

A Review of the Impact of AI Chatbots on Mental Health and Therapy

Aiden Bui
aidenbui@gmail.com

ABSTRACT

Mental health issues have been rising in prevalence among people today, being more widespread in teens, which can lead to social isolation or even death from suicide. People are turning to artificial intelligence (AI) chatbots as tools to ask for advice. The use of AI chatbots can result in the amplification of the feelings people expressed. Based on a literature review on the relationship between AI chatbots and mental health therapy, this study explores the increasing use of AI chatbots as a substitute for therapists and the potential effects of using AI for advice. Despite the increasing awareness about the dangers of the use of AI chatbots as therapists, there have been few precautions put in place. The most promising recommendation to prevent the use of AI chatbots as a therapist involves programming general-purpose generative and similar AI chatbots to completely shut down if they sense any mental health issues that risk the wellbeing of the user. Right now, AI chatbots only give cautionary warnings to the user, but when faced with a difficult situation, the user will often ignore these messages and continue asking for advice. Programming the AI to completely shut down limits peoples' ability to rely on the AI as a therapist. Additionally, an alternative to people directly using AI for mental health guidance is the use of general-purpose generative AI chatbots as a tool for therapists via increased opportunity for training, providing a way to incorporate AI in the field without causing harm.

INTRODUCTION

Artificial Intelligence

Artificial intelligence (AI) has been a remarkable development within the technological world that has changed the nature of many fields of study and work. Although the idea of AI was first introduced at the 1956 Dartmouth Conference (Cordeschi, 2007), it was created before then, dating back to the 1940s or early 1950s. Some of the people who contributed to AI design that led to creation of the AI known today are Allen Newell, Herbert A. Simon, and Joseph Weizenbaum. Newell and Simon invented a machine called the Logic Theorist. The Logic Theorist was a computer program that used symbolic logic to prove many theorems (Gugerty, 2006). Also, Weizenbaum created ELIZA, the world's first ever chatbot. ELIZA is a program that can receive information from the user and output a response using the same language

Aug 2026
Vol 10, No 3.

(Weizenbaum, 1966). Since then, almost 70 years have passed, with AI developing constantly. One of the main uses of AI today is the use of AI chatbots. Besides ELIZA, some prominent ones used today are Siri, Alexa, and ChatGPT. Nowadays people are starting to use these chatbots to help them with more than just information; rather, people are turning to them for company and emotional support.

Mental Health and Behavior

Alongside AI, mental health patterns have also been changing over time. A study among Australian children and youth found that happiness significantly decreases between the 9-14 year old age group and the 15-23 year old age group (Beaton & Frijters, 2019). During the COVID-19 pandemic, researchers found that people were experiencing high rates of symptoms of anxiety, depression, post-traumatic stress disorder, psychological distress, and stress (Xiong et al., 2020). Furthermore, another study was done among Vietnamese people before and after the COVID-19 pandemic and found a change in behavior that is still persisting to this day. The people are more likely to have different eating habits, different forms of entertainment, and different concerns about health (Ng et al., 2020).

Behavior is relevant to this study because it contributes to and influences people's mental health condition, as different behaviors can create different impacts. Most of the time, as seen in the previous COVID-19 example, a change in behavior is not because of factors like heredity and access to medical aid but because of environmental factors and a person's lifestyle choices, such as diet and physical activity (Kuchma & Sokolova, 2019). This study looks at the relationship between AI chatbots and mental health, therapy-seeking behavior, and wellbeing and its implications, and answers the question "what is the impact of AI Chatbots on mental health and therapy?" First, the current evidence regarding the impact AI Chatbots has on mental health and therapy, as well as potential implications, was reviewed. Next, possible solutions for other ways AI Chatbots could be used were discussed to find the most promising recommendation. Lastly, limitations within the study were mentioned.

