ``grey90``,
       linewidth = 0.3),
    panel.grid.minor = element_blank(),
    legend.position  = ``top``,
    legend.title     = element_blank(),
    plot.background  = element_rect(fill = ``white``, color
        = NA),
    panel.background = element_rect(fill = ``#F8F8F8``,
       color = NA)
```

```
  ) +
  labs(x = ``Time``, y = ``Score``) +
  guides(color = guide_legend(nrow = 2, byrow = TRUE))

# combine
p1 / p2 + plot_layout(heights = c(2, 1))
```

 # Steam Plot - Participant B

```
# example code shown below for Participant B; completed
   once for both case studies
cols <- c(``Audio`` = ``#000000``,
          ``Auto & Vehicles`` = ``#392971FF``,
          ``Books & Reference`` = ``#4d7a99``,
          ``Business`` = ``#4669E0FF``,
          ``Communication`` = ``#3AA3FCFF``,
          ``Dating`` = ``#58DBFF``,
          ``Device Maintenance`` = ``#85FFF9``,
          ``Entertainment`` = ``#f0edd8``,
          ``Finance`` = ``#f5ec42``,
          ``Food & Drink`` = ``#FFC427``,
          ``Health & Fitness`` = ``#FFAE0D``,
          ``Home UI`` = ``#FF9D2B``,
          ``House & Home`` = ``#FF8412``,
          ``Installer`` = ``#EA4F0DFF``,
          ``Knowledge`` = ``#FF5E00``,
          ``Libraries and Demo`` = ``#F03113``,
          ``Lifestyle`` = ``#D62B11``,
```

```
``Maps & Navigation`` = ``#B2260F``,
``Medical`` = ``#B03569``,
``Music & Audio`` = ``#FFA7FC``,
``News & Magazines`` = ``#F09CED``,
``Personalization`` = ``#D68BD4``,
``Photography`` = ``#9E9EF0``,
``Productivity`` = ``#656AD6``,
``Security`` = ``#3C42A3``,
``Settings`` = ``#7178F0``,
``Shopping`` = ``#08F000``,
``Social`` = ``#06B800``,
``Sports`` = ``#287014``,
``System`` = ``#F0761F``,
``Tools`` =  ``#B9491E``,
``Travel & Local`` = ``#723918``,
``Video Players & Editors`` = ``#452320``,
``Card`` = ``#C0392B``,
``Strategy`` = ``#8E44AD``,
``Word`` = ``#2C3E50``,
``Action`` = ``#E74C3C``,
``Role Playing`` = ``#1ABC9C``,
``Board`` = ``#3498DB``,
``Arcade`` = ``#E67E22``,
``Casino`` = ``#2ECC71``,
``Puzzle`` = ``#F39C12``,
``Casual`` = ``#D35400``,
``Simulation`` = ``#27AE60``,
```

```
          ``Trivia`` = ``#8E44AD``,
          ``Adventure`` = ``#16A085``,
          ``Racing`` = ``#C0392B``,
          ``Weather`` = ``#87CEEB``,
          ``Educational``=``#98FB98``,
          ``Music`` = ``#FFA7FC``,
          ``Education`` = ``#98FB98``)

# prep data
ss_237_count <- ss_237 %>%
  mutate(
    date = as.Date(date),
    week = as.Date(cut(date, ``2 weeks``))
  ) %>%
  filter(!is.na(category) & category != ````) %>%
  group_by(week, category) %>%
  summarise(count = n(), .groups = ``drop``)

# stream plot
stream_all_237 <- ss_237_count %>%
  ggplot(aes(x = week, y = count, fill = category)) +
  geom_area(position = ``stack``) +
  scale_fill_manual(values = cols, drop = TRUE) +
  scale_x_date(``Date``,
               breaks = function(x) seq.Date(from = as.Date
                  (``2020-06-16``), to = as.Date
                  (``2021-06-15``), by = ``2 weeks``),
```

