Filed Under Seal

Exhibit T

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

PEOPLE OF THE STATE OF CALIFORNIA, et al., Plaintiffs, v. *META PLATFORMS, INC, Instagram, LLC, Meta Payments, Inc., Meta Platforms Technologies, LLC.,* Defendants IN RE: SOCIAL MEDIA ADOLESCENT ADDICTION/PERSONAL INJURY PRODUCTS LIABILITY LITIGATION THIS DOCUMENT RELATES TO: 4:23-cv-05448	MDL No. 3047 Case No. 4:23-cv-05448-YGR Judge: Hon. Yvonne Gonzalez Rogers Magistrate Judge: Hon. Peter H. Kang

TRIAL REPORT OF COLIN GRAY
May 16, 2025

27. I took a series of screenshots and videos on an Apple iPhone 16 Pro using the most current versions of the apps as of March and April 2025.[2] As I worked through each task, I used this trail of evidence to take notes that described my findings in each stage of the platform, comparing these early findings to best practices in UX (e.g., discoverability, interpretability) and known dark patterns or attention capture patterns (e.g., noting that a task flow was unnecessarily obstructive). Wherever my findings were ambiguous, I attempted to complete the task a second or third time, noting how a user might discover certain actions in different parts of the interface.

28. The above is the same type of review and data collection methodology I have used in my own peer-reviewed research, for example, Gray, Santos et al., 2021.

c. Identification of Relevant Dark Patterns and Attention Capture Patterns Using Temporal Analysis of Dark Patterns (TADP) Methodology

29. Once I completed my initial platform review, I performed a formal analysis of the platform functionality in order to identify dark patterns and attention capture patterns that were present. To accomplish this analysis, I used *temporal analysis* to guide my expert evaluation of interfaces to identify and characterize dark patterns at multiple scales of interaction. Temporal analysis can be used to identify not only specific points in the interface where dark patterns are present in isolation, but also the cumulative and aggregated impacts of dark patterns on the experience of users across their holistic interaction with a platform (Gray, Mildner, & Gairola, 2025; Gray, Gairola, & Mildner, 2025).

30. This Temporal Analysis of Dark Patterns (TADP) methodology uses expert inspection, synthesis, and visualization to demonstrate how dark patterns occur in layered ways that increase the duration of exposure and result in cumulative effects that may only be felt over time. These characteristics are mapped using user interface elements and aspects of a platform's delivery strategy across three potential scales: *intra-page*, which indicates the presence of dark patterns on a single interactive page experience using annotated wireframes; *inter-page*, which

[2] Versions of the app used during the analysis period were as follows: Instagram iOS versions included 368.0.0 to 375.0.0.

indicates the presence of dark patterns across multiple pages of a user's journey, visually depicting how they seek to achieve their goal using a task flow diagram; and *system*, which indicates the presence of dark patterns across an entire platform delivery strategy, visually indicating subscription, deactivation or deletion, and other relevant user actions using a modified blueprint method. In my analysis, I found the system-level mapping of dark patterns to be most salient, since many of the observed patterns related to attention capture and manipulation of users' engagement over time. Thus, my analysis of Meta's platforms begins with a system level analysis, followed by analysis of key features or touchpoints within the overall user experience that I felt were most salient in describing the overall experience a child or teen would have. Notably, since each platform is complex, my analysis is not exhaustive—and other features could be similarly mapped and contextualized.

31. In my analysis, I identify, contextualize, and classify dark patterns in relation to their presence and role in the overall user experience. This process includes three main stages, consistent with the TADP methodology (Gray, Mildner, & Gairola, 2025; Gray, Gairola, & Mildner, 2025). They are: (i) identification and characterization; (ii) building static and temporal relationships; and (iii) co-occurrence and amplification.