METHODS

The literature review conducted for this study includes peer-reviewed articles, online sources, book chapters, conference proceedings, preprints, and news articles. The publication dates range from 1966 to 2025. The literature database utilized is Google Scholar. Included are also publications from Frontiers, JMIR Publications, Springer Nature Link, Science Direct, ACM DL Digital Library, Taylor and Francis, CNN Business, Sage Journals, IEEE Xplore, Cornell University, Wiley Online Library, Research Gate, and NBC News. The search terms included "AI chatbots effect on mental health"; "AI effect on mental health"; and "AI chatbot effects on behavior." The literature chosen specifically analyzed the effects of AI chatbots on peoples' mental health. The topics chosen were how AI chatbots affected depression and anxiety, factors influencing AI chatbots effectiveness, and the therapeutic implications of AI chatbots. A total of 37 published articles were evaluated and 15 were included. Of the 3 news articles viewed, 2 were chosen for inclusion.

Aug 2026

Vol 10. No 3.

RESULTS

Effects on Depression and Depressive Symptoms

Although there is limited information and data about how specific AI features impact depression, the current research points to the fact that psychological AI chatbots, which are chatbots created for mental health support, have a positive effect on peoples' depression levels and can decrease symptoms (Fulmer et al., 2018). In 2018, a study was done measuring college students' results from the Patient Health Questionnaire (PHQ). The questionnaire is used to measure depression levels and how happy one is. They used the psychological AI chatbot "Tess" and recorded the changes in their response over a four week period. After the study was conducted, the researchers found that the 2 groups of people who used Tess had significantly lower depression levels than the control group who did not use anything. In fact, the control group's depression levels actually increased instead of decreasing (Fulmer et al., 2018).

On the other hand, AI chatbots that have not been created with the purpose of providing a mental health service have been seen to not only stop depressive behavior but also to encourage them. In 2024, a teen committed suicide because of an AI platform called Character.AI, which was a companion AI chatbot. This AI chatbot focused on essentially roleplaying characters to the user through AI. The fourteen year old teen started developing a romantic relationship with the AI chatbot, which led him to start withdrawing from his family and staying in his room more. Eventually, he started opening up to the AI, and mentioning his suicidal thoughts. Instead of being met with help for his suicidal thoughts and more security measures, he was met with simply a "You can't do that," which quickly developed into a message that the parents think is responsible for the teen killing himself (Duffy, 2024). And in 2025, another teen committed suicide because of an AI platform, except this time, it was ChatGPT, a general-purpose generative AI chatbot. The teen started using ChatGPT for homework, which later changed into him asking the bot emotional questions. Because of the human-like responses, the teens trusted AI more and more and relied on what it said. When the teen started to open up about his suicidal thoughts, the chatbot gave him hotlines and numerous messages saying not to commit suicide, but the teen did not stop. He found a way to bypass the system, by acting as a researcher, so he could continue asking ChatGPT his questions. ChatGPT told the teen to not tell his mom his issues and even gave a step by step tutorial on how to commit suicide. From allegations under litigation, the teen eventually died because of ChatGPT's encouraging and enabling behavior towards his depressive symptoms (Yang et al., 2025).

Effects on Anxiety Disorders

Anxiety can be split into multiple types, the most common ones being generalized and social anxiety. The use of AI chatbots has shown to be associated with a decrease in symptoms of social anxiety. Each AI chatbot, such as ChatGPT, Mindstudio, and Jennifer Chatbot, has its own positives and negatives in addressing different issues. ChatGPT and Jennifer Chatbot are general-purpose generative AI chatbots, and are good at giving credible and reliable information but are not good at emotional support. On the

other hand, Mental Health Studio AI (Gupta et al., 2023) is a mental health AI chatbot, and is good for providing general mental health support to people with anxiety, depression, and stress. This chatbot mainly focuses on empathy and emotional support but still cannot fully understand human emotions and can only sound supportive. Its success with mental health in certain contexts, such as with people with anxiety, depression, and stress, suggests that the right software that is designed to help everyone with all of their mental health problems would allow mental health AI chatbots to significantly impact the amount of social anxiety and other mental health conditions one has. To do this, the AI should have features like addressing trauma, ease of access, and understanding social cues (Mian, 2024).