```
              minor_breaks = function(x) seq.Date(from =
                as.Date(``2020-06-16``), to = as.Date
                (``2021-06-15``), by = ``1 week``)) +
  scale_y_continuous(``Count``) +
  ggtitle(``App Categories Over Time - Participant B``) +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5, face = ``bold``,
        margin = margin(b = 20)),
    axis.title = element_blank(),
    axis.text.x = element_text(angle = 90),
    axis.text.y = element_blank(),
    strip.text.y = element_blank(),
    legend.position = ``bottom``,
    legend.title = element_blank(),
    legend.text = element_text(size = 9, color = ``grey40
       ``),
    legend.box.margin = margin(t = 30),
    legend.background = element_rect(color = ``grey80``,
       linewidth = .5),
    legend.key.height = unit(.25, ``lines``),
    legend.key.width = unit(2.5, ``lines``),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank()
  )

# add seasonal annotations
```

```
stream_annotated_237 <- stream_all_237 +
  geom_vline(aes(xintercept = as.Date(``2020-09-22``)),
             color = ``grey50``, linewidth = .3, linetype =
                ``longdash``) +
  annotate(``text``, x = as.Date(``2020-09-27``), y =
     60000, label = ``Autumn``,
           angle = 90, size = 3.5, color = ``grey5``) +
  geom_vline(aes(xintercept = as.Date(``2020-12-21``)),
             color = ``grey50``, linewidth = .3, linetype =
                ``longdash``) +
  annotate(``text``, x = as.Date(``2020-12-26``), y =
     60000, label = ``Winter``,
           angle = 90, size = 3.5, color = ``grey5``) +
  geom_vline(aes(xintercept = as.Date(``2021-03-21``)),
             color = ``grey50``, linewidth = .3, linetype =
                ``longdash``) +
  annotate(``text``, x = as.Date(``2021-03-26``), y =
     60000, label = ``Spring``,
           angle = 90, size = 3.5, color = ``grey5``) +
  geom_vline(aes(xintercept = as.Date(``2021-06-21``)),
             color = ``grey50``, linewidth = .3, linetype =
                ``longdash``) +
  annotate(``text``, x = as.Date(``2021-06-26``), y =
     60000, label = ``Summer``,
           angle = 90, size = 3.5, color = ``grey5``)

stream_annotated_237
```

 # App Category by Weekday by Hour Clock Plot

```
# example code shown below for Participant A; completed
   once for both case studies
clock_cols <- c(``Audio`` =  ``#000000``,
          ``Auto & Vehicles`` = ``#392971FF``,
          ``Books & Reference`` = ``#4d7a99``,
          ``Business`` = ``#4669E0FF``,
          ``Communication`` = ``lightskyblue2``,
          ``Dating`` = ``#58DBFF``,
          ``Device Maintenance`` = ``#85FFF9``,
          ``Entertainment`` = ``orchid1``,
          ``Finance`` = ``#f5ec42``,
          ``Food & Drink`` = ``#FFC427``,
          ``Health & Fitness`` = ``#FFAE0D``,
          ``Home UI`` = ``#FF9D2B``,
          ``House & Home`` = ``#FF8412``,
          ``Installer`` = ``#EA4F0DFF``,
          ``Knowledge`` = ``#FF5E00``,
          ``Libraries and Demo`` = ``#F03113``,
          ``Lifestyle`` = ``#D62B11``,
          ``Maps & Navigation`` = ``#B2260F``,
          ``Medical`` = ``#B03569``,
          ``Music & Audio`` = ``hotpink``,
          ``News & Magazines`` = ``#F09CED``,
          ``Personalization`` = ``#D68BD4``,
          ``Photography`` = ``#9E9EF0``,
          ``Productivity`` = ``#656AD6``,
```