32. As indicated above, the TADP methodology is the one I use in peer-reviewed academic publications (e.g., Gray, Mildner, & Gairola, 2025; Gray, Gairola, & Mildner, 2025).[3]

i. Identification and Characterization

33. In this stage, I identify the presence of specific dark patterns or attention capture patterns that I found in my expert review of Meta's platforms. Definitions of dark patterns that I utilize in this report are built from academic consensus and definitions used in regulatory reports. In the development of this ontology of dark patterns, my colleagues and I have categorized specific dark patterns into high-level (strategic intent), meso-level (design tactics), and low-level (specific UI elements) characterizations, building on multiple academic and regulatory reports

[3] Previous versions of this methodology were grounded in my prior empirical work (e.g., Gray, Santos, et al., 2021).

that outlined common dark patterns and their manifestations (Gray, Santos et al., 2023; 2024). Each instance of a dark pattern is classified according to these structured ontologies to maintain consistent and detailed reporting.

34. Identification and characterization processes encompass high-, meso-, and low-level descriptions whenever possible, though novel patterns may initially only receive high- and meso-level characterizations, prompting the identification of new low-level pattern types. Factors considered in the identification process include readable text, layout, relative size and positioning of UI elements, use of color, typography, text decoration, feedback mechanisms, and task flows that involve UI elements and screens, as well as contextual or medium-specific considerations (Gray, Santos, et al., 2021; Gray, Mildner, & Gairola, 2021; 2025).

35. Attention capture patterns are identified using the taxonomy from Monge Roffarello et al. (2023). A fuller description of these patterns, and their overlap with dark patterns, is discussed in Section V(b). I was also able to leverage the analysis of other academic researchers that have studied Meta's platforms. For instance, Mildner et al. (2023) found a wide range of dark patterns in their assessment of Instagram that triangulate with my own expert review findings. In addition, recent work by Seyson and Willett (2025) has shown the evolution of specific functionality on Instagram from 2010-2024, including the use of dark patterns to obfuscate account deletion—practices which increased over time in their pervasiveness and complexity.

ii. Building Static and Temporal Relationships

36. In this analytical stage, I focus on establishing explicit connections between dark patterns and their specific UI manifestations. This analysis takes into account relationships between dark patterns and visible UI elements, the absence of UI elements, user journey transitions, and the presence of specific functionality within the overall platform delivery strategy. Through this lens, high- and meso-level patterns are contextualized within specific UI components, sequences across screens, or platform touchpoints.

describing the need for features or proposals for new features to address negative user experiences (reactive correction).

e. Alignment of Dark Patterns and Attention Capture Patterns with Relevant Professional Behaviors

40. Finally, my analysis assesses the impact of dark patterns and attention capture patterns on user experiences, particularly emphasizing vulnerabilities among teenage users and their parents or guardians. My focus here is in demonstrating that the combination of dark patterns and attention capture patterns has a cumulative impact, but also that this impact is acutely felt by adolescent users.

41. I align findings from this analysis stage to professional principles from the ACM Code of Ethics, illustrating that a technologist acting consistently with professional guidelines would likely identify these issues proactively (considering potential harms) or, at the very least, reactively (identifying felt harms and working rapidly to correct them). These principles can easily be activated through ethics-focused methods (which I have published in two collections: Chivukula et al., 2024; Chivukula & Gray, 2025), providing workflows and approaches to identify "positive counterfactuals"—or design decisions of the type that would have been considered and/or prioritized by professionals following those principles.

V. The Scope, Impact, and Presence of Dark Patterns

42. The term "dark patterns" was coined in 2010 by UX design practitioner Dr. Harry Brignull (Brignull, 2011). The term was originally proposed as a way for design practitioners to express concern about the state of manipulative and deceptive design practices in the industry, providing a mechanism for practitioners to "name and shame" companies[5] relying upon these patterns in hopes of changing these design practices (Obi et al., 2025). Brignull's initial definition of dark patterns described instances where designers "take our understanding of human psychology and flip it over to the dark side" (Brignull, 2011). In a later popular press

