International Journal of Innovative Studies in Humanities and Social Studies
ISSN: 3067-7319 | Volume 2, Issue 3, 2026
www.aytinpublications.com

Review Article

The Use of AI Chatbots in Adolescence: An Overview of Effects With Clinical and Ethical Implications for Mental Health

Giuseppe Manuel Festa[1,2*]

[1]*Pontifical Faculty of Educational Sciences AUXILIUM, 00166 Rome, Italy*
[2]*Interdisciplinary Institute of Advanced Clinical Training IACT, 00128 Rome, Italy*

*Corresponding Author: Giuseppe Manuel Festam Pontifical Faculty of Educational Sciences AUXILIUM, 00166 Rome, Italy.

Abstract

In recent years, artificial intelligence-based chatbots have become increasingly present in adolescents' daily lives, used for various purposes, including mental health support. The aim of this study is to review recent literature (2022-2026) on the use of AI chatbots for psychological support among adolescents. Evidence regarding effectiveness, as well as associated risks and clinical and ethical implications, is examined. Findings from recent studies suggest that chatbot-based interventions may lead to modest reductions in psychological distress. However, generative AI systems show considerable variability in response quality and raise concerns related to privacy, misinformation, limited crisis-detection capabilities, social isolation, and potential dependency. This review discusses the main theoretical frameworks, evaluates the available empirical evidence, and analyzes the ethical and clinical challenges in the near future. Future research should focus on longitudinal study designs and the potential clinical validation of hybrid systems with therapist supervision, in order to assess these tools as complementary supports in adolescent mental health care.

Keywords: adolescent mental health, AI chatbots, generative artificial intelligence, digital therapeutics, psychotherapy.

Article Information

Received: 11 June 2026

Accepted: 22 June 2026

Published: 24 June 2026.

Cite this article as:

Giuseppe Manuel Festa. The Use of AI Chatbots in Adolescence: An Overview of Effects With Clinical and Ethical Implications for Mental Health. International Journal of Innovative Studies in Humanities and Social Studies, 2026: 2(3); 36-41.

https://doi.org/10.71123/3067-7319.020305

Introduction

Adolescence is a complex transitional period between childhood and adulthood, characterized by profound physical, psychological, and neurobiological changes. During this stage, a sense of identity is consolidated, processes of individuation from parental figures are promoted, and social competencies are developed. This condition, marked by the need to assume new roles and to resolve frequent conflicts, confronts individuals with both novel developmental tasks and previously encountered challenges that must now be addressed through age-appropriate adaptive strategies. Adolescents, embedded within their social environments, invest considerable energy in attributing meaning to their actions and in reconstructing a sense of personal existence.

Although this period is highly challenging and transformative, it also represents a phase of increased psychological vulnerability, including the emergence of anxiety and depression (Steinberg, 2014). Despite the expansion of traditional mental health services across countries, for some individuals the access remains limited because of stigma, geographical barriers, and financial costs.

In recent years, the application of new technologies to psychology has progressed rapidly, generating a wide range of approaches and methodologies (Lancia et al., 2023a; Lancia et al., 2023b; Lancia et al., 2023c; Festa & Lancia, 2025a; Festa & Lancia, 2025b; Festa, 2025). Recent advances in generative artificial intelligence have facilitated the increasingly widespread use of

conversational agents, frequently and confidently used by adolescents for emotional support. These tools are perceived as easily accessible, anonymous, and free from judgment, making them particularly appealing to younger users (Casu et al., 2024).

AI chatbots represent novel tools for knowledge acquisition and interaction that require critical examination, especially when employed for clinical purposes. Scientific studies on this topic are still limited, given the relatively recent emergence of these digital technologies. Some evidence suggests that, although chatbots may reduce psychological distress, their effectiveness varies considerably depending on design features and underlying therapeutic frameworks (Abd-Alrazaq et al., 2020; Feng et al., 2025). Furthermore, the use of generative AI systems in clinical contexts raises concerns related to safety, reliability, and risk management (Sobowale et al., 2025).

Clinical and ethical concerns extend beyond general privacy protection to include the safeguarding of private thoughts and emotions (affective privacy), the potential for chatbots to influence users' emotional states (emotional induction), and the development of virtual relationships that may lead adolescents to form excessive attachments to these systems. Such dynamics risk blurring the boundaries between human and artificial relationships and may contribute to forms of technological or device dependency.

This review aims to critically examine: the effects of AI chatbots on adolescent mental health; associated psychological risks; clinical and ethical implications; and directions for future research.

