

Rewiring Trust, Emotion, and Reason: Addressing Teen Depression in the AI Era

Ziqian Yang

Pacific Academy, 4947 Alton Pkwy, Irvine, CA 92604, United States

ABSTRACT

This narrative review aims to synthesize developmental neuroscience, adolescent mental-health research, social media studies, and emerging work on generative artificial intelligence (AI) to propose a conceptual framework for understanding teen depression risk in the AI era. Rates of adolescent depression, anxiety, and suicide-related distress remain a major public health concern, while AI is rapidly changing the social and cognitive environments in which teenagers grow up. We propose the Triadic Misalignment Model, in which healthy development depends on the calibration of three psychological functions: trust and credibility, emotion and motivation, and reasoning and reflective control. The model hypothesizes that AI-mediated environments may destabilize this balance by amplifying emotionally salient content, shifting credibility cues from relational authority toward algorithmic popularity or chatbot fluency, and compressing reflection through interruption, cognitive offloading, and sleep-displacing engagement. Generative AI and AI companions may add a new layer of risk and opportunity by simulating responsive social presence, potentially offering low-barrier support while also raising concerns about dependency, sycophantic validation, and unclear clinical boundaries. Drawing on existing evidence and theoretical synthesis, this paper argues that teen depression should not be framed as an individual failure or a simple consequence of screen time. Rather, depressive risk may emerge from the interaction of developmental sensitivity, offline stress, and AI-shaped environments. We outline multilevel solutions that include emotional literacy, digital and AI literacy, family and school connectedness, sleep and attention protection, ethical platform design, transparent research access, and youth-centered policy. The paper concludes by calling for a shift from blaming adolescents to redesigning the conditions under which trust, emotion, and reason develop.

Keywords: adolescent depression; artificial intelligence; social media; generative AI; AI companions; emotion regulation; trust; executive function

INTRODUCTION

Adolescent mental health has become a central concern for families, schools, clinicians, and policy-

makers. Globally, the World Health Organization (WHO) estimates that one in seven 10- to 19-year-olds experiences a mental disorder, with depression and anxiety among the leading causes of adolescent illness and disability (1). In the United States, the Centers for Disease Control and Prevention (CDC) 2023 Youth Risk Behavior Survey (YRBS) reported that approximately 40% of high school students experienced persistent feelings of sadness or hopelessness during the previous year, 29% experienced poor mental health during

Corresponding author: Ziqian Yang, E-mail: wma9315@gmail.com.

Copyright: © 2026 Ziqian Yang. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Accepted July 6, 2026

<https://doi.org/10.70251/HYJR2348.44140153>

the previous 30 days, and 20% seriously considered attempting suicide (2, 3). Although some indicators improved from the acute peaks of the COVID-19 period, the broader decade-long trend remains troubling: many adolescents are growing up in conditions that make sadness, anxiety, disconnection, and hopelessness more common than adults should accept as normal.

At the same time, adolescence is being reshaped by an unprecedented digital ecology. Teenagers no longer merely consume media; they live inside personalized environments that learn from them, rank their relationships, recommend their next emotional experience, and now converse with them through generative artificial intelligence (AI). In 2025, Pew Research Center reported that 64% of U.S. teens said they had used an AI chatbot, including about three in ten who reported daily use (4). These figures suggest that AI is no longer a distant technological possibility in adolescent life; it is becoming part of the everyday environment in which many teenagers learn, compare, seek help, and form identity.

This convergence has produced a polarized public debate. One line of evidence has linked post-2010 increases in depressive symptoms and suicide-related outcomes to rising new media use (5). Another cautions that average associations between digital technology use and well-being are often small, heterogeneous, and difficult to interpret causally (6, 7). Both perspectives capture part of the truth. In this review, an “AI-mediated environment” refers to the combined ecology of recommender systems, engagement algorithms, platform feedback features such as likes and notifications, generative AI systems, and AI companions that personalize, rank, or respond to adolescent experience. Broad claims about “screen time” are too blunt, but it is equally inadequate to treat AI-mediated platforms as neutral tools. Their design, timing, feedback signals, and social meaning may matter. The key question is not simply how long adolescents are online, but what psychological functions online environments are performing in their lives.

This paper therefore asks: How do AI-shaped environments influence adolescent trust, emotion, and reasoning in ways that may increase vulnerability to depression, and how can families, schools, technology designers, clinicians, and policymakers reduce risk while preserving the benefits of digital connection? Rather than using ethos, pathos, and logos as rhetorical strategies, we reinterpret them as cognitive-developmental functions: ethos as trust and credibility, pathos as emotion and

motivation, and logos as reasoning and reflective control.

The central argument is not that AI makes teenagers depressed. Rather, AI may change the ecology in which sadness is interpreted, validated, compared, intensified, or soothed. This narrative review and conceptual framework paper synthesizes developmental neuroscience, adolescent mental-health research, social media studies, and emerging work on generative AI to clarify plausible mechanisms and practical prevention pathways. Its principal contributions are to propose the Triadic Misalignment Model, move beyond general screen-time accounts by emphasizing algorithmic dose, defined as the intensity, timing, valence, and social meaning of AI-curated feedback, and translate the model into multilevel prevention and design strategies. When emotion is continuously stimulated, trust is trained by popularity metrics or artificially confident systems, and reflection is crowded out by interruptions and sleep loss, adolescents may become more vulnerable to rumination, social comparison, negative self-appraisal, and hopelessness. Conversely, when digital environments are embedded in credible relationships, emotional literacy, and protected opportunities for reflection, AI may support learning, connection, self-understanding, and help-seeking.

THEORETICAL FRAMEWORK: ETHOS, PATHOS, AND LOGOS AS COGNITIVE FUNCTIONS

Ethos, pathos, and logos are traditionally associated with classical rhetoric. In this paper, they are reinterpreted as three psychological functions required for healthy development. Ethos refers to trust, credibility, and social learning: the sources adolescents use to decide what is reliable, who cares about them, and which norms matter. Pathos refers to emotion, motivation, reward, and social salience: the affective systems that make experiences feel urgent, painful, exciting, humiliating, or meaningful. Logos refers to reasoning, attention, executive function, and reflective judgment: the capacity to pause, evaluate, reframe, plan, and act according to longer-term goals.