[5] Chiang, Oliver. (2010, Sep. 22) Naming and Shaming Sneaky Webs Design Tricks of Facebook, Yahoo and Others. Forbes. https://www.forbes.com/sites/oliverchiang/2010/09/22/naming-and-shaming-sneaky-web-design-tricks-of-facebook-yahoo-and-others/

discoverability of important information (or omits the information entirely), causing them to make uninformed choices, such as agreeing to terms that were not fully disclosed. The use of **social engineering** causes a user to feel the need to "fit in," responding to pressures to perform certain actions that are supported by tactics such as social proof or scarcity to falsely convince a user something is extremely popular (e.g., "99% of users opt-in") or imparting urgency (e.g. "187 other shoppers are looking at this item") so they make hasty decisions.

b. *Meso-level patterns* function as context-agnostic angles of attack, using tactics such as manipulation of the visual choice architecture, utilizing an attentional "roach motel" that is more difficult to get out of than get into (e.g., subscriptions (Sheil et al., 2024)), bad defaults that favor shareholders over users, (e.g., dark as opposed to "bright" or privacy preserving patterns (Graßl et al., 2021)) or hiding information (a type of **sneaking** and **interface interference**) that is relevant to a user making an informed choice (e.g., preventing users from understanding the impacts of location settings (Gray, Gairola, & Mildner, 2025)).

c. *Low-level patterns* are context-specific and refer to specific tactics within a user interface, including privacy mazes to split up the information needed to make an informed choice across many connected pages, false hierarchies that confuse the user or privilege one action over another, or what has come to be known in the literature as "privacy zuckering"—named for Meta CEO Mark Zuckerberg—where users are tricked into sharing more personal information than they intend (Bösch et al, 2016).[8]

[8] See Leiser, Mark. (2022). Illuminating Manipulative Design: From "Dark Patterns" to Information Asymmetry and the Repression of Free Choice Under the Unfair Commercial Practices Directive. Loyola Consumer Law Review 34, 484-491.

encouraging a user to lose track of their goals, contributing to a sense of lost time or control, or instilling a sense of regret about the time they have spent on a platform. These criteria relate to the characteristics of dark patterns proposed by Mathur et al. (2021) with a focus on specific kinds of manipulation (e.g., use of psychological biases and automation) and harms (e.g., attentional, addiction-oriented). Importantly, the use of ACDPs is highly effective in increasing user engagement; as noted by Lee et al. (2025) in a recent experiment on social media websites, the introduction of a browser extension that allowed users to turn "off" common ACDPs resulted in much lower levels of distraction (7% versus 28% on the unaltered social media site) and fewer daily minutes of engagement (1.61 hours versus 1.96 hours).

53. The collection of ACDPs is dominated by examples from social media, including Meta's platforms. Monge Roffarello et al.'s (2023) initial typology of ACDPs include a wide range of practices utilized by social media platforms to drive and sustain engagement, such as increasing the capacity for consumption of content (e.g., **infinite scroll**; casino pull-to-refresh; never-ending **auto-play**), reactivation of engagement (e.g., **recapture notifications**, fake social notifications), reducing awareness of time spent (e.g., time fog), increasing certain patterns of engagement (e.g., **playing by appointment**, which refers to being forced to use a digital platform with a consequence of losing achievements if they do not), and making discontinuation of a platform difficult or impossible (e.g., attentional "roach motel"). Thus, while dark patterns generally target a wide range of manipulative, deceptive, and/or coercive behaviors, ACDPs focus specifically on attention, time, and emotional manipulation. The intersection of these two categories is critical in understanding how design practices exploit psychological vulnerabilities—especially the vulnerabilities of adolescents—to contribute to harmful user experiences and to undermine users' control over their digital engagement.

c. <u>Dark Patterns Undermine Consumer Autonomy</u>

54. Dark patterns scholars have described the power that designers and technologists have in modifying users' perceptions of the choice architecture—shaping how users perceive what choices are available to them, which then limits or even determines possible courses of

digital contexts. Two of the most comprehensive studies on consumer behavior have shown strong causal connections between the presence of dark patterns and outcomes that are disadvantageous to users. Luguri and Strahilevitz (2021) showed that "users exposed to mild dark patterns were more than twice as likely to sign up for a dubious service as those assigned to the control group, and users in the aggressive dark pattern condition were almost four times as likely to subscribe." In a large online experiment conducted by the EU Commission, dark patterns were shown to impact "consumers' transactional decision […] l[eading] to inconsistency with their preferences" (Behavioral study on unfair commercial practices in the digital environment: dark patterns and manipulative personalization: final report, 2022). The removal of key attentional triggers also manipulates the information flow and limits discoverability by a user, resulting in substantial declines in user engagement and demonstrating the power of these manipulative design tactics (Lee et al., 2025; Schaffner et al., 2024).

