LINKS GLOBAL
Communicating with Purpose. Powered by Insight.

THOUGHT LEADERSHIP SERIES

Social Media and Children

What the Meta lawsuits tell us that the platforms won't

A Links Global thought leadership article by M. Sanchez, founder of Links Global, a communications firm working at the intersection of trust science, narrative design, and civic engagement. Part of the Links Global article series: Contemporary Challenges. Drawn from 25 years of field practice and evidence-based research in strategic communication, public engagement, and institutional trust.

In Brief

The state lawsuits against Meta did more than allege harm. They surfaced the gap between what a platform knew internally and what it said publicly. This article explains what engagement-optimized platforms are built to do, what Meta's own research showed, and what that means for parents, educators, clinicians, and policymakers. It ends with the communications failure underneath the safety failure.

On October 24, 2023, attorneys general from forty-one states and the District of Columbia filed landmark lawsuits against Meta, alleging that the company knowingly designed Instagram and Facebook to be psychologically harmful to children and adolescents, and that it concealed what its own internal research showed about that harm while publicly claiming to prioritize youth safety.

The lawsuit was not the first legal challenge to social media platforms' treatment of minors. It will not be the last. What made it significant (beyond its scale, beyond the number of jurisdictions involved, beyond the specificity of its allegations) was what it surfaced about the relationship between a technology company's internal knowledge and its public communications.

Meta knew. That is the central allegation, supported by internal documents the litigation has brought into public record. Meta's own researchers had documented, internally, the psychological effects of Instagram on adolescent girls: the association between platform use and body image disturbance, depression, and anxiety. In documents first reported by the Wall Street Journal in 2021, 32 percent of teen girls said that when they felt bad about their bodies, Instagram made them feel worse. The company knew this. It commissioned the research. And then, in its public communications, it characterized its platforms as safe for young users, positioned its safety initiatives as genuine protections rather than regulatory management, and continued optimizing the platform design features (infinite scroll, algorithmic content amplification, engagement notifications) that its own research connected to harm.

This article is not primarily about the litigation. It is about what the litigation reveals: about how social media platforms actually function for children and adolescents, about the gap between platform safety claims and platform safety reality, and about what parents, educators, policymakers, and professionals working with young people need to understand to make informed decisions in the current environment.

What the Platforms Actually Do, and Why

Social media platforms are not neutral communication tools. They are engagement-optimization systems: designed, with considerable technical sophistication, to maximize the time users spend on the platform and the frequency with which they return to it. This is not a secondary feature of the platforms. It is their primary business logic: advertising revenue depends on attention, attention depends on engagement, and engagement is maximized through design features that exploit known psychological mechanisms.

The mechanisms these platforms exploit are well-documented in the behavioral science literature and were clearly understood by platform designers before they were applied to youth users.

Variable reward scheduling. The psychological mechanism underlying slot machine addiction, the unpredictable reward that keeps a behavior going more effectively than consistent rewards, is the same mechanism underlying social media notification systems. The uncertainty of whether the next scroll will produce a rewarding post, the uncertainty of whether the notification is a meaningful social signal or a platform prompt, keeps users returning with the same neurological urgency that keeps gamblers at slot machines. This mechanism is effective across all ages. It is particularly potent during adolescence, when reward sensitivity is neurologically heightened and impulse control systems are still developing.

Social comparison architecture. Platforms optimized for engagement present curated, highlight-reel representations of peers' lives at exactly the developmental moment, adolescence, when social comparison is most psychologically significant and most vulnerable to distortion. The research on social comparison and adolescent mental health predates social media. What social media has done is industrialize social comparison at scale, making it continuous, quantified (through likes, follower counts, engagement metrics), and algorithmically amplified toward content most likely to provoke strong emotional responses, including the negative emotional responses associated with upward social comparison.

Algorithmic amplification of distress-adjacent content. Internal documents from multiple platforms, as reported in litigation and press accounts, have indicated that algorithms trained to maximize engagement systematically amplify content associated with strong emotional arousal: including content related to body image, self-harm, eating disorders, and extreme political content. Not because the algorithms were designed to promote harmful content, but because harmful content, like all

emotionally arousing content, drives the engagement metrics the algorithms are optimizing for. The result is that a young user who interacts with content adjacent to these themes is algorithmically served more of it, not because the platform intends harm, but because the harm-adjacent content is engagement-generating.