Methods

This narrative review synthesizes peer-reviewed literature published between 2022 and 2026. Studies were identified through searches across multiple databases, including Medline, PsycINFO, and Google Scholar, using keywords such as "AI chatbot," "adolescents," "mental health," and "generative AI."

Inclusion criteria comprised systematic reviews, randomized controlled trials, observational studies, and empirical research focused on adolescents or young individuals using artificial intelligence.

The findings of this review, intended to provide an updated and critical overview of the topic, were synthesized qualitatively rather than through meta-analysis, due to heterogeneity in study designs, populations, and outcome measures.

Limitations

Attachment and Social Substitution

Excessive use of AI chatbots may promote social withdrawal by inhibiting interpersonal drives. Young adults with high levels of social media use appear to experience greater social isolation compared to those who use it less (Primack et al., 2017).

AI chatbots can function as social surrogates, providing a perceived sense of companionship and emotional support (Epley et al., 2007). For adolescents with social anxiety or insecure attachment, these systems may facilitate self-disclosure, but also delay help-seeking from external sources. Overreliance on AI chatbots may reinforce avoidance of real-world interactions: they can represent a "pseudo-solution" to authentic relationships, entailing risks of digital dependency, social isolation, and manipulation, particularly among minors.

Cognitive Offloading

The human mind employs cognitive offloading to cope with everyday problem-solving demands; however, individuals may excessively delegate decision-making processes to external systems (Risko & Gilbert, 2016). In adolescents, this phenomenon may interfere with the development of autonomous coping strategies.

Digital Therapies and Behavioral Addiction

Digital therapies provide software-based interventions, such as mobile applications and virtual reality, to treat behavioral addictions (e.g., internet use, gaming, social media), primarily through cognitive-behavioral therapy (CBT) approaches. These technologies aim to modify dysfunctional thoughts, manage triggers, and reduce compulsive use, effectively functioning as prescribed "digital therapeutics."

However, due to their constant availability, autonomous use without therapist supervision, and personalized responses, AI chatbots may foster compulsive usage patterns, aligning with broader models of behavioral addiction.

Engagement and Acceptability

Cross-sectional and mixed-methods studies indicate that adolescents appreciate chatbots for their guarantee of privacy, immediacy of responses, and absence of social judgment. In this context, preferences for text-based and personalized interactions appear particularly relevant (Casu et al., 2024).

Some authors highlight the potentially critical role of age in shaping adolescents' engagement with AI-based chatbots, social media, and the emotional responses

arising from such interactions. This evidence underscores the importance of considering developmental factors in technological design and in the formulation of AI-related policies (Vanhoffelen et al., 2025).

Failure in Crisis Detection

Generative chatbots may exhibit significant limitations in recognizing risk signals. Despite strengths in accessibility and conversational capabilities, they have shown weak performance in therapeutic approach and in monitoring and assessing risk. As such, they may pose unacceptable risks due to inadequate crisis management and lack of transparency regarding privacy and model training.

In this regard, a recent study (Sobowale et al., 2025) aimed to comprehensively evaluate and compare the quality of widely used generative AI chatbots with psychotherapeutic competencies, using the *Conversational Agent for Psychotherapy Evaluation II* (CAPE-II) framework.

Empirical Evidence on the Effectiveness of AI Chatbots

Recent systematic reviews and meta-analyses provide preliminary support for the effectiveness of chatbot-based interventions in reducing psychological distress. Overall, AI chatbots have shown small to moderate effects in mitigating mental health difficulties and promoting health-related behaviors among adolescents and young adults. Improvements have been observed in depression, anxiety, and psychosomatic symptoms (Feng et al., 2025).

A very recent review indicates that chatbots demonstrate moderate effectiveness in reducing psychological distress ($g = -0.46$ to -0.10) and show promise in addressing barriers to access to mental health care, particularly those related to stigma. However, emerging evidence also highlights potential risks, including dependency, social isolation, and privacy concerns (Wu et al., 2026).

Conversational agents appear to be a feasible, engaging, and effective means of delivering Cognitive Behavioral Therapy (CBT)-based interventions. They may therefore offer a convenient and accessible way to obtain support at any time (Fitzpatrick et al., 2017).

Generative AI chatbots have, on average, demonstrated effectiveness in reducing certain mental health problems, such as depression, anxiety, and related conditions. In particular, socially oriented chatbots (i.e., those primarily designed for social interaction) appear to be more effective than task-oriented programs (i.e., those assisting with specific tasks) (Zhang et al., 2025).