58. As noted by Rossi et al. (2024) and Sanchez Chamorro (2024), all users are vulnerable to some extent, but some users are more vulnerable than others. Research on the role of digital platforms for children (whose patterns of use then impact their behaviors during their teen years) show that children, their parents, and their friends are impacted as part of a technology ecology (e.g., Hiniker et al., 2017; Fitton et al., 2021). These findings are echoed by recent scholarship on adolescents specifically (Sanchez Chamorro, et al., 2024). Schäfer et al. (2024) also showed that children were able to identify some relevant mechanisms relating to simple dark patterns, but more complex formulations of dark patterns were missed by the majority of children. These findings illustrate that children are more vulnerable to dark patterns than adults, and that this vulnerability is likely based both in their developmental cognitive ability and their relative degree of digital literacy.

d. Research Has Identified Dark Patterns on Meta's Platforms

59. Mildner et al. (2023) conducted the first comprehensive analysis of dark patterns across several major social networking platforms, including Facebook, Instagram, and TikTok, revealing a pervasive presence of manipulative design practices. In that study, researchers

attention. EUDs, such as **infinite scroll** and **auto-play**, contribute to what has been termed **attention capture**, whereby users are nudged to continue engaging with the platform. Chen and colleagues found these techniques are not only prevalent across VLOPs but also particularly aggressive on social media platforms, concluding: "social media platforms that depend on user attention employed twice as many EUDs as platforms that rely on other business models." These findings further highlight the intersection of dark patterns and attention-capture techniques, showing how these designs exploit users' time, attention, and cognitive load, creating an environment that discourages disengagement and increases the chance for users to experience multiple forms of harm.

61. The findings of these studies are also mirrored by a 2025 report from a consumer protection organization in Germany (the Verbraucherzentrale Bundesverband e.V.) that underscores the growing concerns regarding "Hyper-Engaging Dark Patterns" on social media platforms (Bundesverband der Verbraucherzentralen und Verbraucherverbände, 2025). These manipulative design techniques—including push notifications, countdowns, **infinite scroll**, and **auto-play**—are specifically designed to induce addictive behavior, particularly among young users, and leverage the same mechanisms as those exploited in gambling, such as stimulated variably timed dopamine release. The report highlights how these techniques, often using psychological mechanisms like the fear of missing out (FOMO) and variable rewards, drive users to spend excessive amounts of time on the platforms. The report notes that addictive engagement is especially concerning for minors, who are more vulnerable to the emotional and psychological effects of these designs, and could lead to negative long-term consequences, such as impaired decision-making and addictive behaviors.

62. Researchers have concluded that the ubiquity of dark patterns in digital systems, including on social media platforms, is by design: these patterns are effective at enabling shareholder goals and often represent a focus on business incentives to the detriment of user experience and societal impact (Leiser & Santos, 2024; Mathur et al., 2021). However, the normalization of these practices is out of alignment with the professional responsibilities of

technologists to produce systems that have positive social impact and promote user autonomy and agency (e.g., ACM Code of Ethics; Nielsen, 1995). Instead, the use of dark patterns in these systems results in what Doctorow (2022) refers to as the "enshittification" of the web, creating user experiences that maximize user engagement at the expense of user value. Instead of providing a user with a sense of control, many dark patterns result in a systemic disempowerment that prevents users from reaching their desired goal (Gray, Santos, et al., 2021; Gray, Gairola, & Mildner, 2025).

