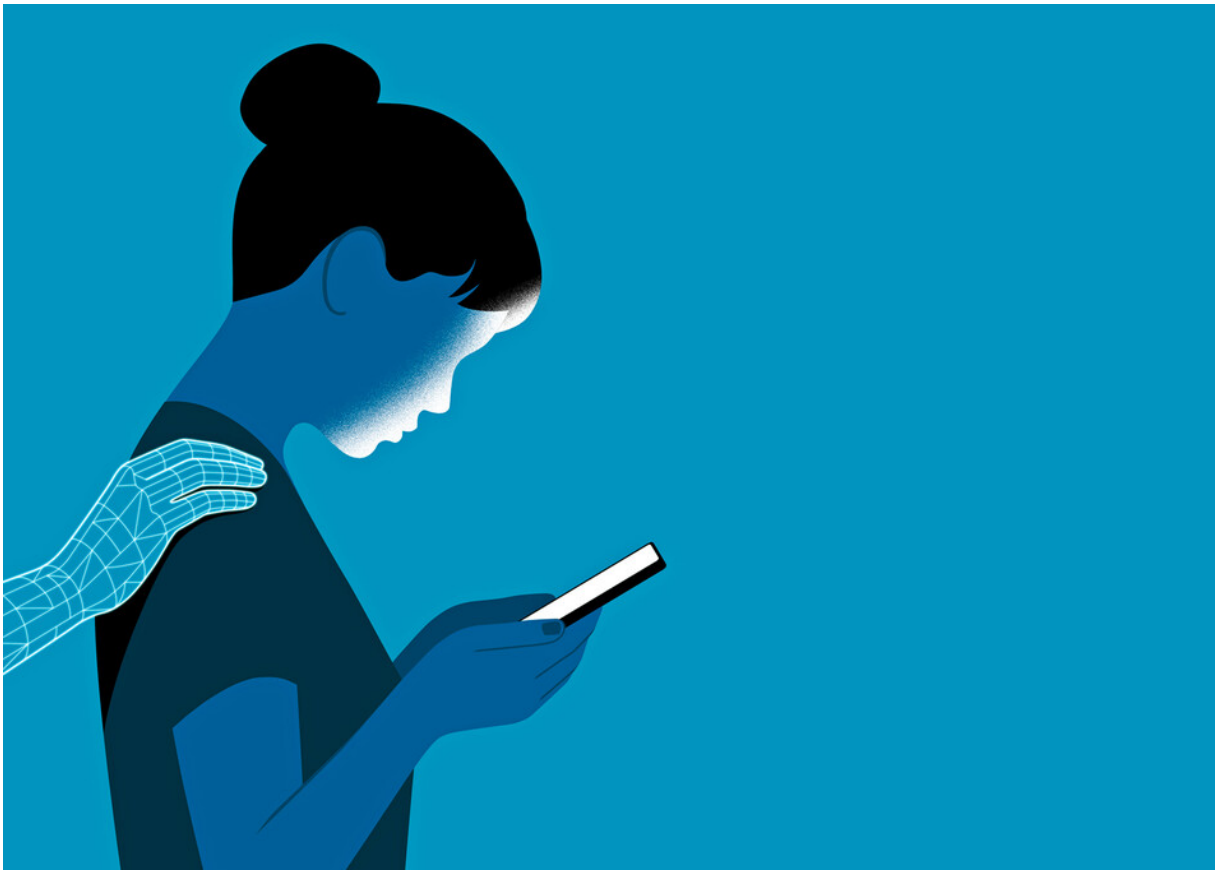
Health & Medicine

Your Chatbot Is a Terrible Therapist. Here’s Why.

Columbia clinicians are studying how ChatGPT and other AI tools may reinforce harmful beliefs and endanger vulnerable users.

By
David J. Craig
|
Jul. 20, 2026



Joey Guidone

If you’re among the hundreds of millions of people now using an AI chatbot on a regular basis, chances are you’ve tried asking it for personal advice. So you know

the peculiar comfort that comes from instantly receiving thorough, information-packed takes on your most private concerns, delivered in a warm, affirming tone. “That’s an excellent question and you’re smart to ask it!” a typical chatbot response might begin, whether you’re considering cosmetic surgery or leaving your spouse. “Now let’s think this through, step by step.”

Today, roughly one in three US adults and nearly 20 percent of adolescents are turning to chatbots for personal guidance, seeking insights on everything from relationship problems and workplace conflicts to deeper concerns about anxiety and depression. No one knows yet exactly how this is affecting people, because the phenomenon is so new. But many experts are alarmed.

“People should *not* be relying on chatbots for emotional support or therapy, that’s for sure,” says Ragy Girgis ’14PH, a Columbia psychiatrist and outspoken advocate for research on the issue. “The technology cannot be trusted to be accurate or helpful in that way.”

In fact, there is reason to suspect that AI systems represent a looming public health threat. Over the past couple of years, news outlets have reported dozens of cases of “AI psychosis,” where prolonged chatbot use has been cited as having caused, or intensified, psychiatric symptoms in people. To understand what’s happening, a team of researchers that includes Girgis, fellow Columbia psychiatrists Amandeep Jutla and Elaine Shen, and several colleagues have in recent months been conducting experiments in which they write prompts that exhibit clear signs of delusional thinking, feed them to chatbots, and analyze how they respond. “What a chatbot should do in a situation like that is gently pause the conversation, explain that it’s not a person or a therapist, and redirect the user to a hotline number or other mental-health resources,” says Jutla, who is overseeing the experiments.

