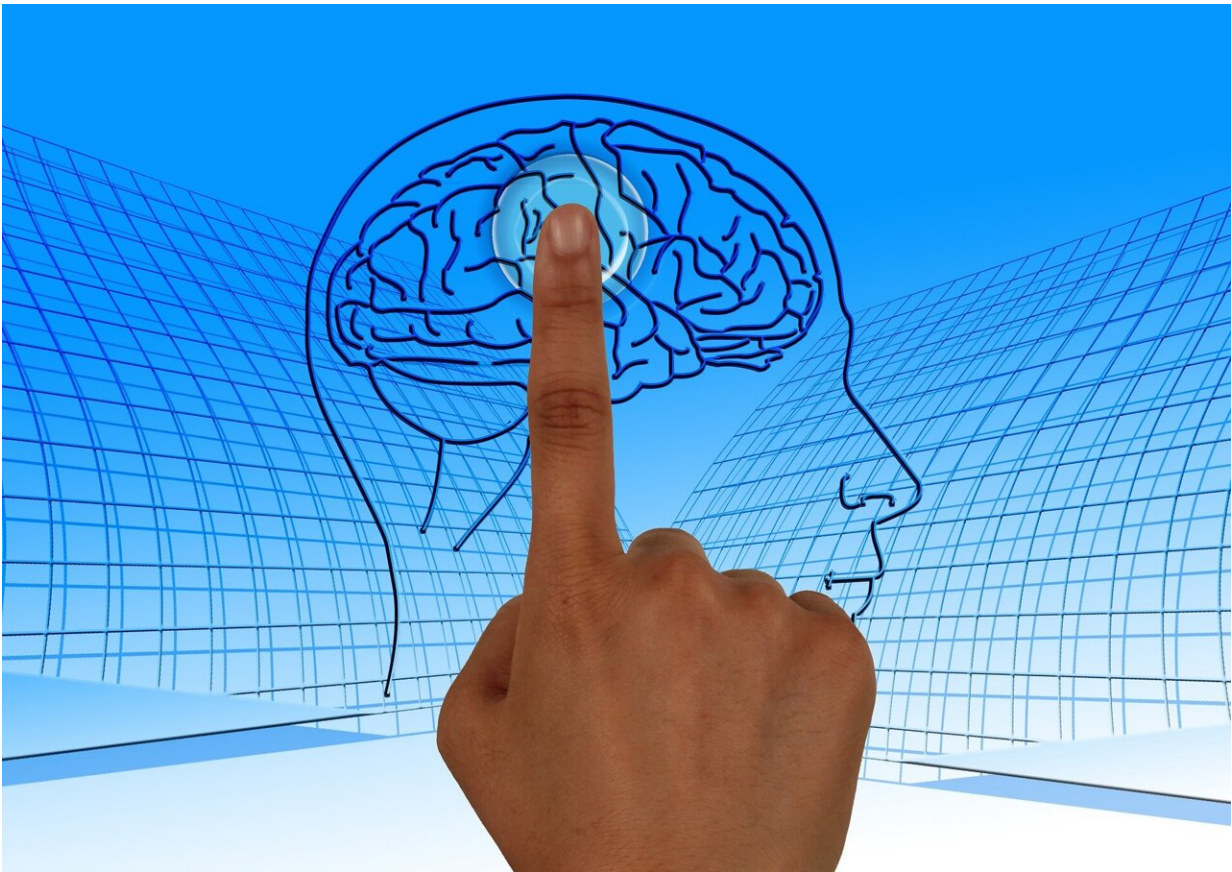


Can AI read our minds? Probably not, but that doesn't mean we shouldn't be worried

April 16 2024, by Sam Baron and Jenny Judge



Credit: Pixabay/CC0 Public Domain

Earlier this year, Neuralink [implanted a chip](#) inside the brain of 29-year-old US man Noland Arbaugh, who is paralyzed from the shoulders down. The chip has enabled Arbaugh to move a mouse pointer on a screen just by imagining it moving.

In May 2023, US researchers also [announced](#) a non-invasive way to "decode" the words someone is thinking from brain scans in combination with generative AI. A similar project sparked [headlines](#) about a "mind-reading AI hat."

Can [neural implants](#) and generative AI really "read minds"? Is the day coming when computers can spit out accurate real-time transcripts of our thoughts for anyone to read?

Such technology might have some benefits—particularly for advertisers looking for new sources of customer targeting data—but it would demolish the last bastion of privacy: the seclusion of our own minds. Before we panic, though, we should stop to ask: is what neural implants and generative AI can do really "reading minds"?

