News and Perspectives

How Does That Large Language Model Make You Feel?

Simon Spichak, JMIR Correspondent

Abstract

People are increasingly turning to commercially available large language models (LLMs) for emotional support. In this *News and Perspectives* article, JMIR Correspondent Simon Spichak reports on the role of LLMs in mental health, speaking with experts about safety concerns, research gaps, and next steps.

Key Takeaways:

- Although there's a lack of data on safety, long-term outcomes, and effectiveness, many people are using large language models (LLMs) for mental health support.
- Some companies are developing clinically validated chatbots for mental health, but these remain untested against general-purpose LLMs.
- Experts urge for more research and understanding from clinicians who are encountering patients that use LLMs for support.

Simon Spichak is the co-founder of Resolvve, a low-cost virtual therapy clinic in Ontario specializing in attention-deficit/hyperactivity disorder (ADHD) and obsessive-compulsive disorder (OCD). He has no commercial interest or association with any products mentioned in this article.

In the mid-1960s, MIT computer scientist Joseph Weizenbaum developed an early chatbot called ELIZA.

It used simple techniques to reflect a user's language, and despite no actual understanding of language, human psychology, or emotions, people interacting with the chatbot were quickly drawn in. Users attributed empathy and understanding to the chatbot, and Weizenbaum's secretary even asked him to leave the room before speaking with the chatbot privately. Later in his life, Weizenbaum warned about the dangers of such chatbots, cautioning that we should never confuse them with humans.

Six decades later, OpenAI launched its flagship large language model (LLM), ChatGPT, prompting other companies to follow suit. Since then, a handful of companies have released several versions of chatbots. Many people are already using LLMs for mental health support. In an October 2025 survey of more than 1800 US adults, 35% reported using AI tools more than once a week for this purpose, including more than 28% who were already seeing a human provider. Researchers are trying to catch up, leaving lots of questions in the air about the effectiveness and safety of these LLMs for mental health.

"We are in the midst of a major mental health crisis that requires systemic solutions, not just technological stopgaps," Arthur Evens Jr, PhD, chief executive officer of the American Psychological Association, said in their new health advisory. "While chatbots seem readily available to offer users support and validation, the ability of these tools to safely guide someone experiencing crisis is limited and unpredictable."

Safety and Effectiveness of General-Purpose LLMs for Therapy

Although LLMs are ubiquitous, it isn't clear how well or for whom they work best, says Shannon Wiltsey Stirman, PhD, a psychologist and professor at Stanford University who serves as the co-director of the Center for Responsible and Effective AI Technology Enhancement for PTSD Treatment (CREATE). A scoping review she coauthored found that most of the studies putting AI models to the test involved few participants and did not use validated measures to assess outcomes.

"There have been a couple of larger and better design[ed] studies, but not nearly enough for us to think they're ready to sort of roll out direct-to-consumer," she says, urging that more testing is needed to know who these models are safe and effective for, and to compare them to the existing standard of treatment.

Though researchers have conducted simulation studies to understand whether chatbots conduct sufficient risk assessment and respond appropriately, these might not reflect real-world interactions.

The guardrails "degrade over time," says Wiltsey Stirman. "If somebody has been interacting with a chatbot for months and months, the way that the chatbot responds to certain disclosures might look very different than it does in a simulation."

General-purpose LLMs "can very often produce direct harm," says Nicholas Jacobson, PhD, associate professor at Dartmouth College. For OCD, where clients have persistent obsessions, he says the nature of LLMs may reinforce detrimental reassurance-seeking behaviors. These models, he said, are also quite poor at helping manage difficulties with

emotional regulation, which are common across mental health disorders.

In a rare few cases, LLMs have been linked with suicide and so-coined "AI psychosis." Though as psychiatrist John Torous, MD, director of the Digital Psychiatry Division at Beth Israel Deaconess Hospital and editor in chief of *JMIR Mental Health,* notes, it doesn't mean that the AI is leading to new-onset delusions. He said there's no definitive evidence yet that interacting with LLMs could lead to new-onset psychosis.

Comparing General-Purpose LLMs to Therapy Chatbots

Jacobson started training proprietary models specific to the mental health context in 2019, developing Therabot as a clinically validated chatbot for therapy. In a randomized controlled trial published in 2025, it was effective in treating depression compared to a waitlist control. Emerging companies like Sonia are building AI chatbots as a "solution for emotional support."

But some experts question whether therapy-specific AI models outperform ChatGPT, Claude, and other LLMs. Torous sees the newer models getting bigger and safer fast, which will make it necessary for companies that offer mental health–specific AIs to test them against publicly available LLMs "to do a better job of justifying themselves."

One of the few published studies pitting these models against each other found no difference between ChatGPT and a specialized LLM built on OpenAI's models. Still, more data is needed for head-to-head comparisons, says Wiltsey Stirman, since the study didn't use "a specially designed foundation model."

Jacobson told JMIR Publications that he and his colleagues have collected internal data showing Therabot is superior but will not test it against other available models due to ethical concerns. The rollout of these models "can be profoundly unsafe and not something that we'll want to facilitate," he says.

Speaking to Patients Using LLMs

Clinicians are likely meeting with clients who are already using LLMs to support their mental health.

Wiltsey Stirman says it's important to keep an open dialogue and understand how clients are using AI. Understanding how these tools work could even help practitioners improve their clients' AI use. "If people are using them for some coaching and accountability and support around a specific goal, the therapist could work with them on how to prompt the chatbot," she says.

> *Wiltsey Stirman says it's important to keep an open dialogue and understand how clients are using AI.*

But Jacobson raised concerns that the advice provided by these chatbots, which don't have specific mental health training, may sometimes be at odds with that of the health care professional they're working with.

The Future of LLMs in Mental Health Care

Other applications are rapidly being developed, including clinician-facing chatbots to aid in diagnosis and assessments, though experts say they're still in the early days of development.

Some researchers are testing whether LLMs could transform clinical notes into handouts for patients, but one study found the process introduces errors. Meanwhile, Utah's drug regulator green-lit a pilot of Legion Health AI, an LLM-powered tool to issue refills for psychiatric prescriptions.

Across the board, experts who spoke with JMIR Publications say it's important to invest in and run high-quality studies examining the impact of LLM use on mental health, but the rapid deployment of direct-to-consumer AI is outpacing the research. These models also carry significant privacy concerns. A 2026 preprint found that 17 of 20 LLMs share raw conversation data with third parties.

A lot of the best data, says Torous, is coming directly from AI companies that publish white papers and preprints instead of going through peer review. He points to a recent blog post from Anthropic as the best data on sycophancy. "Yes, they could be lying, but what else can you do? [You] kind of have to take their word for it," he says.

Keywords: large language models; artificial intelligence; chatbots; mental health; psychotherapy; patient safety; human-computer interaction; simulation; ethics; research design

Please cite as:
Spichak S
How Does That Large Language Model Make You Feel?
J Med Internet Res 2026;28:e105105
URL: https://www.jmir.org/2026/1/e105105
doi: 10.2196/105105