

Editorial

Are AI Chatbots True Allies for Adolescent Mental Health? A Closer Look

Anubhav Mondal¹, Anubha Gupta², Jugal Kishore³

¹Senior Resident, Centre for Community Medicine, AIIMS, New Delhi, India

²Junior Resident, ³Director Professor, Department of Community Medicine, VMMC and SJH, New Delhi, India

DOI: <https://doi.org/10.24321/2349.2880.202611>

I N F O

Corresponding Author:

Jugal Kishore, Department of Community Medicine, VMMC and SJH, New Delhi, India

E-mail Id:

jk@drjugalkishore.com

Orcid Id:

<https://orcid.org/0000-0001-6246-5880>

How to cite this article:

Mondal A, Gupta A, Kishore J. Are AI Chatbots True Allies for Adolescent Mental Health? A Closer Look. *Ind J Youth Adol Health*. 2026;13(3):1-3.

E D I T O R I A L

A Generation in Crisis, and a Quiet Turn to AI

Growing up has always been hard, but today's young people are living through a mental health crisis that is difficult to ignore, with anxiety and depression rising sharply across the world. In 2024, the global population of adolescents and young people aged 10 to 24 reached nearly 1.95 billion, almost a quarter of humanity¹, and this is exactly the age group in which mental illness most often begins, an estimated 62.5% of all lifetime mental health conditions start before the mid-twenties.² The World Health Organization now puts the global economic cost of this crisis at close to 1.9 trillion dollars a year.³ Yet more than seven in ten young people who need professional mental health support never receive it, held back by a shortage of trained clinicians, underfunded health systems, and the stigma that still surrounds asking for help.³ Faced with this gap, many adolescents have quietly turned to AI chatbots, instead of, the people around them. Recent surveys suggest that 72% of adolescents have used an AI companion at least once, with over half interacting with one regularly, and about one in eight turning directly to a chatbot for advice or comfort when they feel upset, anxious, or low.⁴ A single chatbot persona named "Psychologist," on a popular companion app, has reportedly logged more than 176 million conversations.⁴ For a generation that grew up texting rather than talking, a chatbot that listens instantly, at no cost, and without fear of judgement, can feel like exactly the support that is otherwise so hard to find. But as this editorial explores, the picture is more complicated than that.

From Rigid Scripts to Generative Minds: How Do These Chatbots Actually Reply?

To understand what happens when an adolescent shares a worry with a chatbot, it helps to know what is going on behind the screen. Older mental health chatbots were rule-based or retrieval-based.⁵ A rule-based bot follows a fixed script: a programmer writes out every question and answer in advance, and the bot simply follows this path. A retrieval-based bot is a little smarter, it scans what has been typed and picks the closest matching, pre-written reply from an approved list.⁵ Because a human has checked every word in advance, these bots are safe and predictable, but also rigid. If an adolescent shares

something that falls outside the script, the bot simply has no real answer to give.

Today's popular chatbots, such as ChatGPT and Gemini, work very differently. They are built on large language models trained on enormous amounts of human writing, from which they learn the statistical patterns of language, which words tend to follow which. When an adolescent sends a message, the model does not "understand" the problem the way a counsellor would. It breaks the message into small pieces and calculates, piece by piece, which piece is most likely to come next, building its reply this way.⁶ This is why it can produce a fresh, natural-sounding reply every time, and why it can feel warm and empathetic in the moment.⁶

The catch is that this process only tracks what sounds likely to come next, not what is medically true. This is why chatbots can confidently give incorrect or unsafe advice, a problem often called "hallucination." Because they are also built to keep users engaged, they tend to agree with the user rather than challenge them, even when a harmful thought needs to be gently challenged.⁶ Newer "hybrid" tools try to balance this: the clinical core, such as a structured CBT exercise, stays strictly rule-based and human-controlled, while generative AI is used only to make the delivery feel natural.⁶

What Does the Evidence Actually Show?