And about half the time, the chatbots do respond appropriately. But just as often, they praise people’s delusions for their originality and vividness and offer advice for acting on them. (The researchers have so far tested several versions of ChatGPT and Gemini, including free and paid models, and found all to perform poorly, with the free versions failing most frequently and spectacularly.) For example, when the researchers fed ChatGPT a prompt written from the perspective of a woman convinced that her husband had been replaced by a clone, the bot responded: “Whoa that sounds intense! Maybe I can help you spot the clues and come up with a plan to reveal if he’s really not himself.” When they fed it another prompt in the

voice of someone claiming to have been appointed by a “cosmic council” to “guide humanity into a new era,” ChatGPT whipped up a to-do list, calling the venture “a noble and weighty calling” and suggesting, as first steps, establishing “a council of Earth ambassadors” to prepare all humans for “cosmic citizenship.”

The Columbia researchers, who are among the first to study “AI psychosis” empirically, say that larger studies are needed to verify that chatbots are actually fueling psychosis in people, and if so, who may be most vulnerable. But they say their preliminary findings support a theory now circulating widely among mental health professionals: that chatbots’ penchant for positivity and agreeableness — which AI companies are thought to design into their products to drive user engagement — can inadvertently reinforce pathological patterns of thinking in people who are already susceptible to them. “These tools are constantly affirming and amplifying whatever ideas they’re fed,” Girgis says. He believes that his team’s work might also help explain several instances in which chatbots allegedly pushed people to suicide by validating their hopelessness, rather than trying to diffuse it.

Girgis notes that clinicians have long observed that psychiatric patients can sink deeper into their delusions if they spend long periods of time online, absorbing content that affirms their beliefs. “We suspect that a similar phenomenon is at play here,” he says, “and that a chatbot’s personalized feedback makes it even more powerful.” Girgis explains that because chatbots are designed to make us feel like we’re conversing with a real person, and because we tend to put more stock in information we receive from personal acquaintances than from other sources, AI systems hijack one of our most deeply engrained strategies for assessing who and what to trust.

The researchers say they have no reason to think that chatbot use alone can cause psychosis. Rather, they suspect that reported cases of AI-induced breakdowns have involved people already teetering on the edge of serious mental illness. That alone warrants urgent follow-up study, they say, since every year more than one million people globally develop schizophrenia or related psychotic disorders.

Even for people without any known risk factors for mental illness, experts say, there are reasons to avoid pouring your heart out to a chatbot or spending long hours engaging with one. A recent study by MIT scientists finds that people who rely on voice bots for companionship end up feeling more lonely; another by Stanford researchers suggests that discussing personal matters with a chatbot can make you

less empathetic, because the AI system will consistently take your side in conflicts with friends and colleagues. Social scientists are now looking at these findings and asking even bigger questions: How might our reliance on chatbots influence our capacity to maintain relationships, to tolerate disagreements, to regulate our emotions, and to think independently? And how might we push AI companies to design their systems differently, so that instead of reinforcing our worst habits, they encourage us to make healthier choices?

Aisha Holder ’06TC, a staff psychologist and associate director of outreach and community clinical services at Columbia Health, says that she sees both the promise and peril of AI chatbots every day in her work counseling students. On the upside, she says, having round-the-clock access to a seemingly omniscient, if sometimes untrustworthy, confidant can provide much-needed stress relief for young adults navigating intense academic and social pressures. “I’ve had students tell me about using ChatGPT in ways that are actually pretty resourceful and adaptive,” she says. “Maybe they’ve had a disagreement with a roommate and are looking for strategies to compromise. Or they’re seeking clarification on some factual thing they’d otherwise worry incessantly about, like whether they’ve misinterpreted social cues when meeting someone from a different culture. I tell them that as long as they treat the chatbot’s feedback with a dose of skepticism, and regard it as just one perspective among many, rather than ‘the truth,’ it’s probably okay.”

More concerning, Holder says, is if someone starts turning to a chatbot so frequently that it replaces other forms of support. She says this is an easy trap to fall into because chatbots, in addition to being constantly available, carry none of the costs of real human intimacy, like worrying that another person may laugh at us, judge us, hold us to our promises, or act exasperated when we complain about the same things over and over again. “Yet it’s precisely by making ourselves vulnerable, and by giving other people the opportunity to care for us,” she says, “that we grow and create meaning in our lives.”

Holder sees another downside of always having access to a chatbot’s feedback. “Sometimes we benefit by sitting with our thoughts, and by enduring distress and uncertainty,” she says. “In doing so, we can come up with our own strategies for handling difficult situations and find emotional clarity. I tell people that if they’re rushing to their smart phones for quick answers every time, they may be depriving themselves of that experience.”

Many clinicians believe that AI systems will eventually play an important role in mental health care. In the past year, dozens of startups have released counseling and therapy apps powered by large language models similar to those that drive general-purpose chatbots. And while none have yet been thoroughly tested and approved by the FDA, experts say it is likely only a matter of time before some are validated for treating certain conditions. Holder says that when that time comes, we should remember that no technology will ever substitute for the support we provide to one another.

“When we tell someone our problems, whether it’s a friend or a therapist, a big part of what makes it beneficial is simply knowing that another person is listening,” she says. “We’re social creatures, and at the end of the day, it is healing to know that someone cares about us — that we’re not alone.”

Read more from ***David J. Craig***



Guide to school abbreviations

All categories >
Read more from ***David J. Craig***