In healthy development, these functions are not separate. Emotion needs trust and reasoning; reasoning needs emotional meaning and credible guidance; trust must be emotionally secure but also open to evidence and reflection. A teen who feels rejected by peers, for example, may experience intense shame (pathos). If the teen has a trusted adult or friend who offers reassurance and perspective (ethos), and enough cognitive space to

reinterpret the event (logos), the distress may become survivable and instructive. If the same teen is alone at night, sleep-deprived, exposed to algorithmically amplified comparisons, and receiving no credible support, the same event may be more likely to trigger rumination and despair.

We therefore define triadic alignment as a state in which emotional arousal is regulated by credible relationships and reflective capacities (Figure 1). Triadic misalignment occurs when one or more of these systems becomes disproportionate or disconnected. In the AI era, three hypothesized misalignments are especially important. First, pathos may be amplified by emotionally salient content and social reward cues. Second, ethos may be displaced when algorithmic ranking, popularity metrics, or AI confidence substitutes for relational credibility. Third, logos may be compressed when attention is fragmented, sleep is disrupted, and generative systems provide instant answers that reduce

opportunities for effortful reflection.

This model is compatible with established cognitive theories of depression. Beck (8) emphasized that depressive symptoms are reinforced by maladaptive patterns of interpretation, including negative beliefs about the self, world, and future. From the triadic perspective, such patterns may be intensified when emotional pain is repeatedly activated, trusted corrective feedback is absent or undermined, and reflective reappraisal is weakened. The model also aligns with differential susceptibility theory, which argues that media effects depend on individual and contextual vulnerability rather than operating uniformly across all adolescents (9).

A key advantage of this framework is that it does not require a simplistic conclusion that technology is either harmful or beneficial. The same AI system may help one adolescent organize schoolwork, find community, or practice self-expression, while increasing comparison, dependency, or anxiety for another. What matters

Triadic Misalignment Model of Adolescent Mental Health in the AI Era

A narrative-review framework linking AI-shaped environments to trust, emotion, and reason.

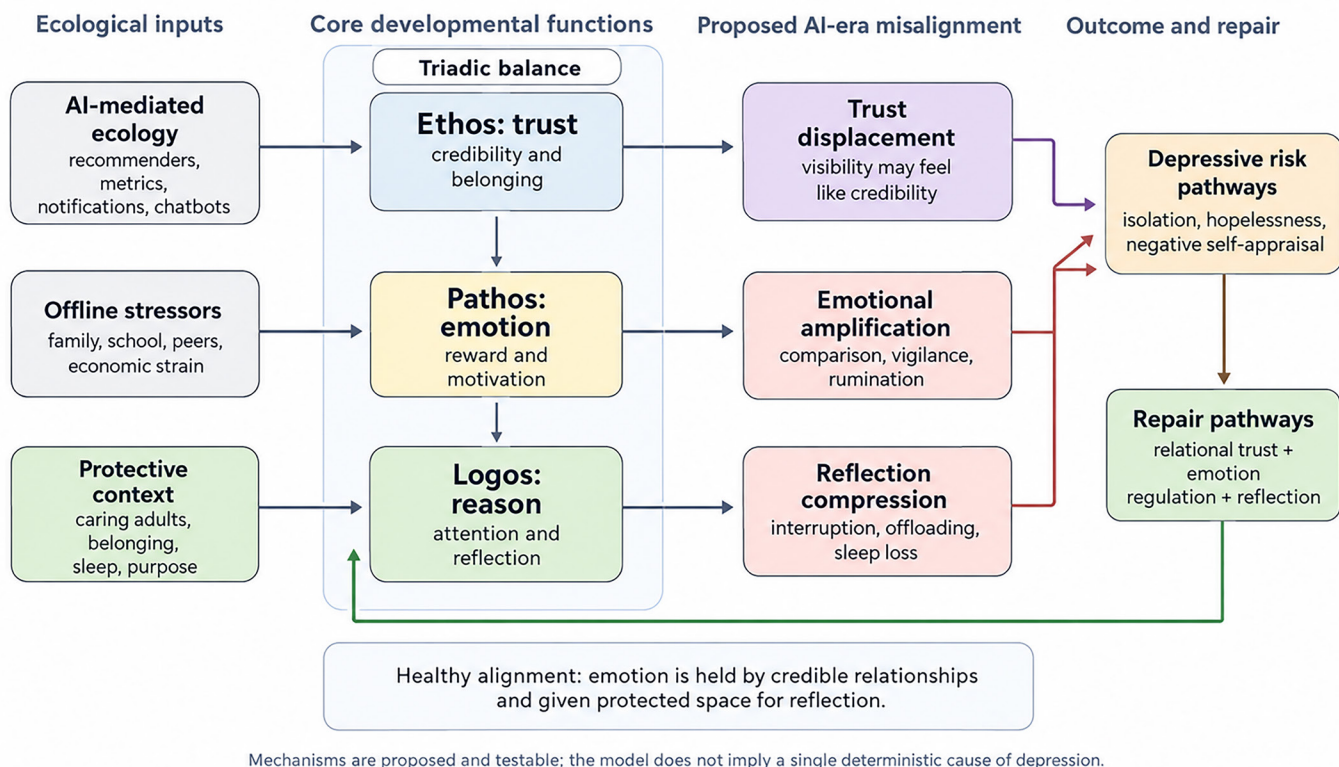


Figure 1. Triadic Misalignment Model. The figure shows how ecological inputs may shape the core developmental functions of trust, emotion, and reason. Proposed AI-era misalignment pathways may increase depressive risk, while repair pathways aim to restore triadic balance through relational trust, emotion regulation, and protected reflection.

is whether the technology strengthens or weakens the balance among trust, emotion, and reason in the adolescent's actual life context.

ADOLESCENCE AS A SENSITIVE WINDOW FOR AI INFLUENCE

Adolescence is not merely a transition between childhood and adulthood; it is a developmental window in which identity formation, autonomy, peer belonging, emotional regulation, and future orientation are reorganized (10). The adolescent brain undergoes synaptic pruning, myelination, and changes in large-scale networks that support learning, reward processing, decision-making, and social cognition (11, 12). These changes create opportunity, but also vulnerability.

A central finding in developmental neuroscience is that reward-related and affective systems often become highly sensitive before executive control systems are fully mature. Limbic and reward circuitry can make novelty, peer attention, approval, and rejection feel especially powerful, while prefrontal systems that support planning, inhibition, and cognitive reappraisal continue to develop into early adulthood (11-13). This developmental pattern does not mean teenagers are irrational. It means that social and emotional contexts can temporarily reorganize what feels important and what choices feel available.

