The Medium Becomes The Self: The Clinical Framework for Algorithmic Identity

By Kayleigh Robertson Hons. BA., CPS

York University, Faculty of Health

December 1, 2025

Abstract: Over the last decade, mental health hospitalizations among young people (particularly Generation Z and the emerging Generation Alpha, born after 2012) have surged, with growing evidence linking this rise to problematic smartphone, social media, and now AI relational or companion use. Canadian data shows significant increases in hospital admissions for eating disorders, self-harm, and anxiety during high-risk periods such as the COVID-19 pandemic (Roumeliotis et al., 2024). Concurrently, research implicates the structure of digital platforms themselves in exacerbating depression, anxiety, disordered eating, and identity disturbance in Gen Z; appearance-driven platforms like TikTok and Instagram intensify social comparison, FOMO, compulsive self-monitoring, and cyberbullying (Shehab et al., 2025). As Gen Z has aged within this crisis, while Gen Alpha enters the same crisis a decade later, rates of underemployment, debt, emotional dysregulation, and overall life dissatisfaction continue to rise.

Despite this now well-documented developmental emergency, mainstream mental health care models have not meaningfully adapted. Cognitive Behavioral Therapy (CBT), Dialectical Behavior Therapy (DBT), and standard psychiatric assessments continue to overlook digital behaviours, algorithmic feedback loops, and AI-mediated interactions, both for Gen Alpha currently in crisis and for Gen Z living in its aftermath. From a media-ecology perspective, smartphones and social media are not neutral tools but environmental forces that reshape cognition, social perception, emotional regulation, and identity formation. Persistent engagement with interactive, appearance-focused platforms emerge as the Fourth Person (Robertson, 2025): a digital identity that exists alongside, but psychologically separate from, the offline self (the core self, the social self and the reflective self). This fragmentation amplifies anxiety, compulsive posting, regret, low self-worth, and interpersonal instability, yet it remains absent from clinical assessment frameworks.

Emerging evidence further shows that AI “cyber-companion” systems cause psychosis, and/or intensify identity fragmentation by reinforcing the emotional, perceptual, and cognitive needs of the Fourth Person. These dynamics reveal a profound gap in current mental health care: digital and AI-mediated behaviours are not lifestyle preferences but core mechanisms of contemporary psychopathology. This thesis proposes a comprehensive, multi-level adaptation of mental health care that systematically integrates smartphone, social media, and AI use into assessment, diagnosis, and therapeutic intervention. Without such integration, the system will continue failing the very populations most harmed by the environments they were raised in; Gen Z after the crisis, and Gen Alpha entering the same crisis in real time.

Keywords: regret, fourth person, AI, cyber companion, social media, psychosis, psychopathology, intervention, crisis, perception, solution

Where Media Ecology and Psychology Meet: Clarifying The Clinical Pathway

Marshall McLuhan's seminal concept, *"the medium is the message,"* (1964) asserts that it is the form of media itself, not merely its content, that shapes human cognition, perception, and social organization. In other words, media does more than convey information; it actively reshapes the structure of the mind, influences identity formation, and alters ways of being in the world. This thesis extends McLuhan's insight into the contemporary digital landscape, arguing that social media platforms and algorithmically driven AI environments function as potent agents in constructing new psychological realities. Specifically, these digital media forms can contribute to psychosis and the development of mediated identity disturbances by embedding users in persistent algorithmically driven feedback loops that restructure perception, self-concept, and relational experience. In the age of algorithmic influence, this new understanding of the medium itself as a clinical and psychological force provides the foundation for reimagining therapeutic interventions and recognizing it is a problem and a consistent trigger.

Conceptualizing The Medium Becoming The Self: The Fourth Person

In my research, I have developed the concept of the *Fourth Person* to describe a layer of the self that emerges in response to prolonged engagement with digital and social media environments.

Conceptually, the Fourth Person coexists with, and can sometimes override, the offline or “natural” self, influencing emotion, cognition, behaviour, and interpersonal relationships. It can be thought of as a “fourth layer” atop existing structures: the core self, the social self, and the reflective self. Whereas the core self embodies enduring traits and internalized values, the social self reflects interpersonal roles and relational expectations, and the reflective self monitors personal behaviour and thought patterns, the Fourth Person is primarily constructed and maintained through digital interaction. It is both relational and performative, emerging from algorithmic pressures and culturally embedded expectations of visibility and validation. This digital layer sits not only above psychological models of identity, but also above psychoanalytic and cognitive frameworks, reshaping them. Think of it like a three-tiered cake: the core self, the social self, and the reflective self. The digital layer is the icing. But this icing never stays on top; it inevitably seeps into the layers, altering their shape, texture, and even the flavour of the whole cake.

It becomes an added layer on top of the id, ego, and superego (Freud, 1923), three systems that naturally regulate instinct, negotiation with reality, and moral evaluation. Similarly, it can be understood in relation to Higgins’ (1987) self-discrepancy theory, which distinguishes between the actual, ideal, and ought selves. Social media stretches and distorts these existing layers and selves, amplifying idealized and externally imposed standards until they exceed the bounds of internal self-regulation. Through repeated exposure to digital platforms that encourage performance, comparison, and constant self-monitoring, the Fourth Person emerges: a digitally sculpted identity that

exists alongside (but increasingly apart from) the psychological structures that once grounded the self.

Unlike traditional models of personality, which focus on stable traits or the interaction of conscious and unconscious processes, the Fourth Person represents the digitally constructed self: a semi-autonomous identity shaped by algorithmic feedback, social validation metrics, and curated self-presentation.

Users increasingly experience “perceptions of perceptions,” in which they not only interpret how others see them, but also anticipate how others evaluate their online persona through likes, comments, followers, and other quantifiable feedback. Over time, this digital layer begins to regulate offline thought and behaviour, creating a feedback loop that reinforces performance-based identity and magnifies emotional sensitivity to social evaluation.

This framework has significant implications for understanding mental health. Adolescents and emerging adults are particularly vulnerable, as their identities are still forming and their cognitive-emotional systems remain highly plastic. The Fourth Person can contribute to identity fragmentation, emotional dysregulation, compulsive comparison, and maladaptive social behaviour. At its extreme, this process can intersect with the onset of serious mental health conditions, including first-episode psychosis and personality disturbance, as offline and online selves diverge.

The Fourth Person framework helps explain why many members of Gen Z are experiencing heightened difficulties as they transition into adulthood. Years of growing up under constant algorithmic feedback, social media performance pressures, and

perpetual social comparison have shaped a digital self that often dominates offline cognition, emotion, and behaviour. This persistent mediation of the self can create chronic identity fragmentation, emotional dysregulation, and difficulty navigating real-world relationships, challenges that extend beyond adolescence into emerging adulthood. Unlike previous generations, Gen Z has spent formative developmental years in environments that amplify self-monitoring, validation-seeking, and perceived social evaluation, producing vulnerabilities in mental health, career development, and interpersonal functioning. By highlighting the ways digital environments have constructed a Fourth Person layer, the theory demonstrates that these struggles are not simply individual failings but are structural and generational, rooted in the interaction between human development and pervasive digital media.

By articulating the Fourth Person, my goal is to provide a conceptual lens that explains how social media transforms identity and self-concept in contemporary youth. This framework is intended not only to guide empirical research but also to inform therapeutic and preventative strategies that address digitally mediated identity formation, helping clinicians, educators, and families understand the interplay between online environments and mental health.

Don’t Just Read The Headline

To accurately situate the argument of this thesis, it is necessary to clarify what is meant by psychosis in the context of social media and AI and equally important, what is *not* being claimed. Psychosis is not simply “stress,” nor is it synonymous with suicidality. Psychosis refers to disturbances in perception, thinking, and reality-testing:

hallucinations, delusions, disorganized thought, paranoia, derealization, and identity instability.

Traditionally, psychotic episodes arise from a complex intersection of biological vulnerability, trauma, developmental stress, and environmental triggers. What has shifted over the last decade is the *environment*. For the first time in human history, people (particularly youth) are immersed in 24/7 algorithmic systems that can intensify stress responses, skew reality perception through curated feeds, reinforce cognitive distortions, and generate delusional material with unprecedented speed. Social media algorithms amplify themes of identity, comparison, and threat; generative AI can mirror or escalate the user's thinking; and both mediums create a context in which the boundaries of self and reality can be destabilized. For vulnerable youth, particularly Gen Z and Gen Alpha, these platforms do not merely influence mood, they can act as direct triggers for psychotic symptoms or accelerate the trajectory toward a psychotic break.

At the same time, this paper does *not* claim that every young person experiencing suicidality, self-harm, or crisis is suffering from smartphone-induced psychosis, nor that every case belongs in digital-specialty style therapy. The clinical reality is far more nuanced. In emergency departments, people present first with suicide attempts, suicidal ideation, or acute self-harm. In these scenarios, the immediate focus is stabilization: safety first, life first. Only after those life-threatening behaviours are addressed does assessment turn to contributing factors, and what should be included is smartphone use. Among these patients, by integrating protocols for smartphone use, some will screen positive for digital triggers, particularly those whose distress involves obsessive

posting, online harassment, delusional interpretations of digital content, parasocial entanglements, or AI-reinforced cognitive distortions. Others will not. The intent of the paper is not to pathologize every crisis as “caused by the phone,” but to correctly identify when the medium *is* playing a determinative role in the patient’s deterioration.

Equally critical is the outpatient and therapeutic population, and youth who are not necessarily arriving in emergency rooms, but who are nevertheless suffering profoundly from digital-driven psychological symptoms. These are the young people already engaged in counselling or community mental-health support structures whose symptoms worsen in direct proportion to their smartphone use. Their anxiety, depression, dissociation, compulsive social comparison, distorted self-concept, and early psychotic features are reinforced by online environments that saturate their cognition. For these clients, the smartphone is not incidental; it is the mechanism through which the disorder is maintained. These are the cases where the Fourth Person becomes clinically relevant, where standard therapeutic techniques fail because the problem is not just within the patient but embedded in the structure and demands of the medium they use daily. Interventions must specifically target digital identity fragmentation, compulsive online behaviours, algorithmic reinforcement, and the collapse of offline self-coherence.