VI. Guidelines Applicable to Technologists

63. Technologists, including designers, engineers, and other computing professionals, typically receive ethical training as part of their formal education.[11] This exposure to ethical concepts, theories, and practices aims to prepare them for the ethical complexities of their roles, ensuring they can make informed decisions that promote positive social outcomes. Like codes of ethics that guide designers and engineers,[12] this exposure is overseen in the United States through accreditation. ABET (formerly known as the Accreditation Board for Engineering and Technology) is the relevant accreditation body.[13] While such codes are typically not binding since most technology fields are not licensed (rare exceptions include the title of "professional engineer"), these codes demonstrate the attention of professional fields in technology and design towards avoiding specific negative outcomes for technology users and society.

a. Codes of Ethics and Accreditation Standards Create Duties of Ethical Engagement

64. Many computing, technology, and design societies have established a code of ethics they expect members of the relevant profession to uphold in their work practices.[14] Such codes, reinforced by accreditation standards relating to ethics, typically emphasize the importance of understanding one's responsibility as a professional and making ethical decisions

[11] Buwert, P. (2018, June); Hess, J. L., & Fore, G. (2018).
[12] Buwert, P. (2018, June).
[13] ABET. https://www.abet.org (last accessed May 15, 2025).
[14] In the course of my research, I have only encountered a very small number that do not.

- "[R]eport any signs of system risks that might result in harm. If leaders do not act to curtail or mitigate such risks, it may be necessary to "blow the whistle" to reduce potential harm."

- **1.4 "Be fair and take action not to discriminate."** Responsibilities include:
  - Ensure that "[t]echnologies and practices [are] as inclusive and accessible as possible and computing professionals should take action to avoid creating systems or technologies that disenfranchise or oppress people."
  - "[Failing] to design for inclusiveness and accessibility may constitute unfair discrimination."

- **1.6 "Respect privacy."** Responsibilities include:
  - "[U]nderstand[ing] the rights and responsibilities associated with the collection and use of personal information" and being "conversant in the various definitions and forms of privacy."
  - "[E]stablish transparent policies and procedures that allow individuals to understand what data is being collected and how it is being used, to give informed consent for automatic data collection, and to review, obtain, correct inaccuracies in, and delete their personal data."

- **2.5 "Give comprehensive and thorough evaluations of computer systems and their impacts, including analysis of possible risks."** Responsibilities include:
  - Recognizing that "[a] system for which future risks cannot be reliably predicted requires frequent reassessment of risk as the system evolves in use, or it should not be deployed."
  - Ensuring that "[a]ny issues that might result in major risk [are] reported to appropriate parties."

- **3.1 "Ensure that the public good is the central concern during all professional computing work."** Responsibilities include:
  - A focus on the "public good […as] an explicit consideration when evaluating tasks associated with research, requirements analysis, design, implementation, testing, validation, deployment, maintenance, retirement, and disposal."
  - Attention to this central concern "no matter which methodologies or techniques they use in their practice."

- **3.7 "Recognize and take special care of systems that become integrated into the infrastructure of society."** Responsibilities include:
  - "Continual monitoring of how society is using a system [to] allow the organization or group to remain consistent with their ethical obligations outlined in the Code."

b. <u>Age-Appropriate Design Codes Emphasize the Well-being of Children</u>

68. The ethical responsibility of technologists is also underscored by the expectations of age-appropriate design decisions, especially when designing for vulnerable populations like minors (users under the age of 18) (Fitton et al., 2021; Radesky & Hiniker, 2021; Sanchez-Chamorro et al, 2024). The risks and impacts of such design decisions on children are already well-known in scholarship (e.g., Fitton et al., 2021; Kim et al., 2025), and I observe these

impacts solely to show that adoption of this research informs laws that now also mandate protections when designing for children.[22] All guidance I have reviewed regarding designing appropriate systems for minors indicates that this group—as a whole—is vulnerable and their needs should be considered when designing any system that this group would be likely to access. While all minors can be considered as vulnerable, there are also specific kinds of protections that are relevant at different developmental stages (e.g., under 13 year olds may need protections that 16 year olds may not).[23]