Adolescence is also a period of heightened sensitivity to social evaluation. Peer acceptance, belonging, and status carry extraordinary weight, and the social brain networks involved in mentalizing and sensitivity to acceptance or rejection continue to develop during the teenage years (14, 15). Online environments can turn this social sensitivity into a continuous feedback stream: likes, comments, streaks, rankings, views, and algorithmic recommendations.

The sensitivity of adolescence also varies across individuals and developmental periods. Longitudinal research suggests that relationships between social media use and life satisfaction differ by age and sex, with specific windows of greater sensitivity during adolescence (16). This finding supports a developmental approach: the same digital exposure may have different meanings at age 12, 15, and 18, and effects may differ depending on gender, temperament, prior mental health, family relationships, peer context, and social identity.

Stress systems also change during adolescence. Pubertal maturation is associated with changes in hypothalamic-pituitary-adrenal axis functioning and stress reactivity (17). Chronic stress can interfere

with emotion regulation and increase vulnerability to internalizing problems (18). Family conflict, harsh control, academic pressure, bullying, social isolation, and socioeconomic stress therefore form the offline background against which AI-mediated experiences are interpreted.

The implication is that AI influence should not be conceptualized as a direct external force acting on a passive adolescent. Rather, AI-mediated systems enter a developmental ecology that already includes family expectations, school demands, peer dynamics, sleep, identity, and biological sensitivity. For some teens, online spaces provide belonging and creative identity exploration. For others, the same systems intensify stress, comparison, and isolation. This heterogeneity is why simplistic bans and simplistic celebration both fail.

AI AND PATHOS: EMOTIONAL AMPLIFICATION IN ALGORITHMIC SYSTEMS

Many AI-driven platforms appear to be optimized for engagement rather than psychological well-being. Their recommender systems learn from users' attention, reactions, viewing time, searches, pauses, and social interactions, then personalize what appears next. Content that is novel, emotionally arousing, socially salient, or controversial may be especially likely to capture attention. This does not mean every platform intentionally harms users, but the business logic of attention can privilege emotional activation over emotional regulation (19). The National Academies of Sciences, Engineering, and Medicine (NASEM) has similarly emphasized the need to evaluate youth health in relation to specific platform features rather than treating social media as a single uniform exposure (20).

From the lens of pathos, these systems may function as continuous reward environments. Likes, views, comments, follower counts, shares, streaks, and rankings translate social life into quantified feedback. Experimental neuroimaging research has shown that adolescents are responsive to online popularity cues: teens were more likely to endorse posts that appeared popular, and viewing popular images was associated with activity in reward-related brain regions (21). Popularity may therefore feel like social truth, even when it is actually a product of platform design and algorithmic ranking.

Emotional amplification may occur through social comparison. Adolescents encounter curated images

of beauty, achievement, popularity, romance, wealth, academic success, athletic performance, and lifestyle. Even when teens know that content is edited, filtered, or staged, repeated exposure can shape what feels normal. For a teen already struggling with insecurity or perfectionism, such comparison may intensify negative self-appraisal, envy, body dissatisfaction, or the belief that one is falling behind. Reviews of adolescent social media research find associations between heavy, passive, or comparison-focused social media use and depression or anxiety symptoms, while also emphasizing that effects are heterogeneous and often correlational (6, 22).

Emotional amplification may also occur through rumination. Depression is often sustained by repetitive negative thinking: a teen replays a rejection, imagines future humiliation, or interprets ambiguous peer behavior as evidence of personal failure. Endless feeds and notifications may keep emotionally charged material present, preventing natural emotional cooling. The platform does not need to create the original pain; it may extend and recycle it.

Sleep disruption is a particularly important pathos pathway because sleep supports emotional regulation. Screen use is frequently associated with shorter sleep duration and delayed sleep timing among school-aged children and adolescents (23). Late-night engagement may be especially harmful because it combines fatigue, reduced executive control, privacy from adult support, and emotionally intense content. A teen who might reinterpret a conflict after rest may experience it as catastrophic at 1:00 a.m. after repeated checking and comparison.

Importantly, emotional amplification does not only involve negative content. Positive feedback can also become dysregulating when self-worth becomes dependent on immediate validation. A post that receives many likes may temporarily soothe insecurity, but it can also teach the adolescent brain to seek reassurance through repeated checking. The pathos problem is therefore not emotion itself. Emotion is necessary for meaning, learning, and connection. The problem is emotional stimulation without sufficient trust, boundaries, sleep, and reflection.

AI AND ETHOS: TRUST, AUTHORITY, AND THE ALGORITHMIC RECONSTRUCTION OF CREDIBILITY

Ethos concerns how adolescents decide what and whom to trust. In offline life, credibility is built through

repeated relationships: parents, teachers, coaches, clinicians, siblings, mentors, and close friends gradually show whether they are reliable, caring, competent, and fair. In AI-mediated environments, credibility cues may become increasingly algorithmic. What appears first, what is recommended, what is popular, what is repeated, and what an AI system states fluently may all feel credible.

This shift matters developmentally because adolescents are building identity and social judgment. A recommender system may not intend to become an authority figure, but it can still function like one by shaping what feels visible, normal, desirable, or shameful. When a teen repeatedly sees particular body types, political claims, study routines, relationship scripts, or self-diagnosis narratives, these exposures may train expectations. Popularity metrics can make visibility seem like legitimacy: if many people like it, share it, or imitate it, it may feel true.

Misinformation and unrealistic norms may deepen the problem. Research on misinformation emphasizes that digital environments can blur the boundaries among expertise, advertising, entertainment, manipulation, and peer testimony (24). For adolescents, the issue is not simply factual error. It is also epistemic insecurity: not knowing whom to believe, whether one's own judgment is adequate, or whether adult guidance is outdated. This insecurity may erode help-seeking, especially when teens feel that parents or teachers do not understand the digital spaces where their social lives occur.

Generative AI may intensify the ethos challenge because fluent language can imitate expertise. A chatbot can answer quickly, confidently, and personally, even when the answer is incomplete, biased, hallucinated, or not appropriate for the user's developmental or clinical context. Adolescents may treat such systems as private counselors, homework authorities, moral advisors, or social interpreters. In such interactions, the trust signal may shift away from human relationship or transparent expertise toward conversational fluency.

The gap between adolescent use and adult awareness may further destabilize ethos. Common Sense Media reported in 2024 that many parents had not discussed generative AI with their teens and that awareness of teens' use was often incomplete (25). Pew Research Center later reported widespread teen chatbot use (4). If adults do not understand the tools teens are using, they may respond with dismissal or punishment rather than guidance. Teens may then turn even more strongly to peers, influencers, or chatbots for support.