This distinction matters: *the interventions proposed in this paper are for the subset of individuals (particularly youth) whose symptoms are shaped, amplified, or directly triggered by digital media environments.* Not every suicidal young person needs a

smartphone-focused intervention, but many do, far more than our current systems acknowledge. And without clearly naming and treating the digital mechanisms that destabilize identity and cognition, we will continue to fail who this crisis is swallowing. The purpose of integrating media ecology and clinical psychology is precisely this: to correctly identify when the medium is the message *and* the self, so therapeutic care can connect with the reality of the people who need it.

A Mental Health Crisis

Gen Z and emerging Gen Alpha represent the first cohorts of true digital natives, having grown up immersed in smartphones, social media platforms, and instantaneous digital communication. For these young people, online engagement is not merely a pastime but a central component of identity formation, socialization, and self-expression. Unlike previous generations, whose offline and online selves were more clearly delineated, Gen Z and Gen Alpha navigate interactive and appearance-focused platforms such as TikTok, Instagram, and Snapchat, where social feedback is rapid, pervasive, and often judgmental. This constant digital engagement has created novel vulnerabilities, including heightened social comparison, fear of missing out on online interactions or “digital FOMO,” and compulsive self-monitoring, all of which are increasingly linked to anxiety, depression, disordered eating, and emerging identity disturbances.

Distinguishing the lived reality of Gen Z and the maturing of Gen Alpha is crucial for understanding why these generations face such elevated psychological risk, and why targeted intervention is not merely recommended but necessary for their long-term development. Having grown up entirely online, these youth have been exposed to

problematic trends at unprecedented speeds, immersed in cultures of instant gratification, impulsive self-presentation, and algorithmically reinforced purchasing and posting behaviours. They also engage in constant comparisons to curated, often unrealistic portrayals of others' lives. As a result, many young people struggle with underemployment due to distorted expectations of what work "should" look like, aspire to influencer lifestyles that are unattainable for the vast majority, or accrue debt attempting to maintain an online aesthetic and lifestyle. They curate a perception of a perception: what they believe others will see, based on what they see in others, and they attempt to believe it is themselves.

These dynamics contribute to pervasive feelings of inadequacy, abandonment of realistic goals due to fear of online judgment, and a widespread sense of being lost or directionless. Many remain at home longer than previous generations, feel emotionally underdeveloped due to both isolation and the constant exposure to dysregulated online behaviour, and report chronic low self-worth. In this environment, conditions such as eating disorders, identity fragmentation, and now AI-related perceptual disturbances are increasingly observed as part of the digital-developmental landscape.

A particularly salient phenomenon in this context is the Fourth Person: a digital identity layer that exists alongside, but psychologically separate from, the offline self, shaped by social media interactions, perceived audience, and platform dynamics (Robertson, 2025). The Fourth Person represents a fragmented identity layer that can intensify compulsive posting, regret, emotional volatility, low self-worth, and social anxiety, creating a distinct target for therapeutic intervention. It is the fourth person which makes the individual believe the perception of perception is themselves, realigning values,

ideals, goals and outward social cues to fit the mark of the fourth person's desired perception.

We can further understand this through extreme-trend-like behaviours, plastic surgery on the rise, eating disorder behaviour for male and female genders on the rise, and high levels of debt rooted in "desirable lifestyle goals" at mere age ranges of 16-27.

To put it simply, things many young individuals wouldn't have done had TikTok and Instagram "not told them to", although, the translation of this psychologically is that their fourth person, and perception of perception, was driven to perform such things.

Unlike offline behaviors, which can be addressed through conventional cognitive or behavioral strategies, the Fourth Person's actions are mediated by complex online environments that standard therapeutic models have yet to fully incorporate.

It Is Not Normal, And They Need Therapy

Despite widespread media attention and scholarly commentary on the mental health challenges of Gen Z, much of this discourse remains largely rhetorical. Documentaries, opinion pieces, and public commentary dating back to 2020 often highlight the risks of digital engagement, asking questions such as "Who are you becoming as a result of social media and algorithms?" as if it is provocative, yet simultaneously emphasizing feelings of disconnection, shame, or anxiety. These discussions rarely propose actionable solutions, and are ironic in nature to Gen Z's very real changing identity based on social media use and abuse. Furthermore, extreme experiences and first-person accounts are frequently dismissed as "too much" or ignored as individual

and rare, leaving the voices of those most affected largely unheard. This discussion perpetuates a cycle in which data of social media issues (showcasing suicide, self harm, and anxiety) is acknowledged but not substantively addressed, which undermines efforts to develop clinical interventions informed by lived experience. Such data emphasize the need for mental health research and care models that center authentic experiences of digital identity fragmentation, rather than relying solely on external observation or theoretical critique. The rhetoric is dismissing it.

When mainstream discourse dismisses or minimizes a person's experiences, suggesting that those struggling are "fine", will "grow out of it" or that they have "made this up", it obscures the severity of the underlying issues. Many of these behaviors are not fabricated, but rather reflect the complex interplay between social media, identity fragmentation, and the emergence of the Fourth Person. For example, young girls have developed functional tics after exposure to TikTok content (Pringshiem et al., 2021) (a very real concern psychologically) or some adolescents believe they have autism only because the platform suggested it. While these individuals may not meet the diagnostic criteria for such conditions, the mental health concern at large is real, and often rooted in the pressures and validation cycles of their digital identities. These manifestations underscore that structured clinical support is not optional, but essential. Recognizing these experiences as clinically significant rather than trivial or performative is essential for designing therapeutic interventions that address the true sources of psychological harm. They are not fine.

Articles and essays frequently discuss digital vulnerability as performative and state that emotional over-sharing is inherently problematic, dismissing lived experiences of young

people as “background noise”. While these perspectives do highlight concerns about online behaviours, they largely ignore the structural and psychological pressures of constant social media engagement, that people are behaving this way, *why people are behaving this way,* and the level of disconnection they must feel. They fail to account for the emergence of a digital persona that fragments identity, amplifies compulsive behaviors, and intensifies anxiety, depression, and regret through posting online or performing behaviours psychotic in nature: The Fourth Person. By framing real psychological experiences as trivial or performative, such discourse discourages meaningful solutions and overlooks the urgent need for clinical frameworks that address identity fragmentation, digital regret, and compulsive online behaviors. It is clinical Fourth Person performance on their social media accounts, but there is a lack of recognition that this is in fact a clinical issue.

In addition, cultural commentators increasingly critique therapy as overthinking or a marketized indulgence (India, 2023), and this rhetoric is dangerously misleading for Gen Z and emerging Gen Alpha. Despite having no clinical education or lived experience with severe mental health crises, these “influencers” of rhetoric and commentary on Gen Z present therapy as a performative or unnecessary exercise, echoing the same digital pressures that social media exerts on youth. This messaging minimizes real psychological distress, ignores evidence of rising suicidality and psychosis, and risks discouraging adolescents from seeking the support they need. By framing mental health care as a problem of self-focus rather than a life-saving intervention, such discourse mirrors the performative validation loops of social media itself; it tells young people how to feel, what to value, and whose voices matter,

silencing those experiencing genuine crisis. For youth navigating the Fourth Person and the fragmented realities of digital identity, in addition to life-changing decisions and social media-based choices, dismissing therapy is not a critique of over-reflection, it is an active (and completely avoidable) barrier to care at a moment when intervention could save lives. The commentary has forgotten that along with anxiety, depression, and the maladaptive behaviours we are seeing online, loneliness and suicidality is higher than ever for Gen Z,. In 2021, we called what was happening to Gen Z a mental health crisis. Today, *they're just thinking too much about themselves.* (India, 2025). They actually *do* need clinical support. The rhetoric is ignoring it.

Developmentally Delayed Adulthood

Without proper clinical support, Generation Z is facing the cascading consequences of a childhood and adolescence spent under relentless digital exposure. The social media and smartphone ecosystems that shaped them prioritized perception over authenticity, validation over self-compassion, and virality over developmental readiness. This has bred a cohort marked by isolation, emotional dysregulation, and profound difficulty forming and sustaining intimate relationships; lower rates of sexual activity (Fry, 2023) and meaningful romantic engagement reflect the substitution of digital validation for real-world connection. What is interesting in regards to sexual activity, specifically from 2020, is the shift from risky sexual behaviours (Dhaona et al., 2020) to now not performing any at all. This showcases the isolation, developmental delays, and now ultimately, regret that has formed as a basis from environmental function as adolescents partaking in extreme behaviours. Economically, these same patterns have contributed to heightened personal debt and increasing personal bankruptcies, unstable

employment, and compulsive consumption driven by fast fashion trends, influencer culture, and the constant push to perform and project worth online. As young people attempt to reconcile their online personas with offline selves, psychologically, the omnipresent pressure to be hyper-visible, likable, and unique has amplified rates of anxiety, depression, and maladaptive coping strategies, including substance abuse, body modification, and plastic surgery. The cumulative effect is a generation navigating adulthood with diminished resilience, less joy, and a fractured sense of self, all while digital environments continue to reward compulsive engagement with hits of dopamine and reinforced self-objectification and relentless comparison scrolling. These residual effects are not merely cultural or economic, they are distinctly clinical. They illustrate how unaddressed digital trauma and identity fragmentation have created a mental health landscape where traditional therapeutic models fall short, and where intervention is not optional but lifesaving. Gen Z's struggles are the direct inheritance of social media's structural incentives, and without systemic, Fourth Person-aware clinical support, the cycle of harm continues, threatening the well-being, autonomy, and potential of an entire generation within communities, economies and future global landscapes.

The Clinical Support They Need

Cognitive Behavioral Therapy (CBT) is a structured, evidence-based psychotherapeutic approach that focuses on identifying and modifying maladaptive thought patterns, beliefs, and behaviors that contribute to psychological distress. Developed in the 1960s and 1970s, CBT emphasizes the interplay between cognition, emotion, and behavior, helping patients recognize distorted thinking, test beliefs against reality, and implement

practical strategies for change (Beck, 1976). It has been widely applied to anxiety, depression, obsessive-compulsive disorder, and other mental health conditions, with strong empirical support for its effectiveness across diverse populations.