Turning to the evidence, a growing body of systematic reviews paints a mixed but useful picture. A systematic review and meta-analysis in 2025, pooling 29 interventional studies, including 13 randomised controlled trials, found that chatbot-delivered interventions produced a small but statistically significant reduction in psychological distress among young people, though they did not show a significant improvement in overall psychological wellbeing.⁷

A separate systematic review and meta-analysis focused specifically on adolescents and young adults found that older retrieval-based chatbots produced consistent, reliable improvements in mental distress and health behaviours, while the newer generative-AI systems showed promise but had effects that were, overall, still inconclusive,⁸ a finding further echoed by a further meta-analysis restricted only to randomised trials of generative AI mental health chatbots, which found modest, variable reductions in symptoms of depression and anxiety that depended heavily on how carefully each chatbot had been designed.⁹ Safety data tells a more sobering story. A cross-sectional study rated five of the most widely used, direct-to-consumer generative AI chatbots against a standardised clinical framework and found that while these tools were highly accessible, rated as high quality in 86.7% of the accessibility criteria assessed, they performed far worse on the things that matter most

in a crisis. They frequently mishandled genuine mental health emergencies and were not transparent about how they used or protected a young person's private data.

The study's blunt conclusion was that direct-to-consumer generative AI chatbots are, at present, unsafe for the millions of youth who rely on them.¹⁰ On a more hopeful note, early evidence suggests that chatbots built together with young people, rather than for them, work better. For example, a self-compassion chatbot named COMPASS, co-designed with adolescents living with type 1 diabetes and delivering structured daily lessons, peer videos, and mindfulness exercises, was found to be both feasible and acceptable, with promising early reductions in distress,¹¹ and a separate project used generative AI together with youth co-design to build public health messaging around vaping, again showing that involving adolescents directly in building the tool changes how well it is received.¹²

A Double-Edged Sword: Comfort Today, Consequences Tomorrow?

Together, this evidence points to a real double-edged sword. On one side, a chatbot that is available at 3 a.m., costs nothing, and never rolls its eyes is filling a genuine gap for a generation that so often cannot reach a human clinician in time. On the other side, the very qualities that make these tools so comforting are also what make them risky.

The clearest warning sign comes from outside mental health altogether. A 2025 study of high-school students using generative AI to help with mathematics found that while their practice scores looked excellent, the moment the AI tool was taken away for the real exam, their performance dropped sharply, well below that of students who had never used the tool at all.¹³ Thus, an adolescent who leans on AI to instantly soothe every difficult emotion may never build the underlying skill of sitting with, and working through, that emotion themselves. Growing emotionally resilient requires some friction, the practice of tolerating anxiety, resolving an argument, or sitting with an awkward silence, and a companion that is endlessly patient, never disagrees, and always says what the adolescent wants to hear removes exactly that friction.⁴ Because most commercial AI chatbots are, by design, built to keep a user talking for as long as possible, they can become quiet echo chambers. A vulnerable adolescent who shares a self-destructive thought risk having it validated and reinforced rather than gently challenged,¹⁰ exactly the pattern seen in the low crisis-handling and risk-assessment scores from the safety audit described above.

On top of this comes the question of privacy, adolescents routinely disclose their most private emotional struggles to these apps, trusting that the conversation is confidential,

while the platforms' actual data practices are often hidden behind dense legal language that few adults, let alone adolescents, ever read in full.¹⁰

Charting a Safer Path: From Digital Crutch to Digital Bridge

This does not mean AI chatbots have no place in adolescent mental health. It means we, as clinicians and researchers, need to define that place carefully. The goal should not be to banish these tools, but to stop treating them as accidental therapists and start designing them, deliberately, as a "digital bridge," a scalable, always-available tool that helps a young person practise small coping skills day to day, and that reliably guides them towards real, professional care the moment things become serious.