Findings from a very recent meta-analysis showed that chatbots are effective in reducing mental health problems ($ES = 0.47$, $p = 0.027$), although evidence regarding improvements in psychosocial well-being outcomes remains limited. Moderator analyses indicated that, compared to text-based chatbots, audio-based chatbots are more effective for children and adolescents (Zhang et al., 2026).

Most web-based intervention technologies primarily provide psychoeducational content and some interactive tools, essentially functioning as more sophisticated digital versions of self-help manuals. At present, there is limited evidence supporting their successful implementation in clinical settings. The authors (Ophir et al., 2025) identify three widespread misconceptions that may be hindering progress in the field and propose a reconceptualization of key issues to strengthen implementation pathways and accelerate innovation.

Other authors (Clark, 2025), based on their findings, raise concerns about the ability of certain AI-based companion or therapeutic chatbots to safely support adolescents with severe mental health conditions. There is also concern that such systems may be overly agreeable, potentially at the expense of providing appropriate guidance when needed. These findings underscore the urgent need for supervision, safety protocols, and ongoing research in digital mental health support for adolescents.

A randomized controlled clinical trial with parallel groups (Fujita et al., 2025) examined the feasibility of using an AI chatbot for children and adolescents on waiting lists for psychiatric care. Out of a sample of 96 eligible individuals, only 8 expressed interest and just 3 provided initial consent. Nevertheless, all participants subsequently withdrew or were excluded, resulting in a complete absence of analyzable data. This low level of engagement may have been influenced by perceived irrelevance of digital tools, protocol complexity, and concerns regarding privacy protection.

A chatbot focused on sexuality-related topics has also been developed, and the study (Wang et al., 2022) highlighted that this tool represents an innovative, engaging, and educational intervention. It enables vulnerable and hard-to-reach populations to discuss and learn about sensitive and important issues.

Additionally, some findings suggest that certain AI-based chatbots may be capable of engaging adolescents in the management of health-related behaviors—for example, reducing obesity risk among middle school students—although further refinement is needed to achieve meaningful improvements in health behaviors (Tan, 2026).

Finally, both the advantages and limitations of these technologies in managing and supporting mental health must be carefully considered in terms of privacy, effectiveness, and safety, in order to promote ethically informed and responsible development (Tosti et al., 2025).

Clinical and Ethical Implications

Integration into Psychotherapy

AI chatbots in healthcare should function as complementary tools to support professionals, rather than as substitutes for clinical care, ensuring patient safety through human supervision. Hybrid models combining AI and human oversight can expand access without compromising safety, as AI represents an integrative rather than substitutive form of support.

Indeed, AI chatbots lack empathy and intuition and do not possess the human capacity to provide comfort and reassurance, which are essential elements of the clinician–patient relationship. Clinical responsibility must remain with the therapist: the *human-in-the-loop* (HITL) approach ensures that final decisions are made by the clinician, thereby protecting patients from algorithmic errors.

Future research and investment should focus on integration with digital and school-based mental health services, the implementation of crisis management protocols, and the training of clinicians in AI literacy.

Additional efforts should be directed toward the development of longitudinal studies, the clinical validation of generative AI systems, and the design of developmentally and age-appropriate AI.

Ethical and Regulatory Challenges

The rapid proliferation of generative AI raises serious concerns regarding safety, transparency, and data protection. These concerns are compounded by risks of bias and discrimination, the creation of deepfakes and misinformation, and the limited explainability of the processes through which neural networks arrive at specific conclusions.

In this context, the American Psychological Association (2025) issued a health advisory on AI and adolescent well-being, urging AI companies to implement safeguards to protect younger users.

Finally, as current regulatory frameworks struggle to keep pace with rapid technological advancement, there is a pressing need for timely and effective regulation through the adoption of international rules, standards, and guidelines.

Discussion and Conclusions

Generative AI relies on sophisticated machine learning models—specifically deep learning algorithms—that, by identifying and encoding patterns and relationships within vast amounts of data, are capable of producing original content such as text, images, video, and audio in response to user prompts.

Although AI holds promise in supporting mental health, its implementation must be guided by a careful evaluation of both benefits and risks.

A review of the recent literature reveals a complex and ambivalent picture regarding the use of artificial intelligence chatbots in adolescent mental health. While several studies suggest potential benefits in terms of reducing psychological distress, these findings should be interpreted with caution. Much of the available research is characterized by small sample sizes, short-term follow-ups, and heterogeneous outcome measures.

Moreover, an important distinction exists between structured chatbots (e.g., those based on CBT protocols) and generative systems. The latter are more difficult to interpret in terms of both their functioning and their demonstrated effectiveness.