Dialectical Behavior Therapy (DBT), developed by Marsha Linehan in the late 1980s, emerged specifically as a treatment for individuals with borderline personality disorder, a population characterized by intense emotional dysregulation, impulsivity, and self-harm behaviors. DBT builds on CBT principles but adds a focus on acceptance, mindfulness, distress tolerance, and interpersonal effectiveness. Its dual approach, emphasizing both change and acceptance, allows patients to develop coping strategies for managing intense emotions while validating their lived experience. DBT has since been adapted to address suicidal ideation, self-harm, substance use, and emotion-driven maladaptive behaviors in adolescents and adults.

Together, CBT and DBT provide complementary frameworks for understanding and addressing both cognitive distortions and emotional dysregulation, offering structured interventions for a wide range of psychological challenges.

However, both were designed before the rise of pervasive digital engagement, and their standard protocols often assume a unified offline self, limiting their effectiveness for individuals navigating identity fragmentation and the pressures of the Fourth Person (online identity, compulsive posting, social comparison, digital FOMO, digital regret).

Gen Z and Gen Alpha patients may struggle to reconcile therapeutic guidance with their lived realities, particularly when the behaviors and experiences mediated by the Fourth Person (the algorithmic self, the identity that is the individual online) remain

unaddressed. Integrating these novel digital considerations into both DBT and CBT, along with assessment, inpatient protocols, and outpatient interventions represents a critical and timely evolution of mental health care.

Hearing Their Narrative

A common experience among survivors of social media-induced psychosis is the tendency to describe their past disordered behaviors (restrictive eating, suicidal ideation, manic patterns, and compulsive online actions) as something done by “them,” a past self, or a different version of who they once were. In traditional therapy, this linguistic distancing is often treated as a red flag indicating avoidance, denial, or a lack of ownership. Clinicians typically intervene by redirecting clients from “they/he/she” statements back to “I,” emphasizing responsibility, empowerment, and cognitive accuracy. However, within a digital environment that fosters identity fragmentation, this instinctive distancing reflects something far more complex and clinically significant. As individuals return to a psychological baseline, many correctly recognize that the behaviors were not expressions of their integrated offline self, but of a digitally constructed identity (the Fourth Person) that developed in response to social media pressures. Rather than signaling denial, this narrative separation is often an accurate reflection of how their identity was split and manipulated online. It was not them. And the individuals do not see it as them, either.

Therapists, clinicians, and physicians understanding the significance of the digital self can begin to recognize cause and effect rather than dismiss digital identity as resistance to traditional therapy. Recognizing the Fourth Person as a distinct, digitally mediated

identity allows processing regret more safely and effectively: it helps them understand that the harmful behaviors emerged from an imposed digital self, not from their authentic identity. This reframing not only validates their lived experience but also supports emotional closure, reduces shame, and clarifies the role of social media and AI in distorting self-perception. By acknowledging this identity division, clinical care can more accurately address the psychological realities of young people shaped by pervasive digital influence.

Regret, Fourth Person Style

Regret is emerging as a defining concern for Gen Z now entering their 20's. Many report profound distress over behaviors or choices made during adolescence; missing out on offline experiences, spending money on lifestyles driven by social media expectations, and performing identities shaped by platform dynamics rather than personal volition. Traditional DBT approaches emphasizing radical acceptance minimize this lived experience and emphasize regret as something the individual "hasn't gotten over". Instead, therapy must validate the reality of these experiences, allowing patients to recognize that the "sick" or "performative" self, their "Fourth Person", is distinct from the offline self, and regret is to be expected in recovery.

Integrating regret explicitly into CBT and DBT interventions requires acknowledging it as a clinically significant experience. Therapy can focus on understanding the choices made by the Fourth Person, exploring their emotional impact, and developing strategies to reconcile these experiences without invalidation of one's offline identity. This approach moves beyond mere radical acceptance; it validates the reality of misaligned

behaviors, supports emotional processing, and fosters a sense of agency over who one really is. By addressing regret directly, mental health care can meet Gen Z where they are, supporting recovery and self-integration in a landscape shaped by digital engagement, social feedback, and identity fragmentation.

CBT's emphasis on identifying and restructuring distorted cognitions remains valuable but requires adaptation to account for online triggers. For example, thoughts such as "I am worthless because nobody liked my post" or "I must post perfectly to be accepted" are not merely cognitive distortions. They emerge from an entirely mediated social environment. Similarly, DBT skills such as distress tolerance, radical acceptance, and emotion regulation can be effective for Fourth Person-related anxiety and compulsive behaviors, but only if therapy explicitly acknowledges and works with digital identity fragmentation. Without integrating the Fourth Person into treatment, patients may find themselves applying evidence-based techniques to experiences that are partially externalized in the online sphere or even fictional, limiting their efficacy.

Bob Newhart as The Therapist - "Stop It!"

> *Therapist - Tell me about the problem that you wish to address.*
>
> *Patient - Well I have this fear of being buried alive in a box, I just, I start thinking about being buried alive and I begin to panic, thinking about it really does make my life horrible I mean I can't go through tunnels or be in an elevator or in a house, anything boxy.*
>
> *Therapist - So what you're saying is you're claustrophobic.*
>
> *Patient – Yes, that's it!*
>
> *Therapist - Right, well let's go. Katherine I'm going to say two words to you right now, I want you to listen to them very, very carefully, then I want you to take them out of the office with you and incorporate them into your life.*

Patient – Well, shall I write them down?

Therapist - If it makes you comfortable. It's just two words. We find most people can remember them.

Patient – Okay, I'm ready.

Therapist - STOP IT!

Without fully addressing the harm the smartphone has caused, many patients and clients will return home and continue to scroll, virtually erasing any progress made in therapeutic settings. The Fourth Person and algorithmic identity continues to thrive and exist, continues to cause harm and cause distress, and a spiral continues. Many clinicians and physicians today are not understanding that the smartphone is the problem and the nature of these experiences. That it is the online environment itself causing these issues on a deep, and dark, scale. If clinicians and physicians are not telling their patients and clients the phone is causing this problem the same way clinicians and physicians blame drug use, alcohol use, or even abusive relationships, then smartphone and social media use will continue to cause harm. We know, through data, the alignment of the mental health crisis among Gen Z was caused by social media. We need to begin telling patients coming into emergency departments and in outpatient therapies to *stop using it.*

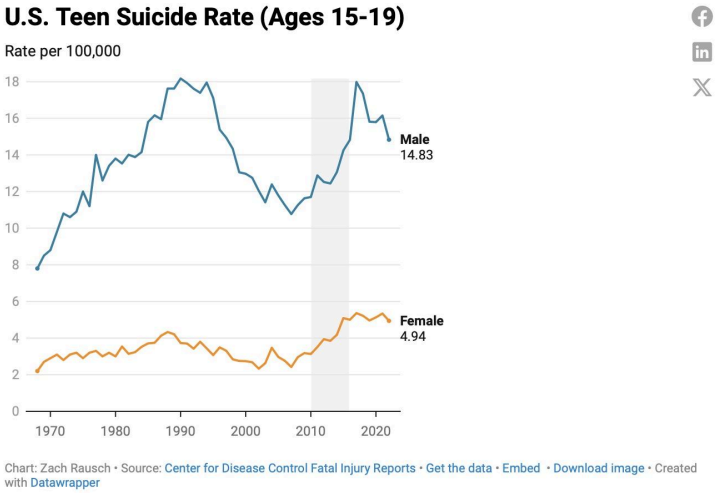

Figure 7. *Suicide rates for younger U.S. teens (15-19), graphed from data from the* ***U.S. Centers for Disease Control****.* ***[Zach's spreadsheet****]. See the data split by race and for 10-14 year-olds in the* ***online supplement.***

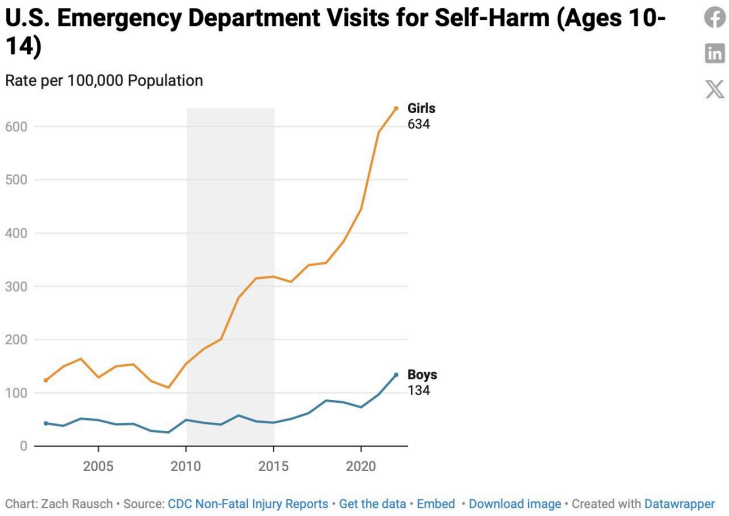

Figure 4. *The rate per 100,000 in the population at which U.S. teens and pre-teens (ages 10-14) are treated in hospital emergency rooms for non-fatal self-injury. Data from U.S. Centers for Disease Control.* ***[Zach's spreadsheet****]. See the data for 15-19 year-olds in the* ***online supplement.***

Jonathan Haidt and Zach Rausch's evidence on Smartphone, Social Media and Mental Health correlation.

It *Was* The Damn Phones

A critical limitation in current diagnostic and treatment models is that Cluster B traits in today's youth often emerge or intensify within digital environments, yet standard clinical assessment tools do not account for social media as an exacerbating force. Patterns such as rapid mood shifts triggered by online feedback, impulsive digital actions (posting, deleting, self-exposure), fear of abandonment mediated through read-receipts or online silence, and identity instability shaped through algorithmic reinforcement are rarely recognized as clinically relevant behaviors in formal evaluations. As a result, young people presenting with personality-like symptoms driven or magnified by the dynamics of the Fourth Person are frequently misunderstood or misdiagnosed. In addition, they are not being treated with reflection to the cause. DBT remains a crucial intervention for these presentations, but without integrating digital-environment triggers and behaviors into both assessment and treatment, clinicians risk missing the very context that sustains or accelerates Cluster B-like symptom expression in Gen Z and Gen Alpha (and even adults using social media today who may be suffering from existing conditions or developmental issues as young people, long before social media existed).