These mechanisms were not hidden. They were documented in behavioral science literature. They were implemented by design teams that included behavioral scientists. And they were applied to platforms that knew, from the beginning, that significant portions of their user base were minors, often with no meaningful age verification, no design differentiation for younger users, and no parental visibility into the content their children were encountering.

What Meta's Internal Research Actually Showed

The internal documents surfaced by the litigation and by prior journalistic investigation are worth examining in specific terms, because the gap between what they show and what the company communicated publicly is the central fact of the accountability question.

Meta's internal research, as reported through litigation documents and investigative journalism, found that Instagram use was associated with body image issues and depression in adolescent girls at rates that were meaningful and that the company's own researchers characterized as concerning. The research documented that users who were already vulnerable (who were experiencing depression, anxiety, or eating disorder symptoms) were disproportionately served content that reinforced those vulnerabilities. The research identified specific features (including the algorithmic amplification of appearance-related content and the social comparison dynamics of the platform) as contributing factors.

According to the states' allegations, this research was conducted internally, reported internally, and acted on internally, not through design changes that reduced the harms the research documented, but through communications strategies designed to manage the public perception of those harms. The company's public statements about youth safety, its voluntary safety initiatives, and its regulatory testimony all occurred after this internal research was in hand.

The legal and ethical implications of this gap, between internal knowledge and public communication, reached a turning point on August 26, 2026, when Meta settled the multistate case mid-trial. According to the California Attorney General, the settlement is worth up to $17 billion over ten years and requires court-supervised design changes for teen accounts; Meta did not admit wrongdoing. Private and school-district cases continue. But the practical implications for the parents, educators, and professionals making decisions about children's social media use do not wait for the litigation to conclude.

What This Means for Parents

The practical implication of what the litigation has surfaced is that the framing most parents have been given for evaluating social media risk for their children is inadequate.

The dominant parental framing (manage screen time, monitor content, educate children about online safety) treats social media risk as primarily a content problem (exposure to inappropriate material) and a time problem (excessive use displacing healthier activities). These are real concerns. They are also secondary to the structural concern that the litigation has made legible: that the platforms themselves are designed to maximize engagement through psychological mechanisms that are harmful to developing minds, regardless of the specific content encountered.

A child can encounter no inappropriate content on Instagram and still be systematically exposed to social comparison dynamics, variable reward schedules, and algorithmic amplification of distress-adjacent content. The harm is in the architecture, not just the content. And parental management strategies focused on content and time address the architecture only obliquely.

This does not mean that all children who use social media will be harmed, or that harm is inevitable, or that social media has no value for young people. It means that the risk assessment framework most parents are using is not calibrated to the actual risk profile the platforms present, and that making genuinely informed decisions requires understanding what the platforms are actually designed to do.

For parents of adolescents: The research on adolescent social media use and mental health is not uniform, but the preponderance of evidence (the U.S. Surgeon General's 2023 advisory cites a longitudinal study of 6,595 U.S. adolescents aged 12 to 15 in which those spending more than three hours a day on social media faced double the risk of poor mental health outcomes, including symptoms of depression and anxiety), particularly for adolescent girls and for users who are already experiencing mental health challenges, suggests meaningful associations between heavy platform use and outcomes including depression, anxiety, and body image disturbance. The age at which children begin using social media, the platforms they use, the visibility parents maintain into their usage patterns, and the family culture around social media use all appear to be relevant protective factors. But the most protective factor, reducing the structural exposure to the engagement mechanisms the platforms deploy, is not available to parents within the platforms' current design.

For parents of younger children: The platforms' age minimum of 13 is not a safety certification. It is a liability management boundary derived from the parental-consent rules for children under 13 in the Children's Online Privacy Protection Act: a regulatory threshold, not a developmental assessment. Children under 13 routinely access platforms with no meaningful age verification. The developmental appropriateness of social media engagement mechanics for children under adolescence is not a settled question, and the platforms' design was not optimized for developmental appropriateness at any age.

What This Means for Educators and Professionals

The educators, school counselors, pediatricians, mental health professionals, and youth-serving practitioners who work with children and adolescents are on the front lines of whatever harms social media platforms generate, often without adequate information about what they are observing or why.

The clinical presentation of social media-related harm in young people does not always announce its origin. A young person presenting with depression, anxiety, social isolation, body image disturbance, or disrupted sleep may be experiencing a confluence of factors of which social media use is one, and not always the most visible one. The professional capacity to assess social media use as a relevant clinical and developmental factor, to have informed conversations with young people and their families about what the platforms are actually doing, and to make referrals and recommendations grounded in the current evidence base is not yet adequately embedded in most professional training.