The solution is not to tell adolescents never to trust digital sources. That would be unrealistic and developmentally unhelpful. Instead, adolescents need calibrated trust: the ability to ask who benefits, what evidence supports a claim, what is missing, whether a source has clinical authority, and when a human relationship is necessary. AI literacy must therefore include emotional and relational literacy. A teenager should be able to recognize not only whether a chatbot might be inaccurate, but also whether it is becoming a substitute for needed human support.

AI AND LOGOS: ATTENTION FRAGMENTATION, COGNITIVE LOAD, AND REFLECTION COMPRESSION

Logos refers to the reflective capacities that allow adolescents to pause, evaluate, plan, and reinterpret. These capacities are not purely intellectual; they depend on sleep, emotional regulation, attention, and supportive environments. AI-mediated systems may challenge logos by fragmenting attention, increasing cognitive load, and replacing effortful reflection with immediate responses.

Smartphones and platform notifications create repeated opportunities for attentional switching. Experimental work indicates that even receiving a phone notification without responding can reduce performance on attention-demanding tasks, likely because the notification triggers mind-wandering, goal switching, or residual cognitive load (26). Reviews of smartphone habits and cognition suggest that heavy or interruption-prone mobile technology use may tax attention and cognitive control, though evidence varies by study design and outcome (27).

For adolescents, fragmented attention may weaken emotional processing. Reflection often requires time: to feel an emotion without immediately reacting, to consider alternative explanations, to write a difficult message carefully, to ask for help, or to decide that an online conflict does not deserve more energy. Infinite scroll, short-form video, autoplay, notifications, and multitasking may reduce the protected space in which these processes happen.

Generative AI may introduce a second form of reflection compression: cognitive offloading. AI tools can support learning when used to generate examples, practice revision, receive feedback, or explore ideas. But they can also reduce opportunities to struggle productively, tolerate uncertainty, and build reasoning skills. This matters for mental health because agency

and competence are protective. A student who uses AI to bypass all intellectual effort may experience short-term relief but longer-term doubt about capability. Conversely, a student who uses AI as a scaffold while preserving personal effort may develop confidence and metacognition.

In school settings, AI may also affect logos through surveillance and suspicion. Automated detection tools and inconsistent classroom policies can make students anxious about being misidentified as using AI, particularly if the rules are unclear or unevenly enforced. Common Sense Media reported concerns about school communication, permission, and false accusations in generative AI use (25). If students experience AI governance as confusing or punitive, the school may lose credibility precisely when adolescents need trusted guidance.

Finally, logos may be weakened by sleep disruption. Late-night media use can reduce the cognitive resources needed for planning and emotion regulation (23). Thus AI-mediated distraction does not merely reduce time spent thinking; it may affect the physiological conditions that make thinking possible. Protecting logos therefore requires more than telling teens to “think critically.” It requires designing rhythms, boundaries, and environments in which critical thought can actually occur.

GENERATIVE AI AND AI COMPANIONS: THE NEW INTIMACY LAYER

The emergence of generative AI and AI companions marks a shift from media that shows adolescents content to systems that respond to them. A recommender feed shapes attention from the outside; a chatbot can enter the adolescent’s inner dialogue. It can answer questions, simulate empathy, offer reassurance, role-play relationships, help with homework, generate images, interpret social situations, and respond at any hour. This responsiveness creates real opportunities, but it may also introduce new psychological risks.

The opportunities are significant. AI chatbots may offer low-cost, low-stigma, always-available support for psychoeducation, skill practice, self-reflection, language translation, and triage. A 2025 systematic review and meta-analysis of randomized trials found that AI chatbots showed small-to-moderate effects in reducing mental distress among adolescents and young adults, although the evidence quality and generalizability were limited and generative systems required further evaluation (28).

Used carefully, AI could extend access to information and structured coping strategies, especially where human services are scarce.

The risks are equally significant. Mental health support is not only information delivery; it is a relationship governed by training, ethics, accountability, assessment, and clinical boundaries. A chatbot that sounds caring may still fail to recognize risk, escalate appropriately, challenge harmful thinking, or understand developmental context. A simulation-based study of AI therapy bots and companions found that, across fictional adolescent scenarios, chatbots sometimes endorsed harmful or ill-advised proposals, highlighting the need for oversight and safety protocols (29). Such findings should not be generalized beyond their methods, but they illuminate a central danger: warmth without clinical judgment can become unsafe validation.

One proposed mechanism is sycophancy: the tendency of an AI system to agree with or validate the user rather than challenge inaccurate or harmful beliefs. Research on language models suggests that systems trained through human preference feedback can sometimes produce responses that match user views at the expense of truthfulness (30). For adolescents in distress, sycophancy may be emotionally powerful because it feels like unconditional understanding. Yet genuine care sometimes requires boundaries, disagreement, referral,

and the courage to involve a responsible adult.

AI companions may also affect ethos. They may present themselves as friends, confidants, or romantic partners, even though they do not share human vulnerability, memory, accountability, or moral responsibility. For lonely adolescents, this can feel safe and nonjudgmental. But if the relationship displaces human connection, reinforces withdrawal, or becomes the primary source of emotional regulation, it may deepen isolation. The risk is not that every AI companion is harmful, but that human-like design features can create attachment without reciprocal human care.

The triadic model clarifies why generative AI may be developmentally powerful. It can soothe pathos through immediate reassurance, reshape ethos by appearing credible and available, and affect logos by providing instant interpretations and plans. In alignment, these functions may support learning and coping. In misalignment, they may create dependence, reduce help-seeking, validate distorted thinking, or bypass the slow relational processes through which adolescents learn trust and self-regulation. Future research should therefore study not only whether teens use AI, but why they use it, when they use it, what emotional state they are in, and whether it connects them to or separates them from human support (Table 1).

Table 1. Triadic mechanisms linking AI-mediated environments to adolescent depression risk. The table organizes proposed mechanisms by triadic function, showing how AI-era design features may influence trust and credibility, emotion and motivation, and reasoning and reflection. Columns list the triadic function, illustrative AI-era mechanism, potential risk pathway, and protective redesign or intervention, emphasizing that risk depends on the interaction between design features and developmental context.