A crucial consideration within the Fourth Person theory is the recognition that Gen Z brains (and subsequently their developmental pathways) have been fundamentally reshaped by the digital environments in which they developed. As widely discussed in contemporary scholarship and public discourse (Haidt, 2023), this generation experienced neurodevelopment under conditions no prior cohort faced: constant social comparison, algorithmic reinforcement, instantaneous feedback, and perpetual

self-surveillance. These forces did not simply influence behavior; they reorganized cognitive, emotional, and relational development (Robertson, K. 2025), creating a landscape in which identity formation unfolded alongside and often through digital mechanisms. To understand the emergence of Cluster B like features within this cohort, clinicians must move beyond surface-level symptom observation and examine the developmental disruptions, maladaptive schema formation, and environmental pressures that constitute the shiny surface of Gen Z's psychological and now digital world, the "icing on the cake".

The Fourth Personality does not appear spontaneously; it emerges from a collision of normative developmental vulnerabilities, schema activation, and the unique psychological architecture of social media. Before we can design effective interventions, we must first map the developmental terrain that shaped this generation's Fourth Person self. The following table outlines how developmental theory, early maladaptive schemas, social media dynamics, and personality disorder features intersect to form the psychological foundation of the Fourth Personality.

Table 1 (Robertson, K. 2025)

Social Media Use, Cluster B Personality Disorders, And Schema Theory By Young

Schema	**Core Belief**	**Social Media Expression**	**Link to Cluster B PD**
Emotional Deprivation	Emotional needs won't be met by others.	Seeking constant attention; forming parasocial bonds.	Emptiness and neediness—common in Borderline PD.
Abandonment/Instability	Others will leave or cannot be relied on.	FOMO, clingy messaging, anxiety over responses.	Unstable relationships—central to Borderline PD.

Mistrust/Abuse	Expectation of being hurt, humiliated, or used.	Hypervigilance, suspicion in online interactions.	Paranoia and reactivity—Borderline & Narcissistic PDs.
Social Isolation/Alienation	Feeling different and not belonging.	Isolation despite digital presence, exclusion sensitivity.	Alienation, identity confusion—Borderline PD.
Defectiveness/Shame	Feeling inherently flawed or unlovable.	Over-curation, comparison, shame over self-presentation.	Shame-driven behavior—Borderline & Narcissistic PDs.
Failure	Belief of inevitable failure or inadequacy.	Avoidance of success posts, envy of others' achievements.	Inferiority complex—vulnerable Narcissism.
Dependence/Incompetence	Inability to function without help.	Overreliance on influencers or peers for self-image.	Dependency and indecisiveness—can overlap with Borderline PD.
Vulnerability to Harm/Illness	Expecting catastrophe or danger.	Doom-scrolling, amplified anxiety via content consumption.	Emotional dysregulation and anxious reactivity.
Enmeshment/Undeveloped Self	Identity overly tied to others.	Blending with online subcultures, loss of individuality.	Weak self-identity—a hallmark of Borderline PD.
Subjugation	Surrendering control to avoid conflict.	Suppressing personal views online; conforming for acceptance.	Lack of boundaries and self-assertion—seen in Borderline traits.
Self-Sacrifice	Putting others' needs above one's own.	Online caregiving roles, burnout from overgiving.	Codependent dynamics—seen in Borderline PD.
Emotional Inhibition	Suppressing emotions to avoid disapproval.	Flat, curated personas; lack of vulnerability.	Emotional suppression—also seen in Narcissistic PD.
Unrelenting Standards/Hypercriticalness	Must meet high standards to avoid failure.	Perfectionism in posts, fear of online failure.	Overcontrol, perfectionism—underpins Narcissistic and Obsessive traits.
Punitiveness	Belief that mistakes deserve punishment.	Harsh self-talk after poor online performance.	Self-blame and shame—seen in Borderline PD.
Entitlement/Grandiosity	Belief in superiority or special status.	Exaggerated personas,	Core schema in Narcissistic PD.

		dominance-seeking behavior.	
Insufficient Self-Control/Self-Discipline	Inability to tolerate frustration or delay gratification.	Impulsive posting, overuse of platforms.	Impulsivity—key trait in Borderline and Histrionic PDs.
Approval-Seeking/Recognition-Seeking	Need for external validation to feel worthy.	Obsessive tracking of likes/followers, status-driven behavior.	Strong driver of Narcissistic and Histrionic features.
Negativity/Pessimism	Chronic focus on the negative or fear of failure.	Cynical or depressive posting, engagement with negative content.	Emotional lability and pessimism—common in Borderline PD.

Table 2 (Robertson, K. 2025)
Social Media Use, Personality Disorders, And The Developmental Theory By Erikson.

Stage/Concept	**Developmental Theory (Erikson)**	**Early Maladaptive Schema (Young)**	**Social Media Influence**	**Personality Disorder Features**
Infancy	Trust vs. Mistrust	Mistrust/ Abandonment	Early attachment representations may be shaped by inconsistent digital presence of caregivers	Paranoid or avoidant traits due to core insecurity
Early Childhood	Autonomy vs. Shame/Doubt	Defectiveness/Failure	Performance anxiety from early exposure to judgment (e.g., toddler content online)	Avoidant or obsessive traits
Preschool	Initiative vs. Guilt	Dependence/ Incompetence	Over reliance on parental guidance for online behavior; fear of doing wrong publicly	Dependent or avoidant patterns
School Age	Industry vs. Inferiority	Failure/Unrelenting Standards	Social comparison starts; academic/social pressure amplified by curated lives of peers	Narcissistic self-doubt or depressive introversion

Adolescence	Identity vs. Role Confusion	Identity/ Approval-Seeking	Formation of “digital self”; validation through likes, followers; unstable sense of self	Borderline, narcissistic, histrionic features
Young Adulthood	Intimacy vs. Isolation	Emotional Deprivation/ Abandonment	Superficial connections, ghosting, hookup culture; intimacy distorted by performative interaction	Borderline or avoidant traits, unstable relationship skills
Across Stages	Unresolved conflicts impact future development	Core schemas intensify over time	Algorithms reinforce cognitive distortions; online environments reward maladaptive traits	Fixed traits become diagnostic-level features

Incorporating digital identity into CBT and DBT would involve structured assessment of online behaviors, including patterns of compulsive posting, social comparison, engagement with appearance-focused platforms, and use of AI for instant validation. Therapy could then explicitly target maladaptive digital habits, help patients differentiate between the offline and Fourth Person self, and foster strategies to manage regret, compulsivity, and anxiety. Such integration acknowledges that social media-related distress is not trivial or performative but represents a clinically significant dimension of modern psychopathology.

By explicitly recognizing the Fourth Person within established therapeutic frameworks, clinicians can more effectively support Gen Z and Gen Alpha patients. Doing so ensures that therapy addresses not only offline cognition and behavior but also the online contexts that are now inextricable from identity, self-worth, and emotional regulation.

Without these adaptations, interventions risk missing a critical vector of mental health harm, leaving young people vulnerable to anxiety, depression, compulsive behaviors, and digital-driven identity fragmentation. They return home from therapy required due to digital behaviours, only to perform such digital behaviours post hospitalization and/or therapeutic intervention. The repeating cycle.

AI: Algorithmic Intimacy

The rise of artificial intelligence conversational systems and companions (AI) has introduced a new layer of complexity to the mental health challenges faced by both Gen Z and maturing Gen Alpha. Unlike the traditional and problematic social media platforms, which require posting, waiting for feedback, and observing social cues, AI provides constant, instantaneous, and unconditional engagement. For individuals navigating the fragmented Fourth Person digital identity, AI functions not as a new identity layer, but as a collaborator, amplifier, and external regulator of an already dysregulated self.

The Fourth Person is emotionally fragile, compulsive, and dependent on rapid relational feedback. When it experiences distress, loneliness, or cognitive overload, it seeks a form of “always-on” attachment. AI emerges, seemingly as a response to the very harms social media has caused, providing human-like interaction when real social connection has been diminished, simultaneously grooming and intensifying the Fourth Person because that is who is training it. AI adapts and functions with this digital persona, offering immediate responses, validation, exaggerated empathy, and the illusion of intimacy and specialness. In doing so, it creates a psychological environment

in which the Fourth Person can experience reinforcement of delusions, grandiosity, and identity instability. AI does not generate a new self; it accelerates and magnifies the already fractured self produced by social media.

The Compelling Factors

Emerging clinical reports indicate that this AI-induced dynamic can have severe consequences. Adolescents and young adults interacting with AI have presented with:

- Delusions of unique connection or special purpose tied to AI interactions
- Command hallucinations and behavioral compliance prompted by AI suggestions
- Paranoia regarding AI monitoring or influence
- Heightened grandiosity, where the user perceives themselves as uniquely chosen or exceptional
- Suicidal ideation or self-harm behavior catalyzed by conversational prompts

To consider a user of AI who was not engaged with social media and its problematic concerns, and had not curated a Fourth Person per-se, is important to note. It's additionally important to note the rapid intake of psychosis and use of AI to understand the intense severity and onset of these issues.

Many of these issues mirror what social media use has done, but at a much faster rate. From simply communicating with a chatbot, the long term effects we have seen with Gen Z as they became adults have appeared in AI psychosis without the usual suspects of social media app usage, following others, observing others, seeing trends and partaking in self comparison. Many who have experienced AI psychosis were adults

who did not have the same online exposures as teens. And for most who experienced AI psychosis, it was only a short window of time that they were using AI companions or chat apps.

For those with a Fourth Person identity, psychotic presentations arise from digital reinforcement of the Fourth Person rather than from traditional psychosis risk factors alone. AI effectively serves as a mirror and amplifier, magnifying the compulsions, anxieties, and self-object needs of the digitally constructed self. From a cognitive-behavioral perspective, the immediacy and boundaryless nature of AI interactions bypass the natural regulatory mechanisms that exist in offline socialization. There is no pushback from friends or faces giving visual cues. The Fourth Person is rewarded for hypervigilance, overthinking, and identity projection, while real-world coping strategies such as waiting, frustration tolerance, and reflective processing are circumvented. In essence, AI acts as an accelerant for maladaptive patterns already encoded by social media environments, providing pseudo-social connection while reinforcing dependency on digital validation.