69. For example, California's Age-Appropriate Design Code Act (2022) mandates that companies designing for children consider their well-being and protect them from potentially harmful design practices, such as dark patterns.[24] Characteristics of the design code include, but are not limited to: (1) ensuring that defaults are always the most private (Section 1.a.8); (2) making age-appropriate experiences based on knowledge of specific age ranges (Section 1.a.5); (3) conducting impact assessments to consider whether children could experience or be exposed to harm, including through sustained use of a platform (1798.99.31.a.1.A); and (4) creating meaningful tools to allow a guardian or parent to monitor a

[22] By "children," I refer to users under the age of 18 that should be afforded protections based on their age and abilities. As noted in the California Age-Appropriate Design Code Act (AADCA), "Children should be afforded protections not only by online products and services specifically directed at them, but by all online products and services they are likely to access. In order to help support the design of online products, services, and features, businesses should take into account the unique needs of different age ranges, including the following developmental stages: 0 to 5 years of age or 'preliterate and early literacy'; 6 to 9 years of age or 'core primary school years'; 10 to 12 years of age or 'transition years'; 13 to 15 years of age or 'early teens'; and 16 to 17 years of age or 'approaching adulthood.'" California Age-Appropriate Design Code Act (AB-2273).

[23] Guidance from the California AADCA, UK ICO, and UN all indicate protected status for minors under the age of 18. The California AADCA also identifies protections aligned with specific age ranges under the age of 18. California Age-Appropriate Design Code Act (AB-2273). Smartphone apps similarly consider the needs of different minor groups; the age rating categories currently proposed by Apple for the App Store are divided in this way (4+, 9+, 13+, 16+, 18+). *Helping Protect Kids Online*. (2025, Feb). Apple. https://developer.apple.com/support/downloads/Helping-Protect-Kids-Online-2025.pdf.

[24] California Age-Appropriate Design Code Act (AB-2273). https://leginfo.legislature.ca.gov/faces/billCompareClient.xhtml?bill_id=202120220AB2273&showamends=false#:~:text=(c)%20(1)%20If,the%20requirements%20of%20this%20title. In March 2025, I understand the California AADCA was preliminarily enjoined. I express no opinion on the constitutionality of the Act, but upon review of the design principles indicated within, I believe these principles are sound, able to be implemented by designers, and consistent with ethical norms.

responding to emerging evidence of harm. Ethical design practice is not just about preventing known dangers—it is about embedding care, accountability, and responsiveness into the entire lifecycle of a system.

74. In the case of Meta's platforms, that did not happen. Meta had the ability to anticipate the impact of their design decisions. It had information showing that large numbers of adolescents were using their platforms and that their functionality would create harms—especially for young users still developing self-regulation and digital literacy. And yet, Meta chose not to act, even in cases where its own evidence clearly indicated the harms these platforms were promulgating.