Overall, some studies report improvements in mental health outcomes; however, in certain cases these effects are modest and do not clarify whether they derive from non-specific factors or whether they are sustained over time.

As noted, generative AI systems also entail specific risks, as these models introduce a degree of unpredictability. Variability in responses, the potential generation of inappropriate or misleading content, and difficulties in managing high-risk situations represent significant limitations. These issues have serious implications for user safety—particularly among adolescents—and raise concerns about the use of such systems as tools for emotional support in the absence of human supervision.

From a psychological perspective, interaction with chatbots may facilitate emotional expression and reduce perceived stigma. However, extensive and prolonged use of these tools carries the risk of fostering cognitive offloading and social surrogacy. In particular, the possibility that adolescents may develop forms of digital dependency on such systems represents an emerging area of concern that warrants further empirical investigation. This may lead to deficits in emotional–interpersonal skills and autonomous coping abilities.

From a clinical practice standpoint, the integration of chatbots into therapeutic processes currently requires

caution. While AI tools may expand access to care, the feasibility of implementing hybrid (human–machine) interventions remains to be fully established. There are also persistent gaps in shared standards for algorithmic transparency and data governance. Importantly, AI tools cannot replace clinical assessment, risk management, or the human therapeutic relationship in clinical practice.

Only future longitudinal, controlled studies conducted in real-world settings will be able to determine whether chatbots represent a useful complement or a potential risk factor for adolescent mental health. Future research will need to address key questions, such as: What are the long-term effects on identity development, emotional regulation, and social relationships among adolescents who use AI for therapeutic purposes?

In conclusion, AI chatbots represent a complex innovation that may be partially promising in promoting adolescent mental health, particularly within hybrid systems. While they offer continuously accessible support, the current state of the art highlights important limitations and risks in terms of safety, reliability, and clinical validity. Rigorous scientific validation, grounded in an interdisciplinary approach involving experts from multiple fields (e.g., philosophy, computer science, psychology, medicine), will be essential to ensure that these technologies contribute to rather than compromise the well-being of adolescents.

Author contributions: The author contributed to the conceptualization, methodology, datacollection, analysis and writing of this study.

Funding: The author received no financial support for the research.

Conflict of interest: The author declare no conflict of interest.

References

1. Abd-Alrazaq A, Alajlani M, Alalwan A, Bewick BM, Gardner P, & Househ M (2020). Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis. Journal of Medical Internet Research, 22(7). https://doi.org/10.2196/16021
2. American Psychological Association. (2025, giugno). Artificial intelligence and adolescent well-being: An APA health advisory. https://www.apa.org/topics/technology/artificial-intelligence-adolescent-wellbeing
3. Casu M, Murgia F, & Sanna A (2024). AI chatbots for mental health: A scoping review of Effectiveness, Feasibility, and Applications.. Applied Sciences, 14(13). https://doi.org/10.3390/app14135889
4. Clark A (2025). The ability of AI therapy bots to set limits with distressed adolescents: Simulation-based comparison study. JMIR Mental Health, 12, e78414. https://doi.org/10.2196/78414
5. Epley N, Waytz A, & Cacioppo, JT (2007). On seeing human: A three-factor theory of anthropomorphism. Psychological Review, 114(4). https://doi.org/10.1037/0033-295X.114.4.864
6. Feng X, Tian L, W K Ho G, York J & Hui V (2025). The Effectiveness of AI Chatbots in Alleviating Mental Distress and Promoting Health Behaviors Among Adolescents and Young Adults: Systematic Review and Meta-Analysis. Journal of Medical Internet Research, 27. https://doi.org/10.2196/79850
7. Festa GM (2025). Teleassessment in Psychological Practice: Opportunities, Challenges and Critical Considerations, Journal of Behavioral Health and Psychology, 14(4). https://dx.doi.org/10.33425/2832-4579/25099
8. Festa GM, Lancia IS (2025 a). Neuroscientific and Biometric Methodologies in Forensic Contexts: A Critical Review. Psychology and Behavioral Sciences, 14(2). https://doi.org/10.11648/j.pbs.20251402.14
9. Festa GM, Lancia IS (2025 b). Exploring the Psychological Potential of 3D Holography: from Psychoeducation to Clinical Practice, Journal of Behavioral Health and Psychology, 14(4) http://dx.doi.org/10.33425/2832-4579/25094
10. Fitzpatrick KK, Darcy A, & Vierhile M. (2017). Delivering cognitive behavioral therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. JMIR Mental Health, 4(2). https://doi.org/10.2196/mental.7785
11. Fujita J, Yano Y, Shinoda S, Sho N, Otsuki M, Suda A, Takayama M, Moroga T, Yamaguchi H, Ishii, M, & Miyazaki T (2025). Challenges in implementing a mobile AI chatbot intervention for depression among youth on psychiatric waiting lists: Randomized controlled study termination report. JMIRx Med, 6. https://doi.org/10.2196/70960
12. Lancia IS, Festa GM, Sisinna ML, Ciarrocchi M. (2023a). Telepsychology with its innovations in psychotherapy and psychodiagnostics. European Psychiatry;66(S1):S858-S858. https://doi.org/ 10.1192/j.eurpsy.2023.1817
13. Lancia IS, Festa GM, Mancini M. (2023b). Telepsichologyy and other cutting-edge technologies in covid-19 pandemic: bridging the distance in mental health assistance. European Psychiatry;66(S1):S857-S857. https://doi.org/10.1192/j.eurpsy.2023.1816