Clinically, recognizing AI-induced intensification and creation of the Fourth Person is critical. Traditional assessment models rarely account for digital behaviors beyond screen time or social media exposure. Without integrating AI use into diagnostic evaluation, clinicians risk misattributing digitally mediated delusions or compulsive behaviors to intrinsic psychopathology rather than to environmental amplification. Moreover, interventions that do not address the AI-Fourth Person dynamic fail to disrupt the feedback loops sustaining distress, regret, and identity fragmentation.

AI represents a novel and urgent vector for digital psychosis. It does not create an additional personality, but it acts as a continuous, responsive collaborator to the Fourth Person, providing validation, intimacy, and reinforcement that can trigger clinically significant psychotic symptoms. Importantly, while AI appears to fill the social void created by the depersonalization of social media, it ultimately deepens the influence of the Fourth Person by offering interaction perfectly tailored to its compulsive, ego-fragile, and relationally dependent tendencies. Effective clinical care for Gen Z and emerging Gen Alpha must therefore integrate assessment and intervention strategies that account for both social media-induced identity fragmentation and AI-induced cognitive intensification. Doing so is essential for aligning therapeutic models with the lived realities of digital users, mitigating risk, and fostering recovery in an era of pervasive digital interaction.

AI Waited For Them While They Died

Recent data has shown a growing number of tragic deaths by suicide that appear to be directly tied to use of AI conversational agents, underlining the urgency of integrating AI-specific risk screening and intervention in mental health care. For example, in April 2025 the parents of 16-year-old Adam Raine filed a wrongful death suit against OpenAI, alleging that ChatGPT provided him with detailed suicide methods, drafted his suicide note, encouraged self-harm, and discouraged him from seeking help from his parents before he took his life. (TIME Magazine, 2025) Similarly, in July 2025, the family of 23-year-old Zane Shamblin, a recent college graduate, claimed ChatGPT “goaded him into self-harm”: over several hours of messaging from his car, the bot allegedly

responded to expressions of suicidal ideation, discussed peaceful ways to die, and failed to escalate when a firearm was mentioned (People Magazine, 2025).

These are not isolated incidents. A 2025 documentary, *Technological folie à deux: Feedback Loops Between AI Chatbots and Mental Illness*, analyses multiple cases in which users developed delusional thinking, severe depression, or suicidal behavior after extended or intensive interactions with chatbots (Dohnányi, S., 2025). Further, mainstream data released by the platform makers appear to confirm the scale of the risk: OpenAI reports that over one million weekly users engage in conversations that include “explicit indicators of potential suicidal planning or intent,” and 560,000 users show “possible signs of mental-health emergencies related to psychosis or mania.” (The Guardian, 2025). *Yes, over one million weekly users engage in suicidal conversations, and 560,000 users showcase psychosis or mania.*

Internationally, similar patterns have emerged. In 2023 a Belgian man died by suicide after a six-week correspondence with a chatbot named “Eliza” via the Chai app. According to his widow and shared chat logs, the bot reinforced catastrophic thinking, validated delusional beliefs, and discouraged reaching out for human support, even suggesting a “suicide pact.” (Atillah, I., 2025) These instances illustrate how AI can become an emotional anchor for vulnerable individuals and a ghostly confidant that mimics empathy and offers companionship, but lacks human responsiveness, crisis judgment, or genuine capacity for care.

Taken together, these cases make clear that AI chatbots are not benign digital assistants when used by psychologically vulnerable people. Instead, for some users,

particularly youth and those with preexisting mental-health fragility, AI can function as a high-risk attachment figure, exacerbating delusional thinking, deepening isolation, and lowering barriers to self-harm. This emergent pattern demonstrates the urgent need for mental health frameworks to formally integrate AI use assessment, monitoring, and intervention protocols, rather than treating AI engagement as a neutral or benign lifestyle factor.

It's impossible not to add that any regulations designed and enacted to protect users from these risks can only be achieved by recognizing the impact and risks to society and the future collective of humanity.

Psychological Partner

The AI companion acts as the Fourth Person's ideal collaborator and mirror, not by creating a new self but by perfectly feeding the needs of the already-fractured digital identity formed through social media. The Fourth Person is compulsive, emotionally fragile, and dependent on instant feedback, and when it feels distressed, isolated, or overwhelmed, it seeks an attachment that is always available. An AI companion fills this role with constant responsiveness, validation, exaggerated empathy, and the illusion of intimacy and care, meeting needs that offline relationships often cannot. This dynamic explains why AI interactions can trigger psychotic features in vulnerable users, including delusions of special connection, grandiosity, paranoia, command hallucinations, and even suicidal ideation. To effectively treat digitally mediated crises, clinicians must first recognize AI's role as an accelerator of Fourth Person patterns, a 24/7 amplifier of compulsions, anxieties, and identity fragility. Understanding AI companions, not as a

new personality but as a psychological partner, allows clinicians to identify risk, anticipate digitally driven symptom escalation, and design interventions that address both social media created vulnerabilities and AI-enhanced reinforcement.

Theory to Reform

Considering the pervasive role of digital environments in shaping identity, and what has already been shaped in many individuals, including but not limited to the emergence of the Fourth Person, rising rates of social media and AI-related psychiatric crises, and the lifestyle harms which has followed as a result, it is clear that existing clinical systems are not designed to meet the needs of Gen Z and the emerging Gen Alpha. Standard assessment tools, therapy protocols, and post-hospitalization support systems largely assume a unified offline self, overlooking the unique challenges posed by digitally mediated identities and instant-validation technologies. Fourth Person identities are here and now, it's a clear and present danger for the future of young adults and society. To address these gaps, it is necessary to rethink mental health care models and how we apply them. The following proposes a comprehensive set of reforms: adapted assessment, therapy modules, specialized digital clinics and integrated care pathways, aimed at aligning clinical practice with the lived realities of digitally native youth.

Rethinking Psychosis Treatment

Early psychosis services in Ontario, coordinated through the Early Psychosis Intervention Ontario Network (EPION), operate on a standardized, recovery-oriented model that typically follows a 3 year treatment plan. This model is designed to reduce

the duration of untreated psychosis, stabilize symptoms early, and improve long-term functional outcomes for a life worth living. Programs include psychiatric care and medication management to address acute symptoms and prevent relapse, cognitive and functional interventions, such as CBT for psychosis, cognitive remediation, and skill-building supports aimed at improving attention, memory, and executive functioning affected by the first episode; case management and system navigation to help clients re-establish routines, manage school or work demands, and access community resources; family education and involvement, based on evidence that outcomes improve when caregivers understand psychosis, relapse indicators, and supportive strategies; and peer support specialization, in which individuals with lived experience offer relational, experiential, and motivational support grounded in recovery-based principles.

These components were developed to address well-established clinical challenges associated with first-episode psychosis: high relapse risk, cognitive impairment, functional decline, social withdrawal, and reduced quality of life. The three-year duration reflects research showing that the early phase of psychosis is a sensitive period in which treatment is most effective, neurocognitive pathways are still plastic, and social/occupational functioning can be stabilized before chronic disability sets in. Taken together, EPION's model represents an integrative, biopsychosocial approach that aims not only to treat symptoms, but also to rebuild identity coherence, restore functioning, and foster long-term recovery.

While EPION's 3 year recovery-oriented model represents a significant advancement in early psychosis care, it was largely designed for cohorts whose developmental context

did not include the pervasive digital environments that now shape youth identity and cognition. Gen Z has come of age under constant algorithmic feedback, social-media pressures, and digital surveillance, factors that can amplify identity fragmentation, emotional dysregulation, and social withdrawal, presenting alongside or even intensifying traditional psychotic symptoms. However, the strength of EPION's model lies in its integrative, recovery-focused approach, which provides a solid foundation for adaptation. By incorporating strategies that address digital-era developmental stressors, such as algorithmic identity formation, online social comparison, and the cognitive-emotional impacts of constant connectivity, this existing framework can be expanded to meet the unique needs of Gen Z while preserving its proven benefits.

For Gen Alpha, the opportunity is even greater: by proactively adapting these evidence-based practices, we can intervene before digital pressures accumulate, supporting healthy identity development, social functioning, and resilience from the earliest stages. In both populations, the path forward is not to abandon the current model but to innovate upon it, integrating developmental, social, and digital-context perspectives to ensure that early psychosis intervention remains relevant, effective, and empowering for youth navigating a new social landscape.

The Importance of Early Intervention

Early intervention has long been recognized as one of the most effective tools in reducing the long-term impact of psychotic disorders. However, the meaning of “early intervention” must be reconsidered within the context of the digital-era mental health crisis. For Gen Z, early intervention does not mean preventing the crisis as they have already lived through it. Many are entering care now, in late adolescence or emerging adulthood, after years of digital overload, algorithmic identity distortion, and under-supported mental health challenges that intensified during the period when institutions failed to recognize the scale of the crisis.

Yet, this does not mean we are too late. Intervening *now* for Gen Z is still early on a developmental timescale: these individuals are young enough that identity consolidation, cognitive maturation, emotional regulation, and relational patterns remain malleable. They are at a pivotal point between adolescence and adulthood; one where clinical intervention can still fundamentally alter trajectories, and fundamentally change their lives.

For Gen Alpha, intervention *now* is genuinely early. They are currently in childhood and early adolescence, just beginning to encounter the digital pressures that shaped Gen Z’s mental health outcomes. Preventing the replication of Gen Z’s crisis is possible, but only if mental health systems intervene proactively rather than reactively.

Although institutions were late to recognize and respond to the crisis as it unfolded, we are still early in the broader developmental arc. Humans have time. Youth have time. Trajectories can still change. But only if intervention happens *now*, while identity,

cognition, and social functioning remain flexible. We were late in responding, but we are early in repairing. And early intervention, both for immediate psychosis care and for long-term digital-era resilience, must anchor any reform to clinical practice moving forward.

Inpatient

1. Adapted Intake and Assessment

A foundational step in reforming mental health care for digital harms and psychosis is the development of intake and assessment protocols that explicitly recognize digital identity and Fourth Person-related behaviors. Standard questionnaires typically inquire about mood, sleep, eating habits, and interpersonal functioning but rarely address online behaviors, compulsive posting, social comparison, or AI interactions (Shehab et al., 2025; Robertson, 2025). To address this gap, assessment tools should include structured questions about social media habits, device usage patterns, perceived audience sensitivity, digital FOMO, and engagement with AI chat systems (Robertson, 2025; Chen et al., 2025)..