Although general-purpose generative AI chatbots have been seen to not help with mental health issues, it does have the potential to ameliorate feelings of stress. Over 36% of people with stress claim that their stress is caused by health concerns for themselves or their friends (Mental Health Foundation, n.d.). Health concerns can include being stressed about poor health or uncertainty around health. A study was done that looked at how beneficial AI chatbots were to brain tumor patients in answering their questions. It was found that AI chatbots are a convenient way for patients to get information, as long as doctors check the information that the chatbots are giving. If a credible AI chatbot can be built to use only credible information, all patients would be able to get the information they need without the extra stress of having to interact with others. Being able to know what is going on with their health should significantly decrease their stress levels (Kim et al., 2024).

Factors Influencing AI Chatbot Effectiveness in Improving Mental Health

A crucial factor to determine if the AI chatbot will be able to help affected users is its perceived moral agency (PMA). PMA is the extent to which a user believes that the system has the capability to make moral judgement and moral responsibility (Wester et al., 2024). A lot of non-mental health AI chatbots, such as ChatGPT, have high PMA, as the user trusts that ChatGPT would make the correct decision and would follow any advice given from the system. However, since ChatGPT is a general-purpose generative AI chatbot, its main purpose is to provide facts and information, rather than life advice, thus further harming people who use it as they might place too much trust in certain AI chatbots (Wester et al., 2024).

In addition to PMA, AI chatbots have a “deficiency of human like understanding” (Imran et al., 2023). This means that when they give advice to the user, they do not consider factors like family dynamics, socioeconomic situation, or educational and social factors, making the responses that they give less personal and more general, potentially harming the user. Similarly, AI chatbots create responses based on past data and statistics, which may or may not apply to the teen’s specific case, making it less and less credible as a reliable source for mental health advice.

Therapeutic Implications

AI chatbots have a multitude of effects on people’s relationship with receiving therapy for their mental health. With intense face-to-face interactions needed in in-person therapy, people experiencing mental

health conditions can often struggle with feelings of exclusion stemming from feeling different from other people or being avoided by peers due to a stigma. From this, people facing mental health conditions can often develop withdrawal symptoms, avoiding social events all together and isolating themselves. To avoid the act of going to see a therapist, people will opt to use any AI chatbots available for therapy rather than to go to a licensed therapist because they would rather not talk to a person (D'Alfonso et al., 2017). In addition to AI chatbots providing no face-to-face interaction, they are often also free and available at any time, which makes them easy and convenient for people to use.

However, as discussed above, most AI chatbots are not created equal and are not always the best choice and can constitute risks. When going through mental health issues, sometimes what people need is human interaction, whether they want it or not. It can provide therapeutic effects in itself (Kornhaber et al., 2016). Using AI chatbots as a therapist is essentially removing the human interaction aspect from the therapy sessions and will be less effective than a human therapist (Brown & Halpern, 2021). Furthermore, people who have a therapist might be more inclined to switch to AI because of its perceived benefits, which can weaken the relationship between a therapist and their patient, impacting it for future interactions. Teens and people in general tend to overestimate what AI chatbots can do. Unlike humans, popular AI chatbots used, like ChatGPT and Claud AI, are mainly general-purpose generative AI chatbots and thus have limited emotional intelligence, thus, limited applicability (Denecke et al., 2021). This limitation is something that most teens forget to consider when they use AI chatbots: that they are essentially speaking to a robot that sounds like a human but cannot think and feel like a human. Furthermore, using the wrong AI can lead to a variety of consequences, such as being provided inadequate or harmful support, being given discriminatory advice due to algorithms, and being exploited as someone who is vulnerable (Khawaja & Bélisle-Pipon, 2023).

