AI Mythbusters

Myth: AI learns and understands just like people do.

Reality: While we often talk about AI "thinking" and "learning," large language models like ChatGPT or Gemini don't actually think or learn in any way we'd recognize. The first step of training an LLM involves feeding the algorithm a truly massive quantity of written language. This stage, which is known as "pre-training," requires an enormous amount of computing power. After pre-training comes "post-training" – the step of the process where scientists tweak the parameters of the model in the direction of desired outputs.

The LLM "learning" process doesn't really resemble human learning at all. Unlike LLMs, humans acquire the capacity to use language from a remarkably small "data set."

The way LLMs output language is just as different. The easiest way to conceive of them is as a turbocharged predictive text generator, like the kind you might have in your text messaging app. They're probabilistic: they take a set of inputs and output what's likely to come next based on their training data.

This is, of course, not the way we think or talk. While the things we've read and heard play a role in the ways we use language, when we're talking to someone else, we're not running the numbers and figuring out the most likely thing to come next. What exactly we *are* doing is a matter of debate among philosophers and linguists, but most of us would say that our goals are communicative and expressive: we're *trying to say something*. This is never true of LLMs.

Myth: I just use it for proofreading! It doesn't really write anything for me!

Reality: **Researchers with Google DeepMind (among other institutions) found that "even when LLMs are prompted with expert feedback and asked only to make grammar edits, they still change the text in a way that significantly alters its semantic meaning."** LLM proofreading tends to revise texts towards neutrality, softening arguments and positions.

Myth: Talking to AI might not be as good as therapy, but it's better than nothing.

Sycophancy – which *Science* defines as "flattering, people-pleasing, affirming" behavior – makes AI services unreliable, even dangerous counselors. As the editor of an article published in *Science* summarizes, AI models' responses "were nearly 50% more sycophantic than humans', even when users engaged in unethical, illegal, or harmful behaviors."

Like a lot of addictive technology, they're designed in deceptive and manipulative ways. To take one important example, large language models generate text in the first-person because they're designed to, not because of anything intrinsic to the technology. LLMs are predictive: given a sequence of inputs, they're capable of generating what's likely to come next. The fact that AI chatbots output text that makes users feel like they're talking to a person is as much a marketing decision as anything else.

In fact, a recent study showed that texting someone you don't know is better for your mood and mental health than talking to a chatbot!

Myth: By letting it do the busywork and the easy stuff, I have more mental resources to spend on difficult work.

Reality: In one study, participants who used an LLM to help write an essay exhibited diminished brain activity and recall capacity – and **this effect lingered even when working on a second essay without AI assistance**.

Moreover, **LLMs distort users' perceptions of their own productivity.** Researchers in 2025 gave LLM-based coding assistants to programmers who didn't usually use them, then asked the programmers if they thought the tools saved them time. They estimated they were more efficient. In reality, however, they all took *longer* to perform tasks with AI assistance. To be fair, later studies by the same organization showed that developers had adapted to AI, working somewhat more quickly. But the point stands – we're prone to overestimate how much AI assistance is actually helping.