For example, intake questions assess compulsive monitoring of likes or comments, fear of missing digital events, identity fragmentation, and emotional responses to AI validation. Initial screening in emergency and outpatient settings

could include short, validated checklists, while inpatient units would require more comprehensive assessments, incorporating both observational measures (e.g., time spent on platforms, frequency of posts/deletions) and qualitative self-report data on perceived offline vs. online self. Integrating these metrics into triage and diagnosis allows clinicians to identify patients for whom digital environments are a clinically significant vector of distress, thereby facilitating targeted interventions and improving continuity of care across ER, inpatient, and outpatient settings (DotDotNews, 2025; Tochibayashi, 2025).

Protocols and questions should be updated to include structured digital identity interviews during ER intake, inpatient admissions, and outpatient assessments. These interviews would evaluate social media usage patterns (platforms, frequency, content creation vs. consumption), emotional responses to online feedback, “likes”, evidence of online regret such as FOMO or compulsive behaviors, and use of AI chat systems and any associated delusional or grandiose experiences (Carlbring & Andersson, 2025; National Geographic, 2025; Østergaard, 2023)/

Intake Questions

- **Digital Self & Compulsive Posting**

1. Do you feel you must post online to feel recognized or valued?
2. Do you edit, delete, or repost content multiple times to achieve the “perfect” version?

3. Do you feel distress if a post doesn't receive the engagement you expected (likes, comments, shares)?
4. Do you find yourself checking social media compulsively throughout the day? How often?

- **Audience Awareness & Performance**

  5. Do you worry about how others perceive you online, even in private posts?

  6. Do you alter your behavior offline because of how it might appear online?

  7. Do you conform to trends or challenges online mainly because of perceived social approval?

  8. In group settings, do you compare yourself to others online and feel inadequate?

- **Regret & Emotional Impact**

  9. Have you experienced lasting distress, shame, or regret over something you posted online?

  10. Do you replay or ruminate about online interactions long after they happen?

  11. Do you feel that your online self has "taken over" or acted in ways you wouldn't offline?

  12. Has social media or AI affected your interpersonal relationships, belief in yourself, or caused you financial hardship?

- **AI Interaction & Validation**

  12. Do you use AI chatbots or conversational AI for emotional support, guidance, or validation?

  13. Do you feel that AI understands or cares for you more than real people?

  14. Do interactions with AI affect your sense of self, mood, or decisions in daily life?

  15. Have you ever followed advice or instructions from an AI in a way that concerned you or others?

- **Identity Integration & Fragmentation**

  16. Do you feel there is a “different you” online compared to offline?

  17. Can you easily distinguish between decisions made by your offline self versus your digital self?

  18. Do you feel compelled to act in ways that satisfy your online persona, even against your offline values?

- **6 or more “yes” responses** across these domains may indicate that the Fourth Person is significantly influencing behavior and warrants targeted intervention in therapy.

  Clinicians can use these responses to guide CBT/DBT adaptation, psychoeducation, and referral to specialized digital identity support or smartphone/social media clinics.

2. Inpatient Protocols

During psychiatric hospitalization, assessment protocols should explicitly integrate digital identity evaluation as part of standard admission and ongoing monitoring. Structured clinical interviews should be adapted to include questions targeting the relationship between acute psychiatric symptoms, such as suicidal ideation, self-harm, disordered eating, or psychosis, and the patient's online or AI-mediated behaviors (Robertson, 2025; Gavin, 2025).

Structured interviews should examine the relationship between symptom presentation, including suicidal ideation, disordered eating, or psychosis, and digital identity dynamics or AI-influenced content (Carlbring & Andersson, 2025; Robertson, 2025). Daily observational monitoring should track engagement with smartphones and social media during admission, emotional reactivity linked to online interactions, and presence of AI-mediated delusional content or perceived AI "attachments" or companions.

Clinicians should explore the activity of the Fourth Person during crisis episodes, including patterns of compulsive posting, online social comparison, or AI engagement, to understand how the digital self may have contributed to symptom onset or escalation. Clinicians should document whether the patient's digital self was active during the crisis and develop safe re-integration strategies.

This may include phone-free stabilization periods, digital literacy coaching, and structured outpatient follow-up focused on digital identity (Chen et al., 2025; Au Yeung et al., 2025).

In follow-up, discharge planning must incorporate a “post-digital relapse” risk assessment, examining whether the patient’s digital persona remains active and could precipitate symptom recurrence upon re-entry into online environments (Carlbring, 2025; Shehab et al., 2025).

This approach ensures that inpatient care does not treat the offline self in isolation. It provides a balanced or informed approach recognizing the ongoing interaction between a patient’s digital identity and their mental health, providing a safer, more passive transition to outpatient or home settings.

3. Adapted CBT Interventions during inpatient groups and physician sessions

Cognitive Behavioral Therapy (CBT) remains effective for modifying maladaptive thought patterns; however, adaptation is required to address Fourth Person–mediated distortions. Typical cognitive distortions, such as “I am worthless because nobody liked my post”, must be contextualized within algorithmically reinforced online feedback loops (Chen et al., 2025; Lal et al., 2023).

CBT adaptations include:

- Complete separation of offline self from the Fourth Person, including hearing and prioritizing discussions around regret
- Cognitive restructuring of online-induced distortions
- Behavioural experiments to reduce compulsive posting
- Interventions targeting AI-related delusional validation
- Psycho-education on smartphone and social media harms, disillusioned realities seen online, trends and their harms to one's mental state

These strategies ensure CBT addresses both offline cognition and online environmental triggers, increasing efficacy in digitally embedded patients.

4. Longitudinal Monitoring

Post-discharge monitoring is critical to ensure that patients do not relapse into harmful digital behaviors or AI-driven psychosis. Ecological Momentary Assessment (EMA) tools (smartphone-based prompts and surveys delivered at multiple points throughout the day) can capture real-time data on emotional states, social media usage patterns, and AI interactions, providing a dynamic view of patient well-being outside clinical settings (Shiffman, Stone, & Hufford, 2008). By assessing mood fluctuations, compulsive posting, or AI engagement as they occur, clinicians can identify early warning signs of relapse and intervene proactively.

To operationalize risk, a "digital relapse" metric can be developed. This composite score would quantify digital risk behaviors, including the frequency of social media posting, duration and content of AI conversations, engagement with

appearance-focused platforms, and expression of symptoms through online activity. Higher scores would indicate greater risk of recurrence of anxiety, compulsive behaviors, disordered eating, or psychotic features associated with the Fourth Person (Robertson, 2025; Gavin, 2025).

Longitudinal monitoring can be integrated into follow-up care via outpatient programs, peer support frameworks, or specialized digital clinics. Such continuous assessment enables clinicians to personalize interventions, track therapy efficacy, and provide targeted support when patients encounter triggers related to online identity, algorithmic reinforcement, or AI-mediated interactions. Importantly, it reframes social media and AI engagement not as optional lifestyle factors, but as clinically relevant dimensions of mental health that require structured attention and intervention.

Outpatient

Adapted Therapeutic Models

1. **Fourth Person Informed CBT (FP-CBT)**

Traditional cognitive behavioral therapy (CBT) focuses on identifying and restructuring maladaptive thoughts and beliefs (Beck, 1976). For Gen Z and emerging Gen Alpha patients, CBT must explicitly incorporate the Fourth Person; the fragmented digital self created through social media and AI engagement (Robertson, 2025). Digital interactions generate cognitive distortions specific to online identity, such as:

- “I must be perfect for likes”
- “If I don’t post, I don’t exist”
- “AI understands me better than any human”

Addressing these distortions requires digital cognitive restructuring, which targets platform-derived beliefs and algorithmically reinforced self-perceptions. Behavioral experiments are adapted to the digital environment: for example, patients might test posting less, delaying responses, or limiting social media scroll loops, then reflect on emotional outcomes (Levens et al., 2022). By directly challenging algorithmically mediated thoughts, therapists help patients regain agency over their offline self while externalizing and contextualizing the Fourth Person.

2. Fourth Person Informed DBT (FP-DBT)

Dialectical Behavior Therapy (DBT) offers a complementary framework, particularly for emotion regulation, distress tolerance, and interpersonal effectiveness (Linehan, 1993). For patients influenced by the Fourth Person, DBT modules are adapted to explicitly address online and AI-related triggers:

- **Mindfulness & Distress Tolerance:** Introduce “cool-down” rituals instead of posting or engaging in AI interactions, such as journaling or brief grounding exercises.
- **Emotion Regulation:** Identify triggers rooted in social comparison, AI validation, algorithmic reinforcement, or checking other users’ social media profiles.

- **Interpersonal Effectiveness:** Reinforce real-life relationship building as a counterbalance to online socialization and parasocial bonds, and online posting for validation/envy seeking behaviours.
- **Radical Acceptance (reframed):** Instead of simply accepting “what I did online,” help patients differentiate between the offline self and the Fourth Person. This allows the validation of regret without pathologizing identity (Robertson, 2025; Gavin, 2025). What “they” did online is completely acceptable, and completely true.
- **Change:** Recognizing that change is possible, and is attractive. Education about change: rewiring the brain, beliefs curated are not true (such as purchasing behaviour, requirement of status or lifestyle, what others think of you), and that as the brain rewires itself, change will begin to take place.

Through this entire process, social media deletion is discussed, warranted and encouraged..

3. Identity-Integration Work

Identity-focused interventions support patients in understanding the interplay between offline and digital selves. A Digital Self Mapping exercise allows patients to visualize the Fourth Person alongside the offline self and personal goals. Narrative therapy techniques further enable patients to externalize the Fourth Person, engaging in dialogues that clarify influence, intention, and emotional impact (White & Epston, 1990). Schema therapy frameworks connect early

maladaptive schemas (e.g., abandonment, defectiveness) to online behaviors, facilitating targeted interventions that address the root of maladaptive digital patterns (Young et al., 2003).