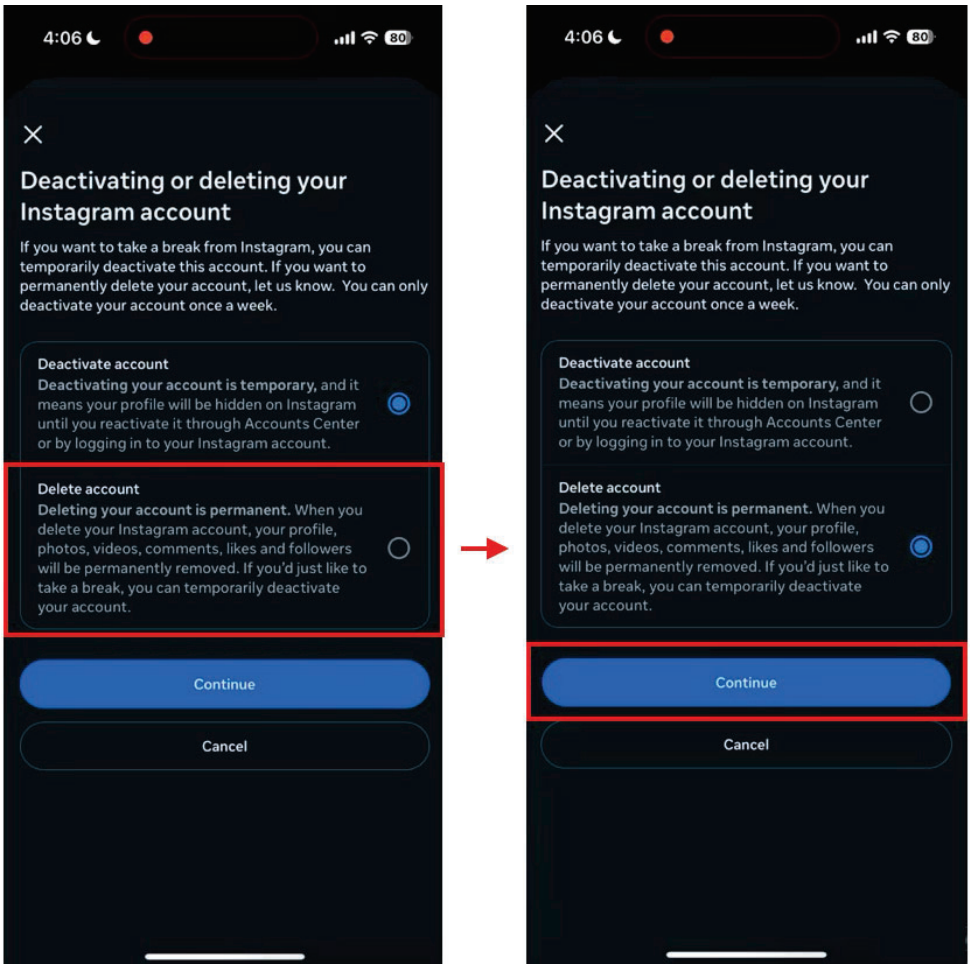
75. These decisions could have leveraged existing resources to take the complexity of digital systems into account. For example, Meta could have used a range of ethics-focused methods (Chivukula et al., 2024; Chivukula & Gray, 2025) or traditional user studies to better understand the potential impacts of its design decisions on user behavior and well-being. In cases where information on potential harms or risks to users already existed (such as through the outcomes of user studies that were performed or utilization of external research), Meta could have turned this into action—making design decisions to minimize further harm to users.

76. Yet in each case, Meta designed (and continued to re-design) systems that failed to meet the principles and standards set out by computing professional societies. None of the platforms implemented sufficient protective defaults for adolescent users, and in many cases, defaults actively placed youth at risk by making their accounts public, enabling contact from strangers, and/or exposing them to addictive content feeds. Notification systems across all platforms employed recapture strategies and nagging, overriding user intent to disengage. Account deletion and supervision flows created additional friction, deterring users from seeking to escape addictive or other problematic use on Meta's platforms.

77. These harms were foreseeable. Meta had access to internal user research, external studies, and regulatory scrutiny indicating that adolescents were particularly vulnerable to the types of engagement loops and feedback systems these platforms engineered. For instance, Meta

94. To cancel or permanently delete an Instagram account, a user must follow numerous steps to even locate the start of the deletion flow. The inclusion of additional steps, many of which are unnecessary, is a similar tactic that has been used to suppress certain kinds of user action in other contexts through the inclusion of multiple dark patterns experienced over time (Gray et al., 2025). Even once the user is able to find an option to delete, there is still another entire flow to work through in order to reach their goal.

- The first option presented (screen #7) is not deletion but rather **pre-selected deactivation**, a reversible action that is framed as safer and temporary, suggesting a path of lower resistance. This preselection is an example of the meso-level pattern **bad defaults**, falling under the broader **interface interference** strategy. A neutral approach would not pre-select either option. Additionally, at this stage, the user is not informed that the delete option has a time delay and is not immediately permanent. This represents the **sneaking** high-level dark pattern strategy, whereby important information is withheld from the user.



manipulative content. This combined effect leaves little opportunity for teen users to pause, understand, and mitigate the psychological and social risks of these settings. Rather than scaffolding intentional or mindful use, Meta's sign-up and onboarding process encourages quick dependency, exploiting the developmental vulnerabilities of teen users in the pursuit of growth and retention metrics.