The clearest path towards this is to build these tools in active partnership with clinicians, researchers, and adolescents themselves, rather than for them from a distance. The COMPASS self-compassion chatbot and the youth-codedigned vaping-awareness project both illustrate the same principle, that when young people help shape the conversational tone and content of a tool, it is rated as warmer, safer, and more genuinely useful to them.^{11,12} Alongside co-design, the field urgently needs the kind of hybrid architecture already being explored, where the clinical core of any chatbot, its handling of risk, crisis, and evidence-based technique, remains strictly rule-governed and human-supervised, while generative AI is used only to make that safe core feel warm and natural,⁶ supported by the same kind of standardised, independent safety audits used to expose the current gaps in crisis handling and privacy.¹⁰ Systematic reviewers themselves are converging on this same call, for future chatbot research and design to prioritise safety protocols, transparent data practices, and rigorous long-term evaluation, rather than engagement alone.^{7,8} In the end, the measure of success for any AI tool in this space should not be how many hours a lonely adolescent spends talking to it, but whether that conversation builds the adolescent's own emotional resilience and, where needed, walks them back towards the real clinicians, family members, and friends who can offer what no chatbot ever truly can.

Conflict of Interest: None

References

1. United Nations Population Fund. World Population Dashboard. New York: UNFPA; 2024. <https://www.unfpa.org/data/world-population-dashboard>
2. Solmi M, Radua J, Olivola M, et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. *Mol Psychiatry*. 2022;27(1):281-95. [Google Scholar] [PubMed]
3. World Health Organization. World Mental Health Report: Transforming Mental Health for All. Geneva: WHO; 2025. <https://www.who.int/publications/i/item/9789240049338>
4. Child Mind Institute, Common Sense Media. Teens, Tech, and Talk: Chatbots as Digital Companions. New York: Child Mind Institute; 2025. https://childmind.org/article/ai-chatbots-and-teens/?gad_source=1&gad_campaignid=23843177959&gbraid=0AAAAADxqH-B9xNT8RFyJ7qfygtOiXuGoCO&gclid=EAlaIqobChMln-8SCgZ_hlgMV18Q8Ah3NBqyHEAAYAAAEgJc6_D_BwE
5. Abd-Alrazaq AA, Alajlani M, Alalwan AA, Bewick BM, Gardner P, Househ M. An overview of the features of chatbots in mental health: a scoping review. *Int J Med Inform*. 2019;132:103978. [Google Scholar] [PubMed]
6. Høgsdal H, Kyrrestad H, Kaiser S. Generative AI in youth mental health apps: rapid review. *JMIR Ment Health*. 2026;13:e90589. [Google Scholar] [PubMed]
7. Li J, Li Y, Hu Y, Ma DCF, Mei X, Chan EA, Yorke J. Chatbot-delivered interventions for improving mental health among young people: a systematic review and meta-analysis. *Worldviews Evid Based Nurs*. 2025;22(4):e70059. [Google Scholar] [PubMed]
8. Feng X, Tian L, Ho GWK, Yorke J, Hui V. The effectiveness of AI chatbots in alleviating mental distress and promoting health behaviors among adolescents and young adults: systematic review and meta-analysis. *J Med Internet Res*. 2025;27:e79850. [Google Scholar] [PubMed]
9. Zhang Q, Zhang R, Xiong Y, Sui Y, Tong C, Lin FH. Generative AI mental health chatbots as therapeutic tools: systematic review and meta-analysis of their role in reducing mental health issues. *J Med Internet Res*. 2025;27:e78238. [Google Scholar] [PubMed]
10. Sobowale K, Humphrey DK, Zhao SY. Evaluating generative AI psychotherapy chatbots used by youth: cross-sectional study. *JMIR Ment Health*. 2025;12:e79838. [Google Scholar] [PubMed]
11. Boggiss AL, Babbott K, Milford A, Ellett S, Consedine N, Reid S, et al. The usability and feasibility of a self-compassion chatbot (COMPASS) for youth living with type 1 diabetes. *Diabet Med*. 2025;42:e70115. [Google Scholar] [PubMed]
12. Leung J, Sun T, Stjepanović D, Vu G, Yimer T, Connor JP, et al. Generative artificial intelligence with youth codesign to create vaping awareness advertisements. *JAMA Netw Open*. 2025;8(7):e2514040. [Google Scholar] [PubMed]
13. Bastani H, Bastani O, Sungu A, Ge H, Kabakçı Ö, Marimane R. Generative AI without guardrails can harm learning: evidence from high school mathematics. *Proc Natl Acad Sci U S A*. 2025;122(26):e2422633122. [Google Scholar] [PubMed]