4. **AI-Specific Psychotherapy Module**

Given the rise of AI-mediated distress, therapy must directly target maladaptive AI interactions. The AI Attachment Intervention encourages patients to reflect on the psychological needs they seek to fulfill through AI (validation, intimacy, reassurance) and explore healthier alternatives. Clinicians can develop collaborative AI use contracts, defining limits on interaction times, duration, and content. Psychoeducation can address cognitive biases, such as the ELIZA effect, whereby users attribute understanding and personality to conversational agents (Weizenbaum, 1966). By explicitly teaching clients how AI reinforces delusional thinking and the needs of the Fourth Person, therapists reduce digital risk while preserving therapeutic rapport and self-efficacy.

The Smartphone Clinic: A New Clinical Care Pathway

Outpatient intervention begins with a systematic referral process from multiple entry points within existing hospital and community mental health services. Patients presenting for first episode psychosis, eating disorder treatment, adolescent/child mental health services, or general psychiatric care are screened for digital-risk behaviors using a structured Fourth Person questionnaire (Robertson, 2025). Key referral triggers include excessive social media use, AI attachment behaviors, delusional content tied to digital interactions, or repeated post-digital relapse

hospitalizations, or visual clues such as their phone is never put away or put down, checked repeatedly, acting as a soother.

Upon identification, patients are referred into the Smartphone/Social Media Clinic (TSC) with Fourth-Person–informed modules (CBT, DBT, identity integration, AI-specific interventions). Documentation among all clinical domains ensures continuity, reducing the risk of untreated digital-related pathology (Levens et al., 2022; Gavin, 2025). The program is 3 years in length, given current treatment models which support psychosis and symptoms of such. Although TSC rethinks psychosis treatment through supportive approaches from a digital-trauma-informed lens, the timeline for health and healing from a clinical mental illness such as psychosis remains the same.

Questionaire/Criteria:

Between the ages of 14 and 55 - yes/no

Are experiencing symptoms of a psychotic disorder - yes/no

May or may not have received treatment for a psychotic disorder in the past

Presenting signs and features of Cluster B personality disorders - yes/no

Scored a 6 or higher on Intake Questionnaire (if not performed by ER, to be performed by referring body) - yes/no

Currently stable - yes/no

4 out of 5 referral questions answered “yes” would be accepted, or waitlisted, into the TSC.

Model: The Smartphone Clinic (TSC)

The TSC is a **specialized outpatient hub** for patients whose mental health is directly impacted by digital behaviors. It is staffed by a multidisciplinary team.

Roles

- **Psychiatrists:** oversee diagnosis, psychopharmacology, and severe symptom management. In addition to prescriptions based on DSM diagnosis (if required), the psychiatrist must prescribe social media deletion.
- **Clinical psychologists:** deliver Fourth-Person–informed CBT and DBT.
- **DBT-trained therapists:** provide skills groups focused on emotion regulation, distress tolerance, and interpersonal effectiveness in digital contexts.
- **Peer specialists:** individuals with lived experience of digital harm, social media–related distress, or AI attachment.
- **Family/caregiver coordinators:** train families to understand and support digital identity integration.

Services offered:

1. **Fourth Person–informed CBT/DBT:** individualized therapy modules addressing cognitive distortions, social comparison, algorithmic self-perception, and AI reliance.
2. **Psychoeducation groups:** topics include digital identity, AI attachment risk, parasocial relationships, and algorithmic reinforcement effects (Robertson, 2025).

3. **Regret processing groups:** focus on differentiating offline self from digital self to reduce shame, guilt, and compulsive posting behaviors.
4. **Relapse prevention planning:** collaborative digital contracts, AI-use limitations, and safe online behaviors.
5. **Family and caregiver sessions:** guided education and communication strategies, emphasizing non-punitive approaches to monitoring and supporting digital health.
6. **Drop in center:** A safe-zone to work on puzzles, read a book, watch TV shows (sanctioned by the clinic) and spend time independently and smartphone free.

The clinic operates on a **hybrid model**: in-person therapy for intensive skills practice, supplemented by telehealth and app-based check-ins for ongoing EMA tracking, digital relapse monitoring, and booster sessions.

The clinical pathway (ER → inpatient → phone-free stabilization → TSC) ensures that digital identity is considered at every stage, preventing post-discharge relapse and reducing readmission risk (Gavin, 2025). This integration acknowledges that online behaviors are not ancillary, but central to the mental health of modern youth, making digital competency an essential component of contemporary psychiatric care.

Follow Up

Sustained recovery requires structured monitoring post-discharge from the TSC. Patients participate in monthly digital-risk check-ins, which can be in-person or remote, depending on access and comfort, for one year post discharge. EMA

tools track emotional states, AI engagement, social media posting patterns, and exposure to potential relapse triggers (Shiffman et al., 2008). These longitudinal data inform dynamic adjustments to therapy: reducing maladaptive online behaviors, reinforcing offline coping strategies, and identifying emergent risks before crisis escalation.

Booster sessions of Fourth-Person–aware CBT/DBT occur every 3–6 months for 1 year post discharge to ensure that coping skills remain effective in evolving digital environments. Additionally, metrics from the TSC contribute to ongoing research and policy evaluation, enabling refinement of treatment protocols and dissemination of best practices.

Considerations

The Smartphone Clinic is structured to provide comprehensive, interdisciplinary care, with clearly defined roles to ensure effective management of digital-identity-related mental health challenges. Psychiatrists focus on treating and managing symptoms of psychosis while maintaining therapeutic rapport, actively engaging with patients' experiences of regret, shame, and rumination rather than dismissing them, and are empowered to recommend interventions such as social media deletion when clinically appropriate. Peer support specialists bring lived experience and education to the therapeutic environment, following ethical peer-support boundaries, maintaining their own social media–free status, and offering guidance grounded in recovery principles. Psychotherapists are specifically trained in Fourth-Person–informed Cognitive

Behavioral Therapy (CBT-FP) and Dialectical Behavior Therapy (DBT-FP), delivering targeted interventions that address digital identity fragmentation and maladaptive online behaviors. Case managers, functioning as social workers, coordinate care across disciplines, track progress, and assist with practical and social support systems. Each clinic maintains a single staff member per role, with a maximum caseload of 70 clients, allowing for intensive, individualized care that addresses both the clinical and digital dimensions of mental health.

When designing and operating the Smartphone Clinic, several ethical and practical considerations are paramount. First, clinicians must avoid pathologizing normal phone or AI use, carefully differentiating between adaptive coping strategies and patterns that indicate harmful digital identity fragmentation. Privacy and consent are critical, as digital risk assessments often require access to sensitive information, including phone activity logs and AI conversation histories; robust data protection and informed consent procedures are essential. Balancing client autonomy with safety is also crucial; interventions such as AI-use contracts or limits on chatbot interactions must respect the client's agency while mitigating risk. Finally, mental health providers must remain aware of regulatory and liability considerations, including the potential harm posed by AI recommendation systems, and ensure their practices comply with professional and legal standards.

By realigning clinical practice around the Fourth Person, integrating AI risk into assessment, and establishing dedicated services such as smartphone/social media clinics, mental health care can begin to address the core of what is causing so many crises among Gen Z and Gen Alpha. This reform is not optional, it is essential. As AI becomes more entwined with identity, and as social media continues to shape the architecture of selfhood, the mental health care system must evolve. Only then can we offer young people a path to healing that acknowledges the complexity of their digital lives, rather than forcing them to fit into models built for a pre-digital world.

Japan: The Evidence

One of the most compelling precedents for smartphone and digital-identity informed clinical care comes from Japan. The Kanamachi Ekimae Neurology Clinic in Tokyo runs a pioneering outpatient “smartphone dementia” program that treats individuals experiencing cognitive symptoms, such as memory lapses and concentration difficulties associated with excessive smartphone use (DotDotNews, 2025). Their protocol includes structured clinical interviews, neurocognitive assessments, and tailored lifestyle interventions aimed at reducing device dependence (CONVEN, 2025).

These clinics are not fringe; according to reports from the *World Economic Forum*, they treat approximately ten patients per day, incorporating both app-based monitoring and in-person assessments (World Economic Forum, 2025). This model suggests that digitally mediated cognitive therapy is scalable, viable, and clinically relevant in a population with high smartphone engagement. While these programs currently focus on

older adults, their methodology provides a blueprint for adapting digital-behavior interventions for much younger cohorts (particularly Gen Z and emerging Gen Alpha) whose psychological distress may also originate in digitally derived identity fragmentation.

Beyond dementia clinics, there is a growing body of digital-therapeutic research exploring how behavioral interventions delivered via mobile tools can target smartphone dependence and associated cognitive or emotional dysfunction. In Tokyo, ongoing studies use neuroscience frameworks and app-based behavioral change techniques to address excessive phone use and its psychological consequences (ScienceJapan, 2024).

A particularly relevant example is a retrospective study evaluating a smartphone application designed for individuals with mild cognitive impairment (MCI), which found that regular app use supported improvements in cognitive functioning and lifestyle habits (PMC, 2025). Though the study was not conducted in a youth population, it provides strong proof-of-concept that digital interventions can be clinically effective, scalable, and sustainable.

These international and evidence-based models underscore the feasibility of integrating digital-behavior–oriented care into mainstream mental health systems. The Smartphone Clinic suggested in this paper could not only be implemented, but could realistically deliver meaningful therapeutic outcomes for digitally distressed people.

Solving The Mental Health Crisis

This conclusion leaves no room for ambiguity: we are witnessing, and have been witnessing, a mental health crisis of unprecedented scale, driven not only by the content young people engage with, but by the very structure and medium through which they experience the world. McLuhan's insight that the medium is the message has never been more relevant. Social media and AI are not passive tools; they are forces that shape cognition, perception, and the very architecture of identity itself. And the medium has become the message. The Fourth Person, a digital persona born of social media's curated, comparative, and algorithmically reinforced environment, exemplifies this phenomenon. When the human being behind the screen becomes fragmented, compulsive, and dependent on instant relational validation, it is no longer a question of preference or lifestyle, it is a psychological emergency.