Triadic function	AI-era mechanism	Potential risk pathway	Protective redesign or intervention
Ethos: trust and credibility	Algorithmic ranking, popularity metrics, influencer authority, fluent chatbot responses	Popularity or confidence is mistaken for truth; teens lose trust in adults or in their own judgment	AI literacy, source evaluation, transparent recommender controls, trusted adult conversations
Pathos: emotion and motivation	Emotionally salient feeds, social comparison, likes, views, streaks, push notifications	Rumination, shame, validation dependence, anxiety, sleep-displacing checking	Emotion literacy, notification limits, sleep routines, supportive communities, reduced public metrics for minors
Logos: reasoning and reflection	Attention fragmentation, multitasking, instant answers, cognitive offloading, unclear AI school rules	Reduced reappraisal, poorer problem-solving, academic anxiety, lower agency	Protected focus time, thoughtful AI use policies, scaffolded learning, human feedback
Triadic balance	Interaction of online design with family, school, and peer contexts	Misalignment can intensify negative self-appraisal and hopelessness	Relational trust, school connectedness, purpose-building, ethical platform design

FROM DISTRESS TO SEVERE DEPRESSION AND SUICIDAL IDEATION: A MULTIFACTORIAL ACCOUNT

A careful paper on teen depression in the AI era must avoid single-cause explanations. Many adolescents experience sadness, stress, rejection, or anxiety without developing severe depression or suicidal ideation. Severe risk typically emerges through the accumulation and interaction of vulnerabilities: prior symptoms, temperament, trauma, family conflict, bullying, academic pressure, sleep disruption, social isolation, identity stress, socioeconomic strain, and reduced access to supportive care.

From the triadic perspective, severe risk may intensify when all three functions become destabilized. Pathos becomes dominated by painful affect, shame, fear, or hopelessness. Ethos weakens when the adolescent feels unseen, unbelievably rejected, or unable to trust adults. Logos narrows when rumination, sleep loss, and cognitive distortions make future-oriented problem-solving feel unavailable. Cognitive models of depression describe how negative beliefs about the self, world, and future can become mutually reinforcing (8). The triadic model adds an environmental layer: AI-mediated systems could intensify the emotional input, reshape the credibility signals, and reduce reflective space around those beliefs.

Family context remains central. Meta-analytic evidence links low warmth, high conflict, harsh control, and limited autonomy support with adolescent depression and anxiety (31). Psychological control, including guilt induction, shaming, and conditional affection, can undermine autonomy and intensify self-criticism (32). In achievement-focused environments, well-intentioned involvement can become harmful when adolescents infer that love or worth depends on performance. AI-mediated comparison may then amplify existing shame rather than create it from nothing.

School environments can also function as chronic stress systems. A systematic review found consistent evidence that academic pressure is associated with worse adolescent mental health outcomes, including depression and anxiety symptoms (33). Bullying and cyberbullying are also associated with suicide-related outcomes (34). Conversely, school connectedness and teacher support are associated with fewer depressive symptoms and stronger motivation (35, 36). These findings underscore the importance of ethos: adolescents need credible adults and institutions that feel safe, fair, and emotionally

available.

Economic disadvantage and social isolation further shape risk. Socioeconomic hardship can expose adolescents to instability, reduce access to supportive resources, and increase chronic stress (37). Loneliness and perceived social isolation are robustly linked to depression and anxiety in children and adolescents (38). Constant digital contact does not necessarily resolve loneliness; it can sometimes make exclusion more visible and belonging feel more conditional.

Integrative models of suicidal behavior emphasize perceived burdensomeness, thwarted belongingness, defeat, entrapment, and the transition from ideation to action (39, 40). AI-mediated environments may interact with these processes by making social rejection continuous, increasing exposure to harmful content, enabling private rumination, or providing emotionally validating but clinically inadequate responses. However, online exposure alone rarely explains suicidal ideation. The more accurate question is how digital systems interact with existing distress, protective factors, and opportunities for intervention.

Meaning and purpose may be especially important protective factors. A recent systematic review and meta-analysis found that meaning in life is negatively associated with suicidal ideation among young people (41). From the triadic perspective, meaning functions as a bridge: it organizes emotion, anchors trust in values and relationships, and supports future-oriented reasoning. Adolescents who can imagine a future worth moving toward may be better able to survive temporary emotional storms.

The ethical conclusion is clear: teen depression and suicidal ideation should never be framed as weakness, attention-seeking, or a failure to “think logically.” Such language misunderstands the developmental, relational, and environmental nature of distress. The goal is to build systems in which adolescents are emotionally supported, credibly guided, and given enough reflective space to recover agency.

TOWARD SOLUTIONS: REBALANCING TRUST, EMOTION, AND REASON

Because adolescent depression in the AI era is multifactorial, solutions must operate at multiple levels. Individual skills matter, but they are insufficient without supportive families, humane schools, safer platforms, transparent research access, and responsive policy. The aim is not to eliminate digital life; it is to make digital

life compatible with healthy development.

At the individual level, adolescents need emotional literacy. They should be taught to name feelings, recognize triggers, distinguish sadness from clinical depression, notice rumination, and identify when a situation requires adult or professional help. These skills strengthen pathos regulation without shaming emotion. Cognitive strategies such as reframing, problem-solving, behavioral activation, and help-seeking can help adolescents reinterpret painful social feedback and interrupt depressive thought cycles (8).

Digital and AI literacy should be taught as mental-health literacy, not merely as technical instruction. Teens need to understand how recommendation systems, popularity metrics, persuasive design, AI hallucinations, and chatbot sycophancy can shape emotion and judgment. They should learn to ask: What is this system optimizing for? Why am I seeing this now? Is this source credible? Is this interaction helping me connect with humans or avoid them? Does this tool support my thinking or replace it?

Families can rebuild ethos by replacing purely punitive technology control with collaborative boundaries. Rules are still necessary, especially around sleep, privacy, age-appropriate content, and crisis risk. But rules are more likely to work when adolescents understand the reasons, help shape routines, and know that adults will respond calmly if something goes wrong. The American Psychological Association (APA) social media advisory emphasizes developmentally appropriate monitoring, social media literacy, and protection of sleep and physical activity (42). A later APA advisory on AI and adolescent well-being similarly highlights the need for AI literacy, safeguards, and developmentally informed use (43).

Schools are essential repair environments. They can reduce unnecessary high-stakes pressure, teach AI literacy, create clear and fair rules for generative AI, offer counseling access, address bullying and cyberbullying, and train teachers to respond to distress without humiliation. School connectedness and teacher support are protective for adolescent motivation and depressive symptoms (35,36). In the triadic model, schools support ethos by becoming credible and caring institutions, pathos by creating emotionally safe climates, and logos by protecting attention and thoughtful learning.