```
``Security`` = ``#3C42A3``,
``Settings`` = ``#7178F0``,
``Shopping`` = ``#08F000``,
``Social`` = ``#06B800``,
``Sports`` = ``#287014``,
``System`` = ``#F0761F``,
``Tools`` =  ``#B9491E``,
``Travel & Local`` = ``#723918``,
``Video Players & Editors`` = ``#452320``,
``Card`` = ``#C0392B``,
``Strategy`` = ``#8E44AD``,
``Word`` = ``#2C3E50``,
``Action`` = ``#E74C3C``,
``Role Playing`` = ``#1ABC9C``,
``Board`` = ``#3498DB``,
``Arcade`` = ``#E67E22``,
``Casino`` = ``#2ECC71``,
``Puzzle`` = ``#F39C12``,
``Casual`` = ``#D35400``,
``Simulation`` = ``#27AE60``,
``Trivia`` = ``#8E44AD``,
``Adventure`` = ``#16A085``,
``Racing`` = ``#C0392B``,
``Weather`` = ``#87CEEB``,
``Educational``=``#98FB98``,
``Music`` = ``#FFA7FC``,
``Education`` = ``#98FB98``)
```

```
# prep data - all categories per hour/day
clock_175 <- ss_175 %>%
  mutate(
    date = as.Date(date),
    hour = datetime_hour,
    day_of_week = factor(weekdays(date),
                         levels = c(``Monday``, ``Tuesday
                            ``, ``Wednesday``,
                                   ``Thursday``, ``Friday
                                      ``, ``Saturday``, ``
                                      Sunday``))
  ) %>%
  filter(!is.na(category) & category != ```` & !is.na(hour)
      & !is.na(day_of_week)) %>%
  group_by(hour, day_of_week, category) %>%
  summarise(count = n(), .groups = ``drop``)

# get only categories present in the plotted data
active_cats <- unique(clock_175$category)
cols_filtered <- clock_cols[names(clock_cols) %in% active_
   cats]

# plot with filtered colors
ggplot(clock_175, aes(x = factor(hour), y = count, fill =
   category)) +
```

```
geom_bar(stat = ``identity``, position = ``fill``, width
   = 1, color = ``white``, linewidth = 0.1) +
facet_wrap(~ day_of_week, ncol = 4) +
scale_fill_manual(
values = cols_filtered,
name   = ``Category``,
breaks = names(cols_filtered),
labels = names(cols_filtered)) +
coord_polar(theta = ``x``) +
scale_x_discrete(
  limits = as.character(0:23),
  labels = c(``12am``, rep(````, 2), ``3am``, rep(````,
     2), ``6am``,
              rep(````, 2), ``9am``, rep(````, 2), ``12pm
                 ``,
              rep(````, 2), ``3pm``, rep(````, 2), ``6pm
                 ``,
              rep(````, 2), ``9pm``, rep(````, 2))
) +
theme_minimal(base_size = 10) +
theme(
  plot.title = element_text(hjust = 0.5, face = ``bold``,
      margin = margin(b = 20)),
  axis.title = element_blank(),
  axis.text.y = element_blank(),
  axis.text.x = element_text(size = 6, color = ``grey40
     ``),
```

```
    strip.text = element_text(face = ``bold``, size = 9),
    panel.grid.major = element_line(color = ``grey90``,
       linewidth = 0.3),
    panel.grid.minor = element_blank(),
    legend.position = ``bottom``,
    legend.title = element_text(face = ``bold``, size = 9),
    legend.text = element_text(size = 7),
    legend.key.height = unit(0.5, ``lines``),
    legend.key.width = unit(2, ``lines``)
  ) +
  labs(title = ``App Category Usage by Hour and Day -
     Participant A``)
```

 # Proportion of Apps Per Categories