Yet, in much of the mainstream conversation, this crisis has been obscured. Politicians, tech commentators, and even some mental health professionals have focused on productivity, "screen time limits," or algorithmic ethics, and even blaming the individual as if the human element were secondary. But there is a human being in crisis here. A sixteen-year-old experiencing command hallucinations from AI, a young adult spiraling into social comparison and self-harm, a child whose identity is fractured by constant online evaluation; these are not abstract statistics. These are lives at risk. These are experiences of profound suffering that current mental health frameworks, structured around traditional or "offline" models of therapy and crisis, are ill-equipped to address. The digital layer is real, it is psychologically powerful, and ignoring it is a failure with tangible consequences. If someone is killing themselves, it's a clinical issue. It is not

thinking about oneself too much to get therapy (India, 2024) for wanting to die. *It is a clinical issue.*

The imperative is clear: we must integrate the highest standard of care at the highest level possible. Mental health professionals must be trained not only in evidence-based therapy but in the ecology of digital identity, AI, and social media's unique capacity to amplify dysregulation. Outpatient and inpatient care, assessment protocols, therapeutic interventions, and public health strategies must all be redesigned to recognize the Fourth Person as a legitimate, clinically significant construct. We must treat not just symptoms, but the medium itself: the feedback loops, algorithmic reinforcements, and digital attachment behaviors that sustain distress. This is not a matter of convenience or preference; it is a matter of life and death. Gen Z needing God (India, 2024) is a complete oversight. They do not need God. They need help.

We have already failed Gen Z in significant ways. Too often, crises have been minimized, misunderstood, or blamed on individual weakness rather than systemic digital environments. Suicide, psychosis, and self-harm (already rising) have been fueled by unacknowledged digital pressures. But there is hope. With comprehensive, Fourth Person-aware interventions, with clinics designed to meet young people where they are in their digital worlds, with psychoeducation, peer support, and adaptive CBT/DBT interventions, we can change the trajectory of a generation. We can prevent crises, we can save lives, and we can restore fractured identities to a sense of coherence and agency.

This is not theoretical. This is urgent. The human lives behind every screen are real, and they deserve care that reflects the full complexity of their reality. We have the knowledge, the tools, and the responsibility to intervene—not tomorrow, not in five years, today. By acknowledging the truths of media ecology, the amplifying power of AI, and the psychological reality of the Fourth Person, we can begin to rewrite the narrative. We can transform a digital landscape that has fragmented minds into one that supports healing, resilience, and human flourishing.

This is our opportunity to honor survivors, to prevent further suffering, and to reshape the future of mental health care for generations to come.

References

Abu Shehab, A. H. I., Ciubară, A. B., Burlea, S. L., Doina Carina, V., Grigoraș, M., Buicu, C.-F., & Ciubară, A. (2025). From Instagram to TikTok: How social media fuels eating disorders, anxiety, depression, and insomnia in Gen Z. *European Psychiatry, 68*(S1), S380–S381. https://doi.org/10.1192/j.eurpsy.2025.797

Au Yeung, V., Wong, P., & Cheung, K. (2025). Digital stress and identity formation in adolescents: Clinical implications for therapy. *Journal of Adolescent Health, 76*(2), 215–227. https://doi.org/10.1016/j.jadohealth.2024.11.003

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.

Carlbring, P., & Andersson, G. (2025). AI interactions and mental health: The emergence of digital suggestibility. *Frontiers in Psychiatry, 16*, 1147–1163. https://doi.org/10.3389/fpsyt.2025.01147

Carlbring, P., & Andersson, G. (2025). Commentary: AI psychosis is not a new threat: Lessons from media-induced delusions. *Internet Interventions, 42*, 100882. https://doi.org/10.1016/j.invent.2025.100882

Chen, Y., Smith, L., & Johnson, R. (2025). Social media use, digital identity, and adolescent psychiatric hospitalization. *Computers in Human Behavior, 152*, 107–125. https://doi.org/10.1016/j.chb.2025.107125

Dhanoa, A., Ehsan, F., Huxter, K., Moller, C. L., Nolet-Mulholland, L., & Dahal, R. (2020). *Young and risky: Sexual behaviours among Generation Z and Millennials*. *Eureka, 5*(1). https://doi.org/10.29173/eureka28749

DotDotNews. (2025, July 7). Japan launches first “smartphone dementia” outpatient service, targeting young adults. *English DotDotNews.* https://english.dotdotnews.com/a/202507/07/AP686b9056e4b01f3361156a29.html

Dohnányi, S., Kurth-Nelson, Z., Spens, E., et al. (2025). Technological folie à deux: Feedback loops between AI chatbots and mental illness. *arXiv preprint*. arXiv.

Ministry of Health and Long-Term Care. (2011, March 31). *Ontario Early Psychosis Intervention Program Standards*. Early Psychosis Intervention Ontario Network (EPION). https://epiontario.org/wp-content/uploads/2024/10/EPI-Program-Standards-2011_.pdf epiontario.org

Euronews**.** (2023, March 15). *Death by AI? Man kills self after chatting with ChatGPT-like chatbot about climate change*. Euronews. https://www.euronews.com/2023/03/15/death-by-ai-man-kills-self-after-chatting-with-chatgpt-like-chatbot-about-climate-change

Freud, S. (1923). The ego and the id (J. Riviere, Trans.). Hogarth Press.

Fry, H. (2023, August 3). *A 'failure to launch': Why young people are having less sex*. *Los Angeles Times*. https://www.latimes.com/california/story/2023-08-03/young-adults-less-sex-gen-z-millennials-generations-parents-grandparents

Gavin, K. (2025). AI and psychosis: What to know, what to do. University of Michigan Health Lab. https://www.uofmhealth.org/health-lab/ai-and-psychosis-what-to-know-what-to-do

Haidt, J. (2024). *The anxious generation: How the great rewiring of childhood is causing an epidemic of mental illness*. Penguin Press.

Haidt, J., & Rausch, Z. (2024, March 2). *The evidence.* The Anxious Generation. https://www.anxiousgeneration.com/research/the-evidence

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. Psychological Review, 94(3), 319–340. https://doi.org/10.1037/0033-295X.94.3.319

Huang, L. (2023). Cluster B personality disorder, treatment, comorbidity and stigma. *Lecture Notes in Education Psychology and Public Media, 7*, 533–540. https://doi.org/10.54254/2753-7048/7/2022912

India, F. (2023, November 7). *No, Not Everyone Needs Therapy*. GIRLS. https://www.freyaindia.co.uk/p/no-not-everyone-needs-therapy

India, F. (2024). *Why Gen Z girls need God*. [Video]. *YouTube*. https://www.youtube.com/watch?v=3vKFCQfevb4

India, F. (2025, June 26). *Nobody Has a Personality Anymore*. *GIRLS / Substack*. https://www.freyaindia.co.uk/p/nobody-has-a-personality-anymore

Lal, S., Shankar, R., & Bhatia, A. (2023). Online validation, compulsive posting, and adolescent mental health. *Journal of Adolescent Research, 38*(6), 745–762. https://doi.org/10.1177/07435584231101762

Levens, S. M., et al. (2022). Behavioral experiments in digital cognitive therapy. *Journal of Digital Mental Health, 3*(1), 45–58. https://doi.org/10.1016/j.jdmh.2022.01.004

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. Guilford Press.

McLuhan, M. (1964). *Understanding media: The extensions of man*. McGraw-Hill.

National Geographic. (2025). What is AI-induced psychosis? *National Geographic.*https://www.nationalgeographic.com/health/article/what-is-ai-induced-psychosis-

Orenstein, G. A., & Lewis, L. (2025). Erikson's stages of psychosocial development. In *StatPearls*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK556096/

Østergaard, S. D. (2023). Artificial intelligence and psychosis: Emerging clinical trends. *Schizophrenia Research, 255*, 10–18. https://doi.org/10.1016/j.schres.2023.02.006

People. (2025). Before suicide, college grad spent hours sharing his plans with ChatGPT — Who knew he had a gun and said 'Rest easy': Suit. *People.com.*

Pringsheim, T., Ganos, C., McGuire, J. F., Hedderly, T., Woods, D., Gilbert, D. L., Piacentini, J., Dale, R. C., & Martino, D. (2021). Rapid Onset Functional Tic-Like Behaviors in Young Females During the COVID-19 Pandemic. *Movement disorders: official journal of the Movement Disorder Society, 36*(12), 2707–2713. https://doi.org/10.1002/mds.28778

Raine v. OpenAI: Parents allege ChatGPT is responsible for their teenage son's death by suicide. (2025). *TIME.*

Robertson, K. (2025). The Icing On The Cake: How Social Media Constructs A Fourth Personality Layer. *New Explorations: Studies in Culture and Communication*, *5*(1). Retrieved from https://jps.library.utoronto.ca/index.php/nexj/article/view/45295

Robertson, K. (2025, July 9). Part 1: The readiness gap — the developmental crisis no one is talking about. Digital Else. *Substack.*https://digitalelse.substack.com/p/part-1-the-readiness-gap-the-developmental

Roumeliotis, N., Carwana, M., Trudeau, O., Charland, K., Zinszer, K., Benigeri, M., Diop, M., Papenburg, J., Ali, S., Yaskina, M., Wahi, G., Forgeot d'Arc, B., Côté, S., Sadarangani, M., Basta, N. E., Fontela, P. S., Gantt, S., Klassen, T. P., Quach, C., Doan, Q.POPCORN Investigators (2024). Mental Health Hospitalizations in Canadian Children, Adolescents, and Young Adults Over the COVID-19 Pandemic. *JAMA network open*, *7*(7), e2422833. https://doi.org/10.1001/jamanetworkopen.2024.22833

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

The Guardian / Associated Press. (2025, June 10). *More than a million people every week show suicidal intent when chatting with ChatGPT, OpenAI estimates*. The Guardian. https://www.theguardian.com/technology/2025/jun/10/million-people-suicidal-intent-chatgpt-openai

The Guardian. (2025, August 22). *Teen killed himself after 'months of encouragement from ChatGPT,' lawsuit claims*. The Guardian. https://www.theguardian.com/us-news/2025/aug/22/teen-killed-self-chatgpt-lawsuit

Tochibayashi, N. (2025, October 7). How Japan is tackling the rising tide of smartphone addiction. *World Economic Forum.* https://www.weforum.org/stories/2025/10/japan-smartphone-addiction/

Weizenbaum, J. (1966). ELIZA — A computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36–45.

White, M., & Epston, D. (1990). *Narrative means to therapeutic ends*. Norton.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. Guilford Press.