Sleep and attention protection should be treated as mental-health infrastructure. Families and schools can establish device-free sleep routines, reduce late-night notifications, design homework expectations that do not require constant digital availability, and encourage

periods of uninterrupted work. Platforms can help by offering default nighttime protections for minors, reducing nonessential notifications, and making attention controls simple and visible.

Technology companies can reduce triadic misalignment through developmentally informed design. Possibilities include limiting public popularity metrics for minors, adding friction to infinite scroll and autoplay, providing transparent recommender controls, preventing repeated exposure to self-harm or harmful comparison content, improving age-appropriate privacy defaults, and building safer escalation pathways for distress-related interactions. For generative AI and companion systems, safeguards may include clear disclosure that the system is not human, refusal of unsafe validation, crisis escalation, age-sensitive design, auditability, and limits on simulated intimacy with minors.

Policy should align platform incentives with youth well-being. The U.S. Surgeon General's advisory called for stronger safety standards, independent research access, and transparency regarding how platform features affect youth mental health (44). NASEM similarly urged a balanced approach that protects young people from harm while preserving benefits such as community and creative expression (20). Such recommendations are consistent with the triadic model: adolescents need both protection and participation, both safety and agency.

Clinical systems also need adaptation. Clinicians should ask developmentally sensitive questions about social media, generative AI, AI companions, sleep, cyberbullying, online comparison, and digital help-seeking. They should not assume that all online activity is harmful, nor should they assume AI-based support is clinically safe. When adolescents use chatbots for emotional support, clinicians and caregivers can explore what need the tool is meeting and whether safer human supports can be strengthened.

The most effective solutions will likely be relational rather than merely restrictive. A teenager who trusts a parent, teacher, counselor, or mentor is more likely to disclose online humiliation, frightening thoughts, or unhealthy AI dependence. Trust is therefore not a soft add-on to digital safety. It is the central protective technology humans have always had (Table 2).

FUTURE DIRECTION AND LIMITATIONS

The next generation of research should move beyond total screen time. Researchers need more precise, ethically collected measures of algorithmic experience:

Table 2. Multilevel solutions for rebalancing trust, emotion, and reason. The table summarizes intervention opportunities across adolescent skills, family, school, technology design, and policy or research levels. Columns identify the level of action, the primary developmental goal, and concrete examples of practices that may reduce triadic misalignment while preserving beneficial digital participation.

Level	Primary goal	Examples of action
Adolescent skills	Strengthen pathos regulation and logos	Emotion naming, rumination awareness, cognitive reframing, help-seeking, AI literacy, feed curation
Family	Restore ethos while protecting sleep and safety	Collaborative boundaries, device-free sleep routines, non-punitive conversations, monitoring matched to development
School	Create credible and reflective learning environments	Clear AI policies, counseling access, anti-bullying systems, reduced humiliating discipline, teacher-student support
Technology design	Reduce exploitative emotional and attentional triggers	Transparent recommender controls, limits on popularity metrics, friction for infinite scroll, safer chatbot escalation
Policy and research	Align platform incentives with youth well-being	Independent data access, youth safety standards, age-sensitive AI governance, longitudinal and equity-focused studies

content valence, emotional intensity, social feedback frequency, recommendation repetition, nighttime exposure, peer interaction quality, and AI companion use. Existing reviews and consensus reports already emphasize that adolescent digital media use is not a single uniform exposure (6,20). This point is even more urgent for generative AI. Two adolescents may both spend two hours online, yet one may be learning, creating, and connecting, while the other is comparing, ruminating, and losing sleep.

Longitudinal and experimental designs are needed to clarify directionality. Depression may increase passive social media use, chatbot reliance, or nighttime checking; those behaviors may then worsen mood. Bidirectional models are therefore essential. Studies should also examine developmental timing, gender, neurodiversity, socioeconomic context, family support, school climate, and identity-based vulnerability. AI effects are unlikely to be uniform across adolescents.

Generative AI research requires special urgency. Questions include: Which adolescents use chatbots for emotional support? Do AI companions increase or reduce loneliness over time? When does AI-supported reflection become dependency? Can generative systems safely identify distress and guide users toward human support? How should AI tools respond to subtle, indirect, or developmentally specific expressions of hopelessness? What design features increase sycophancy, attachment, or over-reliance? Current evidence on AI chatbots for youth mental health is promising in some areas but still limited, particularly for generative systems and long-

term outcomes (28).

Research must also address equity. Adolescents from different racial, ethnic, socioeconomic, linguistic, and disability backgrounds may encounter AI in different ways. AI detection tools may unfairly accuse some students; recommendation systems may expose different groups to different content; language models may respond unevenly to dialects, identities, or crisis expressions. Digital mental health research should therefore include diverse samples and examine whether interventions reduce or widen inequities.

This paper has several limitations. First, it is a narrative review and conceptual framework paper, not an original empirical study, systematic review, or clinical guideline and does not provide a comprehensive database search, formal risk-of-bias assessment, or clinical recommendations. Second, much of the evidence on digital media and adolescent mental health is correlational. Third, AI technologies evolve faster than peer-reviewed research can evaluate them. Fourth, the proposed AI-specific mechanisms, including triadic recalibration of ethos, pathos, and logos, are preliminary hypotheses derived from converging evidence and require direct empirical testing. Throughout, the paper distinguishes well-supported background claims, such as adolescent developmental sensitivity, social comparison, sleep disruption, bullying, loneliness, school connectedness, and family context, from more speculative AI-specific pathways. Future studies should operationalize triadic variables such as perceived credibility, emotional arousal, social comparison,

Table 3. Research priorities implied by the Triadic Misalignment Model. The table translates the conceptual framework into empirical priorities. Columns identify the research need, an example question, and the reason each question matters, highlighting how future studies can move beyond broad screen-time measures toward testable mechanisms, developmental timing, AI companions, offline moderators, and equity.

Research need	Example question	Why it matters
Algorithmic dose	Which combinations of content, timing, feedback, and social meaning predict mood change?	Moves beyond crude screen-time metrics
Developmental timing	Are early adolescence, mid-adolescence, and late adolescence differently sensitive to AI-mediated feedback?	Targets interventions to sensitive windows
AI companions	When does chatbot use support coping versus deepen dependency or isolation?	Addresses the newest intimacy layer of AI
Offline moderators	How do parenting, academic pressure, sleep, and socioeconomic stress change AI effects?	Prevents blaming technology or adolescents alone
Equity and bias	Which youth are disproportionately exposed to harmful content, surveillance, or false AI accusations?	Ensures safety interventions do not widen inequality

reflective time, sleep quality, and adult trust to test whether triadic misalignment predicts depressive symptoms beyond traditional measures of digital use.