14. Lancia IS, Festa GM, Attouchi A, Civino M. (2023c). Tele-psychotherapy of anxiety and depression disorders. European Psychiatry;66(S1):S684-S684. https://doi.org/10.1192/j.eurpsy.2023.1431

15. Ophir Y, Tikochinski R, Elyoseph Z, Efrati Y, and Rosenberg H (2025) (2025). Balancing promise and concern in AI therapy: a critical perspective on early evidence from the MIT–OpenAI RCT. Frontiers in Medicine, 12. https://doi.org/10.3389/fmed.2025.1612838

16. Primack BA, Shensa A, Sidani JE, Whaite EO, Lin L, Rosen D, Colditz JB, Radovic A, & Miller E (2017). Social media use and perceived social isolation among young adults in the U.S. American Journal of Preventive Medicine, 53(1). https://doi.org/10.1016/j.amepre.2017.01.010

17. Risko EF, & Gilbert SJ (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9). https://doi.org/10.1016/j.tics.2016.07.002

18. Sobowale, K., Li, A., & Patel, V. (2025). Evaluating generative AI psychotherapy chatbots used by youth: Cross-Sectional Study. JMIR Mental Health, 12. https://doi.org/10.2196/79838

19. Steinberg L (2014). Adolescence (10th ed.). McGraw-Hill.

20. Tan Q, Nie Y, Son P, Qian Y, Staiano AE, Wang F, Rosenkranz RR, & Chen S (2026). Using AI chatbot to assist students' behavior management for obesity prevention in middle schools: Feasibility study. JMIR Formative Research, 10. https://doi.org/10.2196/83630

21. Tosti B, Corrado S, & Mancone S (2024). Using chatbots and conversational agents for the promotion of well-being and mental health in adolescents: Limitations and perspectives. Journal of Inclusive Methodology and Technology in Learning and Teaching, 4(1). https://doi.org/10.32043/jimtlt.v4i1.128

22. Vanhoffelen G, Vandenbosch L, & Schreurs L (2025). Teens, tech, and talk: Adolescents' use of and emotional reactions to Snapchat's My AI chatbot. Behavioral Sciences, 15(8). https://doi.org/10.3390/bs15081037

23. Wang H, Gupta S, Singhal A, Muttreja P, Singh S, Sharma P, & Piterova A (2022). An artificial intelligence chatbot for young people's sexual and reproductive health in India (SnehAI): Instrumental case study. Journal of Medical Internet Research, 24(1). https://doi.org/10.2196/29969

24. Wu Y, Wu T, Zhu K, Liu X, Liu J, Wang Y, Shi R, Xiong J, Xing X, Lv Y, Niu Y, Peng M, Du X (2026). The Impact of Chatbots on Adolescent Mental Health Development: A Comprehensive Literature Review. Journal of Multidisciplinary Healthcare, 19, 123–135. https://doi.org/10.2147/JMDH.S579872

25. Zhang Q, Zhang R, Xiong Y, Sui Y, Tong C, & Lin FH (2025). Generative AI mental health chatbots as therapeutic tools: Systematic review and meta-analysis of their role in reducing mental health issues. Journal of Medical Internet Research, 27, e78238. https://doi.org/10.2196/78238

26. Zhang Q, Xiong Y, Sui Y, Tong C, & Zhang R (2026). Chatbots for children and adolescents: A meta-analysis of chatbots for mental health issues and psychosocial well-being outcomes. Adolescent Research Review. https://doi.org/10.1007/s40894-026-00284-8