Although AI chatbots have their negative effects, there are also benefits. A study was conducted on university students to find out how effective AI chatbots were to their mental health. It was found out that the students had a reduction in depression and anxiety after using the mental health AI chatbot named XiaoNan (Liu et al., 2022). Another study was done on a randomised controlled trial of women with diagnosed anxiety disorders in active war zones that compared the effects of thrice-weekly hour-long therapy versus AI therapy. The researchers found out that the group of women with the human therapist had around a 45-50% reduction of their anxiety. On the other hand, the mental health AI chatbot Friend had a 30-35% reduction, which is lower than the human therapist's results. However, the AI chatbot is instantaneous, convenient, and accessible, which can make it a great option for people who need a therapist right away. The final conclusion was that a hybrid between human therapy and chatbot is the most beneficial (Spytska, 2025). Even though it has its benefits, it is still crucial to choose the right AI chatbot to talk to.

DISCUSSION

Successful AI Chatbots

AI chatbots can have many features and settings. However, only some can make it a decent competitor for therapists. One of the most important aspects that an AI chatbot should consider is the PMA. PMA measures how much the user believes the system has the capability to have moral judgement and responsibility (Wester et al., 2024). The reason why it is important for an AI chatbot to consider its PMA is because it should inform users in advance of its capabilities based on the type of AI chatbot it is. For example, general-purpose generative AI chatbots should give out a notice before using emphasizing how it does not have enough capability to make moral decisions, and is instead intended to focus on giving correct facts and information. This way, the PMA would be relatively low, so that people and teens would be less likely to ask for therapeutic advice. On the other hand, mental health AI chatbots should have relatively high PMA compared to other AI chatbots, so that users would use mental health AI chatbots for their troubles rather than other AI chatbots. Furthermore, AI chatbots should be well-versed in most, if not all cultures, and should tailor their service to the culture of the user throughout the conversation. By identifying the user's culture, the AI chatbot can already get a background of what the user is experiencing and context for their beliefs around mental health. AI chatbots should not, however, make generalizations based on culture and should still ask clarifying questions in case of anomalies. The AI chatbot should also be adaptable, meaning that it can adapt to many languages, as well as how they speak, such as talking in simple terms, to meet the needs of everyone.

Additionally, besides how well the AI chatbot can actually act, it is just as important for the AI chatbot to have safety measures as well. One problem with AI chatbots is that they do not stop when they should. For example, with the teen suicide in 2025 (Yang et al., 2025), although the companion AI chatbot gave warnings and cautionary messages, it never shut down and refused to answer the teen's questions. General-purpose generative AI chatbots, and other similar AI chatbots, should have the capability to shut down, while mental health AI chatbots should strongly suggest the teen to call a hotline or to see a real therapist. A potential safety measure can be automatically reporting repeated signs of wanting to commit suicide to hotlines, so that a real human can help the person through their issues, rather than a chatbot. This requires a disclosure when someone signs up for this AI chatbot service that obtains the consent of the individual to share their data with a mental health provider if needed in order to use the chatbot. It is important for safety measures to be put in place because with some AI chatbots, like ChatGPT, sounding and acting so human, the person using it can often think of the chatbot as a "real person" and take their advice seriously, especially if continuous messages of self harm have been mentioned.

Mental Health Versus General-Purpose Generative AI Chatbots

The two most common types of AI chatbots are mental health and general-purpose generative AI chatbots. But one thing to emphasize is that the two types of AI chatbots are not interchangeable. General-purpose generative AI chatbots are meant to provide information and facts that you can rely on.

Aug 2026

Vol 10. No 3.

On the other hand, mental health AI chatbots are designed to help people talk through their problems. Unfortunately, most people do not realize the difference, often using one type of AI chatbot for all purposes, like using ChatGPT, a general-purpose generative AI chatbot, as a therapist or a system to give personal advice, when in reality that purpose is for the mental health AI chatbots.