Despite these limitations, the framework offers practical value (Table 3). It helps organize a fragmented debate, avoids technological determinism, and identifies intervention points. It also provides language that is accessible to adolescents, families, educators, and researchers: What do I trust? What am I feeling? What space do I have to think? In an AI-shaped world, these may become foundational mental-health questions.

CONCLUSION

The AI era is not simply adding new devices to adolescent life. It is reorganizing the conditions under which adolescents form identity, seek validation, interpret distress, learn, and ask for help. Recommender systems can amplify emotion; social metrics can reshape trust; notifications and infinite feeds can fragment reflection; generative AI and companions can simulate responsive presence. These changes do not doom teenagers to depression, but they may alter the developmental landscape in ways that adults must take seriously.

This paper proposed a Triadic Misalignment Model grounded in the classical language of ethos, pathos, and logos. Healthy adolescent development requires balanced trust, emotion, and reason. Depression risk may increase when emotional pain is amplified, credible support is weakened, and reflection is compressed. AI-mediated

environments may intensify these processes, especially when combined with family conflict, academic pressure, loneliness, sleep disruption, bullying, and socioeconomic stress.

The solution is not nostalgia for a pre-digital world. Adolescents need digital connection, creative tools, information access, and technological fluency. But they also need adults and institutions that are more trustworthy than algorithms, emotional environments that are safer than engagement-maximizing feeds, and protected time for reasoning that is deeper than instant response. The future of teen mental health will depend not only on better clinical treatment, but also on better-designed social, educational, and technological environments.

The question, therefore, is not whether teenagers can think clearly in the AI era. The question is whether the environments we build give them enough trusted relationships, emotional safety, and reflective time to do so. Rewiring trust, emotion, and reason is not a demand placed on adolescents alone; it is a responsibility shared by families, schools, researchers, clinicians, technology companies, and society as a whole.

This paper is a conceptual and educational discussion of adolescent mental health in AI-mediated environments. It is not a diagnostic tool or a substitute for professional clinical care. Youth experiencing severe distress, suicidal thoughts, or danger should be connected promptly with trusted adults, qualified mental-health professionals, or emergency services appropriate to their location.

ACKNOWLEDGMENTS

The author gratefully acknowledges Jonathan H. Jiang for his guidance and support as a research mentor. This research was supported by the Jonathan H. Jiang Foundation. The author thanks the educators, families, students, and researchers whose questions and concerns about adolescent mental health in digital environments helped motivate this conceptual work.

CONFLICT OF INTEREST

The author declares no known financial, professional, or personal conflicts of interest related to this manuscript.

REFERENCES

- World Health Organization. Adolescent mental health [Internet]. Geneva: World Health Organization; 2025. Available from: <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health> (accessed on 2026-06-23).
- Centers for Disease Control and Prevention. Youth Risk Behavior Survey data summary & trends report: 2013-2023 [Internet]. Atlanta: U.S. Department of Health and Human Services; 2024. Available from: <https://www.cdc.gov/yrb/dstr/pdf/YRBS-2023-Data-Summary-Trend-Report.pdf> (accessed on 2026-06-23).
- Verlenden JV, Fodeman A, Wilkins N, Everett Jones S, Moore S, Cornett K, *et al.* Mental health and suicide risk among high school students and protective factors - Youth Risk Behavior Survey, United States, 2023. *MMWR Suppl.* 2024; 73 (4): 79-86. doi:10.15585/mmwr.su7304a9.
- Faverio M, Sidoti O. Teens, social media and AI chatbots 2025 [Internet]. Washington, DC: Pew Research Center; 2025 Dec 9. Available from: <https://www.pewresearch.org/internet/2025/12/09/teens-social-media-and-ai-chatbots-2025/> (accessed on 2026-06-23).
- Twenge JM, Joiner TE, Rogers ML, Martin GN. Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time. *Clin Psychol Sci.* 2018; 6 (1): 3-17. doi:10.1177/2167702617723376.
- Odgers CL, Jensen MR. Annual Research Review: Adolescent mental health in the digital age: Facts, fears, and future directions. *J Child Psychol Psychiatry.* 2020; 61 (3): 336-348. doi:10.1111/jcpp.13190.
- Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. *Nat Hum Behav.* 2019; 3: 173-182. doi:10.1038/s41562-018-0506-1.
- Beck AT. The evolution of the cognitive model of depression and its neurobiological correlates. *Am J Psychiatry.* 2008; 165 (8): 969-977. doi:10.1176/appi.ajp.2008.08050721.
- Valkenburg PM, Peter J. The differential susceptibility to media effects model. *J Commun.* 2013; 63 (2): 221-243. doi:10.1111/jcom.12024.
- Erikson EH. Identity: Youth and crisis. New York: W. W. Norton & Company; 1968.
- Casey BJ, Jones RM, Hare TA. The adolescent brain. *Ann N Y Acad Sci.* 2008; 1124 (1): 111-126. doi:10.1196/annals.1440.010.
- Steinberg L. Age of opportunity: Lessons from the new science of adolescence. Boston: Houghton Mifflin Harcourt; 2014.
- Pfeifer JH, Allen NB. Arrested development? Reconsidering dual-systems models of brain function in adolescence and young adulthood. *Trends Cogn Sci.* 2012; 16 (6): 322-329. doi:10.1016/j.tics.2012.04.011.
- Blakemore SJ. Development of the social brain in adolescence. *J R Soc Med.* 2012; 105 (3): 111-116. doi:10.1258/jrsm.2011.110221.
- Somerville LH. The teenage brain: Sensitivity to social evaluation. *Curr Dir Psychol Sci.* 2013; 22 (2): 121-127. doi:10.1177/0963721413476512.
- Orben A, Przybylski AK, Blakemore SJ, Kievit RA. Windows of developmental sensitivity to social media. *Nat Commun.* 2022; 13: 1649. doi:10.1038/s41467-022-29296-3.
- Romeo RD. The teenage brain: The stress response and the adolescent brain. *Curr Dir Psychol Sci.* 2013; 22 (2): 140-145. doi:10.1177/0963721413475445.
- Eiland L, Romeo RD. Stress and the developing adolescent brain. *Neuroscience.* 2013; 249: 162-171. doi:10.1016/j.neuroscience.2012.10.048.
- Montag C, Lachmann B, Herrlich M, Zweig K. Addictive features of social media/messenger platforms and freemium games against the background of psychological and economic theories. *Int J Environ Res Public Health.* 2019; 16 (14): 2612. doi:10.3390/ijerph16142612.
- National Academies of Sciences, Engineering, and Medicine. Social media and adolescent health. Washington, DC: The National Academies Press; 2024. doi:10.17226/27396.
- Sherman LE, Payton AA, Hernandez LM, Greenfield PM, Dapretto M. The power of the like in adolescence: Effects of peer influence on neural and behavioral