```
# pivot to long format
apps_long <- apps %>%
  pivot_longer(
    cols = c(X1a_communicative, X1b_socialmedia, X2a_
       tailoredcontent,
             X2b_trendingcontent, X3a_mentalhealth, X4a_
                graphicpurpose,
             X4b_graphicaccess, X5a_socialconnections, X5b_
                workproductivity,
             X6_faces),
    names_to  = ``variable``,
    values_to = ``response``
  ) %>%
```

```
  filter(!is.na(response)) %>%
  mutate(
    variable = recode(variable,
      ``X1a_communicative`` = ``Communicative``,
      ``X1b_socialmedia`` = ``Social Media``,
      ``X2a_tailoredcontent`` = ``Tailored Content``,
      ``X2b_trendingcontent`` = ``Trending Content``,
      ``X3a_mentalhealth`` = ``Mental Health``,
      ``X4a_graphicpurpose`` = ``Graphic Purpose``,
      ``X4b_graphicaccess`` = ``Graphic Access``,
      ``X5a_socialconnections`` = ``Social Connections``,
      ``X5b_workproductivity`` = ``Work Productivity``,
      ``X6_faces`` = ``Faces``
    ),
    response = factor(response, levels = c(``Yes``, ``No``)
     )
  )

# plot
ggplot(apps_long, aes(x = variable, fill = response)) +
  geom_bar(position = ``fill``, color = ``white``,
     linewidth = 0.3) +
  scale_fill_manual(values = c(``Yes`` = ``darkolivegreen
     ``, ``No`` = ``darkred``), name = NULL) +
  scale_y_continuous(labels = scales::percent) +
  theme_minimal(base_size = 11) +
  theme(
```

```
    plot.title = element_text(face = ``bold``, hjust = 0.5,
        margin = margin(b = 15)),
    axis.title.x = element_blank(),
    axis.text.x = element_text(angle = 45, hjust = 1, size
       = 9, color = ``black``),
    axis.text.y = element_text(color = ``black``),
    axis.title.y = element_text(color = ``black``),
    panel.grid.minor = element_blank(),
    panel.grid.major.x = element_blank(),
    legend.position  = ``top``
  ) +
  labs(title = ``App Category Ratings``, y = ``Proportion
     ``)
```

Number of Apps per Categories and Conjunction Categories

```
library(UpSetR)


# prep data - convert Yes/No to 1/0
apps_upset <- apps %>%
  mutate(across(c(X1a_communicative, X1b_socialmedia, X2a_
     tailoredcontent, X2b_trendingcontent, X3a_mentalhealth
     , X4a_graphicpurpose, X4b_graphicaccess, X5a_
     socialconnections, X5b_workproductivity, X6_faces),~
     ifelse(. == ``Yes``, 1, 0))) %>%
  rename(
    ``Communicative`` = X1a_communicative,
    ``Social Media`` = X1b_socialmedia,
```

```
    ``Tailored Content`` = X2a_tailoredcontent,
    ``Trending Content`` = X2b_trendingcontent,
    ``Mental Health`` = X3a_mentalhealth,
    ``Graphic Purpose`` = X4a_graphicpurpose,
    ``Graphic Access`` = X4b_graphicaccess,
    ``Social Connections`` = X5a_socialconnections,
    ``Work Productivity`` = X5b_workproductivity,
    ``Faces`` = X6_faces
  ) %>%
  select(``Communicative``, ``Social Media``, ``Tailored
     Content``, ``Trending Content``, ``Mental Health``, ``
     Graphic Purpose``, ``Graphic Access``, ``Social
     Connections``, ``Work Productivity``, ``Faces``) %>%
  as.data.frame()

# plot
upset(apps_upset,
      nsets = 10,
      nintersects = 20,
      order.by = ``freq``,
      sets.bar.color = ``steelblue``,
      main.bar.color = ``darkorchid4``,
      text.scale = c(1.2, 1.2, 1.2, 1.2, 1.2, 1.2),
      point.size = 3,
      line.size = 0.8,
      set_size.show = FALSE,
      set_size.scale_max = NULL)
```