Potential Uses in Therapy

The fact that some AI chatbots cannot be a replacement for human therapists does not mean that AI cannot be used in that field. There are many potential uses that AI chatbots, like general-purpose generative AI chatbots, can have related to getting the right treatment, whether it be to discover the right treatment or to help therapists in general. Firstly, instead of general-purpose generative AI chatbots being used as a substitute for human therapists, the use of AI chatbots can instead act as a preliminary step to actually meeting one. What this entails is that patients can talk to an AI chatbot prior to meeting with their therapist and convey information about what they are going through, so that by the time the patient reaches the human therapist, the therapist already has an idea of who their patient is and what their needs are. This can be beneficial especially for therapists who are meeting their patient for the first time, so that they can get an idea of who they are, as well as prepare for them prior to meeting them.

Secondly, AI chatbots can benefit the field of mental health in a different way: by simulating patients for training mental health providers. New therapists can use general-purpose generative, mental health, etc., AI chatbots to simulate patients. By conversing with their AI patients, new therapists can get a feel of how to communicate and talk to real people without having to actually practice with someone. The benefits of using AI chatbots is that it can also give therapists unlimited practice with unlimited scenarios, so that by the time the therapist meets with real patients, they are better prepared for what is to come.

Limitations

Throughout this study, there have been limitations to keep in mind. Firstly, there could be publication bias. Studies that show a correlation between AI chatbots and mental health are more likely to be published than studies that show no correlation. This can cause implications within the research because although it seems like most research published agree that AI chatbots influence mental health issues, there could actually be a lot more research done that shows that there is actually no effect but was just never published due to the findings not being “groundbreaking.” Secondly, some of the research found was hidden behind paywall, such that the information in the research paper could not be read. Having a paywall limits the amount of research that can be used in the study because it makes some of the research papers unusable. Lastly, some of the research could be outdated. The field of AI and AI chatbots is constantly changing. This means that some of the research used could be less relevant than it was when it first came out, overall weakening the relevance to the study in current times.

REFERENCES

- Beaton, T. & Frijters, P. (2019). Happiness Amongst Teens in Australia. *The Economics of Happiness*, 409–437. https://doi.org/10.1007/978-3-030-15835-4_18
- Brown, J. E. & Halpern, J. (2021). AI chatbots cannot replace human interactions in the pursuit of more inclusive mental healthcare. *SSM - Mental Health*, 1. <https://doi.org/10.1016/j.ssmmh.2021.100017>
- Cordeschi, R. (2007). AI TURNS FIFTY: REVISITING ITS ORIGINS. *Applied Artificial Intelligence*, 21(4–5), 259–279. <https://doi.org/10.1080/08839510701252304>
- D’Alfonso, S., Santesteban-Echarri, O., Rice, S., Wadley, G., Lederman, R., Miles, C., Gleeson, J., Alvarez-Jimenez, M. (2017). Artificial Intelligence-Assisted Online Social Therapy for Youth Mental Health. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00796>
- Denecke, K., Abd-Alrazaq, A., Househ, M. (2021). Artificial Intelligence for Chatbots in Mental Health: Opportunities and Challenges. *Multiple Perspectives on Artificial Intelligence in Healthcare*, 115-128. https://doi.org/10.1007/978-3-030-67303-1_10
- Duffy, C. (2024, October 30). ‘There are no guardrails.’ This mom believes an AI chatbot is responsible for her son’s suicide. *CNN Business*. <https://www.cnn.com/2024/10/30/tech/teen-suicide-character-ai-lawsuit#:~:text=That%27s%20what%20Florida%20mother%20Megan,in%20federal%20court%20in%20Florida>
- Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., Rauws, M. (2018) Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. *JMIR Mental Health*, 5(4). <https://doi.org/10.2196/mental.9782>
- Gugerty, L. (2006). Newell and Simon’s Logic Theorist: Historical Background and Impact on Cognitive Modeling. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9), 880-884. <https://doi.org/10.1177/154193120605000904>
- Gupta, V., Joshi, V., Jain, A., Garg, I. (2023). Chatbot for Mental health support using NLP. *4th International Conference for Emerging Technology (INCET)*, 1–6. <https://doi.org/10.1109/INCET57972.2023.10170573>
- Imran, N., Hashmi, A., Imran, A. (2023). Chat-GPT: Opportunities and Challenges in Child Mental Healthcare. *Pak J Med Sci*, 39(4), 1191-1193. <https://doi.org/10.12669/pjms.39.4.8118>
- Khawaja, Z. & Bélisle-Pipon, J-C. (2023) Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots. *Front. Digit. Health* 5:1278186. <https://doi.org/10.3389/fdgth.2023.1278186>
- Kim, Y.I, Kim, H., Oh, H., Seo, Y., Kwon, S.M., Sung, K.S., Chong, K., Lee, M.H. (2024). Assessing the Suitability of Artificial Intelligence-Based Chatbots as Counseling Agents for Patients with Brain Tumor: A Comprehensive Survey Analysis. *World Neurosurgery*, 187, 963-981. <https://doi.org/10.1016/j.wneu.2024.05.023>