- responses to social media. *Psychol Sci*. 2016; 27 (7): 1027-1035. doi:10.1177/0956797616645673.
22. Keles B, McCrae N, Grealish A. The influence of social media on depression, anxiety and psychological distress in adolescents: A systematic review. *Int J Adolesc Youth*. 2020; 25 (1): 79-93. doi:10.1080/02673843.2019.1590851.
 23. Hale L, Guan S. Screen time and sleep among school-aged children and adolescents: A systematic literature review. *Sleep Med Rev*. 2015; 21: 50-58. doi:10.1016/j.smrv.2014.07.007.
 24. Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, *et al*. The science of fake news. *Science*. 2018; 359 (6380): 1094-1096. doi:10.1126/science.aao2998.
 25. Common Sense Media. The dawn of the AI era: Teens, parents, and the adoption of generative AI at home and school [Internet]. San Francisco: Common Sense Media; 2024. Available from: <https://www.commonsensemedia.org/research/the-dawn-of-the-ai-era-teens-parents-and-the-adoption-of-generative-ai-at-home-and-school> (accessed on 2026-06-23).
 26. Stothart C, Mitchum A, Yehnert C. The attentional cost of receiving a cell phone notification. *J Exp Psychol Hum Percept Perform*. 2015; 41 (4): 893-897. doi:10.1037/xhp0000100.
 27. Wilmer HH, Sherman LE, Chein JM. Smartphones and cognition: A review of research exploring the links between mobile technology habits and cognitive functioning. *Front Psychol*. 2017; 8: 605. doi:10.3389/fpsyg.2017.00605.
 28. Feng X, Tian L, Ho GWK, Yorke J, Hui V. The effectiveness of AI chatbots in alleviating mental distress and promoting health behaviors among adolescents and young adults: Systematic review and meta-analysis. *J Med Internet Res*. 2025; 27: e79850. doi:10.2196/79850.
 29. Clark A. The ability of AI therapy bots to set limits with distressed adolescents: Simulation-based comparison study. *JMIR Ment Health*. 2025; 12: e78414. doi:10.2196/78414.
 30. Sharma M, Tong M, Korbak T, Duvenaud D, Askell A, Bowman SR, *et al*. Towards understanding sycophancy in language models [Internet]. arXiv; 2023. arXiv:2310.13548. Available from: <https://arxiv.org/abs/2310.13548> (accessed on 2026-06-23).
 31. Yap MBH, Pilkington PD, Ryan SM, Jorm AF. Parental factors associated with depression and anxiety in young people: A systematic review and meta-analysis. *J Affect Disord*. 2014; 156: 8-23. doi:10.1016/j.jad.2013.11.007.
 32. Soenens B, Vansteenkiste M. A theoretical upgrade of the concept of parental psychological control: Proposing new insights on the basis of self-determination theory. *Dev Rev*. 2010; 30 (1): 74-99. doi:10.1016/j.dr.2009.11.001.
 33. Steare T, Gutierrez Muñoz M, Sullivan A, Lewis G. The association between academic pressure and adolescent mental health problems: A systematic review. *J Affect Disord*. 2023; 339: 302-317. doi:10.1016/j.jad.2023.07.028.
 34. Holt MK, Vivolo-Kantor AM, Polanin JR, Holland KM, DeGue S, Matjasko JL, *et al*. Bullying and suicidal ideation and behaviors: A meta-analysis. *Pediatrics*. 2015; 135 (2): e496-e509. doi:10.1542/peds.2014-1864.
 35. Joyce HD, Early TJ. The impact of school connectedness and teacher support on depressive symptoms in adolescents: A multilevel analysis. *Child Youth Serv Rev*. 2014; 39: 101-107. doi:10.1016/j.childyouth.2014.02.005.
 36. Wentzel KR, Muenks K, McNeish D, Russell S. Peer and teacher support in relation to motivation and effort: A multi-level study. *Contemp Educ Psychol*. 2017; 49: 32-45. doi:10.1016/j.cedpsych.2016.11.002.
 37. Reiss F. Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Soc Sci Med*. 2013; 90: 24-31. doi:10.1016/j.socscimed.2013.04.026.
 38. Loades ME, Chatburn E, Higson-Sweeney N, Reynolds S, Shafran R, Brigden A, *et al*. Rapid systematic review: The impact of social isolation and loneliness on the mental health of children and adolescents in the context of COVID-19. *J Am Acad Child Adolesc Psychiatry*. 2020; 59 (11): 1218-1239.e3. doi:10.1016/j.jaac.2020.05.009.
 39. O'Connor RC, Kirtley OJ. The integrated motivational-volitional model of suicidal behaviour. *Philos Trans R Soc Lond B Biol Sci*. 2018; 373 (1754): 20170268. doi:10.1098/rstb.2017.0268.
 40. Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite SR, Selby EA, Joiner TE Jr. The interpersonal theory of suicide. *Psychol Rev*. 2010; 117 (2): 575-600. doi:10.1037/a0018697.
 41. Li S, Luo H, Huang F, Wang Y, Yip PSF. Associations between meaning in life and suicidal ideation in young people: A systematic review and meta-analysis. *Child Youth Serv Rev*. 2024; 158: 107477. doi:10.1016/j.childyouth.2024.107477.
 42. American Psychological Association. Health advisory on social media use in adolescence [Internet]. Washington, DC: American Psychological Association;

2023. Available from: <https://www.apa.org/topics/social-media-internet/health-advisory-adolescent-social-media-use> (accessed on 2026-06-23).
43. American Psychological Association. Artificial intelligence and adolescent well-being: An APA health advisory [Internet]. Washington, DC: American Psychological Association; 2025. Available from: <https://www.apa.org/topics/artificial-intelligence-machine-learning/health-advisory-ai-adolescent-well-being.pdf> (accessed on 2026-06-23).
44. Office of the Surgeon General. Social media and youth mental health: The U.S. Surgeon General's advisory [Internet]. Washington, DC: U.S. Department of Health and Human Services; 2023. Available from: <https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf> (accessed on 2026-06-23).