School administrators and educators face a specific version of this challenge: the question of how to address social media in educational settings in ways that are evidence-informed, developmentally appropriate, and realistic about what schools can and cannot control in the current environment. Blanket prohibitions on phone use address some of the attention and comparison dynamics during school hours without addressing the broader ecosystem within which students are embedded. Digital citizenship curricula teach some relevant skills without adequately addressing the architectural features of the platforms that make those skills difficult to apply consistently.

There is no simple answer. But there is a better-informed practice: one grounded in honest engagement with what the platforms are actually designed to do, what the evidence shows about effects on developing minds, and what the realistic range of protective interventions is in the current regulatory and technological environment.

What This Means for Policymakers

The litigation against Meta is one mechanism through which accountability for social media's effects on children is being pursued. It is not the only mechanism, and it is unlikely to be adequate on its own.

Legislative and regulatory approaches to social media and children are developing across multiple jurisdictions: age verification requirements, restrictions on features specifically associated with harm (algorithmic amplification of eating disorder content, infinite scroll for minors), transparency requirements for internal research on youth effects, and liability frameworks that create meaningful incentives for platform safety investment rather than safety theater.

None of these approaches is simple. Age verification at scale raises privacy concerns. Feature restrictions require regulatory technical expertise that legislative bodies are still developing. Transparency requirements are only as useful as the capacity to evaluate what they reveal. And the

platforms' global scale means that jurisdiction-specific regulation has limited reach.

But the alternative (continuing to treat social media platform safety for children primarily as a matter of platform self-governance and parental management) has produced the documented outcomes that forty-one state attorneys general characterized as sufficient grounds for litigation. The case for regulatory intervention is not made by the litigation alone. It is made by the body of internal platform research, independent academic research, and clinical observation that the litigation has made more fully visible.

The Communications Failure Underneath the Safety Failure

There is a communications dimension to the Meta litigation that has not received adequate attention: the gap between what platforms know and what they say about youth safety is not just a legal and ethical problem. It is an institutional trust problem with consequences that extend far beyond the specific litigation.

Platform safety communications (the voluntary initiatives, the parental control features, the youth safety pledges, the congressional testimony) were received by parents, educators, and policymakers as good-faith representations of platform practice. They were used to make decisions about children's access to platforms. And if the states' central allegations are accurate, they were not adequate representations of what the platforms' own research showed.

The consequence is not just legal liability for the platforms. It is the further erosion of institutional trust in technology companies as honest actors in public discourse: trust that was already under severe strain and that, once lost, is extraordinarily difficult to rebuild.

For communications professionals working in technology, in public health, in youth services, and in any field where institutional credibility matters: the Meta litigation is a case study in the cost of the gap between institutional knowledge and institutional communication. That gap (wherever it exists, in whatever institution, about whatever issue) is the communications problem that eventually produces the accountability reckoning. The question is not whether it will produce one. It is only when, and at what cost.

Where This Fits

The gap between internal knowledge and public communication is the pattern that destroys institutional trust everywhere, not only in technology. The Intersection of Trust and Communication article explains why, and the ERA™ white paper describes what repair requires once a public has learned not to believe you.

Schedule a conversation with Links Global: www.linksglobal.org

Sources

- Wells, G., Horwitz, J., & Seetharaman, D. (2021, September 14). Facebook knows Instagram is toxic for teen girls, company documents show. *The Wall Street Journal*.
- Office of the U.S. Surgeon General. (2023). *Social media and youth mental health: The U.S. Surgeon General's advisory*. hhs.gov
- California Department of Justice. (2026, August 26). *Attorney General Bonta secures transformative $17 billion settlement with Meta* [Press release]. oag.ca.gov

Ownership and use. © 2026 Links Global. All rights reserved. The Links Global Trust Audit™, Narrative Stewardship Infrastructure™, NSI™, Epistemic Re-Regulation Architecture™, and ERA™ are trademarks of Links Global. You are welcome to read this article, link to it, share it, and quote brief passages with attribution to Links Global and a link to this page. Reproducing it in whole or in substantial part, adapting or translating it, or using the analysis and framing presented here in training, consulting, publications, or derivative works requires prior written permission from Links Global. Please cite as: Sanchez, M. (2026). *Social Media and Children: What the Meta Lawsuits Tell Us That the Platforms Won't*. Links Global. www.linksglobal.org.