103. Instead of building in friction to encourage deliberation, Meta's defaults also offered no prompts for teen users to reflect on who might be able to view or interact with their content. A more ethical design that upholds professional responsibilities towards the end user and society would have provided private-by-default accounts and clear, developmentally appropriate options for adjusting visibility—supporting a teen's autonomy while helping them understand and assess the risks of public exposure. In failing to structure these choices in an intentional and informed way, Meta deprived teen users of a significant measure of control over their privacy and safety, and instead normalized high-risk exposure as a default condition.

v. Social Notifications and Engagement Metrics Amplified Attention Capture and Social Comparison

104. Instagram's notification infrastructure plays a central role in driving user engagement, especially among teen users who are developmentally more vulnerable to persuasive and compulsive interface cues (Sanchez Chamorro et al., 2024; see generally Kim et al., 2025). While Meta has publicly acknowledged the importance of helping young users "manage their time" on the platform, the design and delivery of **recapture notifications**—socially driven alerts that encourage users to return to the app—directly undermine those stated goals. Rather than respecting users' attempts to disengage or self-regulate, Instagram's notification system can have the net result of disrupting periods of intentional disengagement, often leveraging relationships or prior interactions to pull users back into the platform. (META3047MDL-136-00013164 at -13213). The power of these recapture techniques is underscored by both academic research (Herder & Moerland, 2024; Lee et al., 2025; Zhang et al., 2022) and Meta's own internal research (e.g., META3047MDL-044-00171345;

and more robust explanations of the risks of visibility, again in consultation with the teen's parents, from the very beginning. More generally, Meta did not begin implementing privacy-respectful defaults for teens and children until March 2021, even though, at least as early as July 2019, Meta researchers recommended doing so. (META3047MDL-065-00124111 at 119; Castaneda Dep. Ex. 14; Kilstein Dep. Ex. 57).

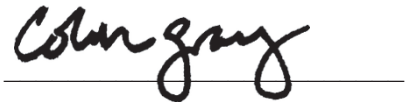
130. **Meaningful Parental Controls**: Meta could have developed parental controls much earlier in the platform's history. And even with the controls it did eventually roll out in 2022 and 2023, the supervisory mechanisms could have been durable and verifiable, including mandatory adult approval for sensitive settings or notification changes, safeguards against alternative account creation, and reliable communication channels that inform adults of critical actions. For instance, guidance from academic research in 2017 on developing self-regulation strategies between parents and children regarding technology tools could have been adopted (Hiniker et al., 2017). Instead, controls were easily disabled and poorly communicated.

131. **Streamlined Exit Options**: For users who sought to delete their accounts, Meta could have offered a clear, linear path to deletion—respecting user agency and mental health. Instead, it created a labyrinth of **forced action** and **obstruction**, making account exit much more difficult than account creation. Design alternatives—such as one-click deletion, shorter and clearly communicated timelines, or deletion flows that prioritized a user's choice—were technically feasible but deprioritized in favor of user retention, as Mark Zuckerberg himself testified (Zuckerberg (Vol. 1) Dep. 207:1-208:23). Meta failed to adopt less harmful alternatives—and took actions to increase friction and the use of dark patterns over time instead of reducing this friction (Seyson & Willett, 2025). It did so despite both information about the problem and available solutions, illustrating ethically problematic deployment of dark patterns to retain users who prefer to leave.

132. **Human-Centered Notification Systems**: Rather than exploiting user anxiety through **recapture notifications** and **social engineering** tactics, Meta could have encouraged users to set boundaries around communication, offering bulk notification control, opt-out

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

Dated: May 16, 2025



Colin M. Gray, PhD