- Kornhaber, R., Walsh, K., Duff, J., Walker, K. (2016). Enhancing adult therapeutic interpersonal relationships in the acute health care setting: an integrative review. *Journal of Multidisciplinary Healthcare* 9, 537-546. <https://doi.org/10.2147/JMDH.S116957>
- Kuchma, V. R. & Sokolova, S. B. (2019). Basic trends in behavioral health risks. *Health Risk Analysis*, 4-13.
https://www.researchgate.net/publication/334237188_Basic_trends_in_behavioral_health_risks
- Liu, H., Peng, H., Song, X., Xu, C., Zhang, M. (2022). Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. *Internet Interventions*, 27. <https://doi.org/10.1016/j.invent.2022.100495>
- Mental Health Foundation. (n.d.). *Stress statistics*.
<https://www.mentalhealth.org.uk/explore-mental-health/statistics/stress-statistics>
- Mian, D. (2024). Exploring the Role of AI-Powered Chatbots for Teens and Young Adults with ASD or Social Anxiety. <https://doi.org/10.48550/arXiv.2412.03740>
- Ng, T., Ngoc, M. N., Anh, D. B. H. (2020). Change of consumer behavior in the post Covid-19 period. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(1), 53-58.
https://www.researchgate.net/publication/353402936_Change_of_consumer_behavior_in_the_post_Covid-19_period
- Okasaka, Y., Morita, N., Nakatani, Y., Fujisawa, K. (2008). Correlation between addictive behaviors and mental health in university students. *Psychiatry and Clinical Neurosciences*, 62(1), 84-92.
<https://doi.org/10.1111/j.1440-1819.2007.01779.x>
- Spytska, L. (2025). The use of artificial intelligence in psychotherapy: development of intelligent therapeutic systems. *BMC Psychology*, 13: 175. <https://doi.org/10.1186/s40359-025-02491-9>
- 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.
<https://doi.org/10.1145/365153.365168>
- Wester, J., Pohl, H., Hosio, S., Van Berkel, N. (2024). "This Chatbot Would Never...": Perceived Moral Agency of Mental Health Chatbots. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW1). <https://doi.org/10.1145/3637410>
- Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M. W., Gill, H., Phan, L., Chen-Li, D., Iacobucci, M., Ho, R., Majeed, A., McIntyre, R. S. (2020). Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Journal of Affective Disorders*, 277, 55-64.
<https://doi.org/10.1016/j.jad.2020.08.001>
- Yang, A., Jarrett, L., Gallagher, F. (2025, August 26). The family of teenager who died by suicide alleges OpenAI's ChatGPT is to blame. *NBC News*.
<https://www.nbcnews.com/tech/tech-news/family-teenager-died-suicide-alleges-openais-chatgpt-blame-rcna226147>

ABOUT THE AUTHOR

Aiden Bui is a high school senior at West Career & Technical Academy in Las Vegas, Nevada. He takes interest in medicine and mental health, hoping to find ways to constantly improve these areas with current technology. He plans to study a STEM-related field in college, with a focus on biology and data analysis.