The brain and the mind

As far as we know, [conscious experience](#) arises from the activity of the brain. This means any conscious mental state should have what philosophers and cognitive scientists call a "neural correlate": a particular pattern of nerve cells (neurons) firing in the brain.

So, for each conscious mental state you can be in—whether it's thinking about the Roman Empire, or imagining a cursor moving—there is some corresponding pattern of activity in your brain.

So, clearly, if a device can track our brain states, it should be able to simply read our minds. Right?

Well, for real-time AI-powered mind-reading to be possible, we need to be able to identify precise, one-to-one correspondences between particular conscious [mental states](#) and brain states. And this may not be possible.

Rough matches

To read a mind from brain activity, one must know precisely which brain states correspond to particular mental states. This means, for example, one needs to distinguish the brain states that correspond to seeing a red rose from the ones that correspond to smelling a red rose, or touching a red rose, or imagining a red rose, or thinking that red roses are your mother's favorite.

One must also distinguish all of those brain states from the brain states that correspond to seeing, smelling, touching, imagining or thinking about some other thing, like a ripe lemon. And so on, for everything else you can perceive, imagine or have thoughts about.

To say this is difficult would be an understatement.

Take [face perception](#) as an example. The conscious perception of a face [involves](#) all sorts of neural activity.

But a great deal of this activity seems to relate to processes that come before or after the conscious perception of the face—things like working memory, selective attention, self-monitoring, task planning and reporting.

Winnowing out those neural processes that are solely and specifically

responsible for the conscious perception of a face is a herculean task, and one that current neuroscience is not close to solving.

Even if this task were accomplished, neuroscientists would still only have found the neural correlates of a certain type of conscious experience: namely, the general experience of a face. They wouldn't thereby have found the neural correlates of the experiences of particular faces.

So, even if astonishing advances were to happen in neuroscience, the would-be mind-reader still wouldn't necessarily be able to tell from a brain scan whether you are seeing Barack Obama, your mother, or a face you don't recognize.

That wouldn't be much to write home about, as far as mind-reading is concerned.

But what about AI?

But don't recent headlines involving neural implants and AI show some mental states can be read, like imagining cursors move and engaging in inner speech?

Not necessarily. Take the neural implants first.

Neural implants are typically designed to help a patient perform a particular task: moving a cursor on a screen, for example. To do that, they don't have to be able to identify exactly the neural processes that are correlated with the intention to move the cursor. They just need to get an approximate fix on the neural processes that tend to go along with those intentions, some of which might actually be underpinning other, related mental acts like task-planning, memory and so on.

Thus, although the success of neural implants is certainly impressive—and future implants are likely to collect [more detailed information](#) about [brain activity](#)— it doesn't show that precise one-to-one mappings between particular mental states and particular brain states have been identified. And so, it doesn't make genuine mind-reading any more likely.

Now take the "decoding" of inner speech by a system comprised of a non-invasive brain scan plus generative AI, as reported in [this study](#). This system was designed to "decode" the contents of continuous narratives from brain scans, when participants were either listening to podcasts, reciting stories in their heads, or watching films. The system isn't very accurate—but still, the fact it did better than random chance at predicting these mental contents is seriously impressive.

So, let's imagine the system could predict continuous narratives from [brain scans](#) with total accuracy. Like the neural implant, the system would only be optimized for that task: it wouldn't be effective at tracking any other mental activity.

How much mental activity could this system monitor? That depends: what proportion of our mental lives consists of imagining, perceiving or otherwise thinking about continuous, well-formed narratives that can be expressed in straightforward language?

Not much.

Our mental lives are flickering, lightning-fast, multiple-stream affairs, involving real-time percepts, memories, expectations and imaginings, all at once. It's hard to see how a transcript produced by even the most fine-tuned brain scanner, coupled to the smartest AI, could capture all of that faithfully.

The future of mind reading

In the past few years, AI development has shown a tendency to vault over seemingly insurmountable hurdles. So it's unwise to rule out the possibility of AI-powered mind-reading entirely.

But given the complexity of our mental lives, and how little we know about the brain—neuroscience is still in its infancy, after all—confident predictions about AI-powered mind-reading should be taken with a grain of salt.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Can AI read our minds? Probably not, but that doesn't mean we shouldn't be worried (2024, April 16) retrieved 2 May 2024 from <https://techxplore.com/news/2024-04-ai-minds-doesnt-shouldnt.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.