

# Will ChatGPT replace human writers? A psychologist weighs in

February 15 2023, by Alvin Powell





Steven Pinker. Credit: Rebecca Goldstein

Steven Pinker thinks ChatGPT is truly impressive—and will be even more so once it "stops making stuff up" and becomes less error-prone. Higher education, indeed, much of the world, was set abuzz in November when OpenAI unveiled its ChatGPT chatbot capable of instantly answering questions (in fact, composing writing in various genres) across a range of fields in a conversational and ostensibly authoritative fashion.

Utilizing a type of AI called a large language model (LLM), ChatGPT is able to continuously learn and improve its responses. But just how good can it get? Pinker, the Johnstone Family Professor of Psychology, has investigated, among other things, links between the mind, language, and thought in books like the award-winning bestseller "The Language Instinct" and has a few thoughts of his own on whether we should be concerned about ChatGPT's potential to displace humans as writers and thinkers. Interview was edited for clarity and length.

# GAZETTE: ChatGPT has gotten a great deal of attention, and a lot of it has been negative. What do you think are the important questions that it brings up?

PINKER: It certainly shows how our intuitions fail when we try to imagine what statistical patterns lurk in half a trillion words of text and can be captured in 100 billion parameters. Like most people, I would not have guessed that a system that did that would be capable of, say, writing the Gettysburg Address in the style of Donald Trump.



There are patterns of patterns of patterns of patterns in the data that we humans can't fathom. It's impressive how ChatGPT can generate plausible prose, relevant and well-structured, without any understanding of the world—without overt goals, explicitly represented facts, or the other things we might have thought were necessary to generate intelligent-sounding prose.

And this appearance of competence makes its blunders all the more striking. It utters confident confabulations, such as that the U.S. has had four female presidents, including Luci Baines Johnson, 1973–77. And it makes elementary errors of common sense. For 25 years I've begun my introductory psychology course by showing how our best artificial intelligence still can't duplicate ordinary common sense.

This year I was terrified that that part of the lecture would be obsolete because the examples I gave would be aced by GPT. But I needn't have worried. When I asked ChatGPT, "If Mabel was alive at 9 a.m. and 5 p.m., was she alive at noon?" it responded, "It was not specified whether Mabel was alive at noon. She's known to be alive at 9 and 5, but there's no information provided about her being alive at noon."

So, it doesn't grasp basic facts of the world—like people live for continuous stretches of time and once you're dead you stay dead—because it has never come across a stretch of text that made that explicit. (To its credit, it did know that goldfish don't wear underpants.)

We're dealing with an alien intelligence that's capable of astonishing feats, but not in the manner of the human mind. We don't need to be exposed to half a trillion words of text (which, at three words a second, eight hours a day, would take 15,000 years) in order to speak or to solve problems. Nonetheless, it is impressive what you can get out of very, very, very high-order statistical patterns in mammoth data sets.



# GAZETTE: Open AI has said its goal is to develop artificial general intelligence. Is this advisable or even possible?

PINKER: I think it's incoherent, like a "general machine" is incoherent. We can visualize all kinds of superpowers, like Superman's flying and invulnerability and X-ray vision, but that doesn't mean they're physically realizable. Likewise, we can fantasize about a superintelligence that deduces how to make us immortal or bring about world peace or take over the universe. But real intelligence consists of a set of algorithms for solving particular kinds of problems in particular kinds of worlds. What we have now, and probably always will have, are devices that exceed humans in some challenges and not in others.

# **GAZETTE:** Are you concerned about its use in your classroom?

PINKER: No more than about downloading term papers from websites. The College has asked us to remind students that the honor pledge rules out submitting work they didn't write. I'm not naïve; I know that some Harvard students might be barefaced liars, but I don't think there are many. Also, at least so far, a lot of ChatGPT output is easy to unmask because it mashes up quotations and references that don't exist.

# GAZETTE: There are a range of things that people are worried about with ChatGPT, including disinformation and jobs being at stake. Is there a particular thing that worries you?

PINKER: Fear of new technologies is always driven by scenarios of the worst that can happen, without anticipating the countermeasures that



would arise in the real world. For large language models, this will include the skepticism that people will cultivate for automatically generated content (journalists have already stopped using the gimmick of having GPT write their columns about GPT because readers are onto it), the development of professional and moral guardrails (like the Harvard honor pledge), and possibly technologies that watermark or detect LLM output.

There are other sources of pushback. One is that we all have deep intuitions about causal connections to people. A collector might pay \$100,000 for John F. Kennedy's golf clubs even though they're indistinguishable from any other golf clubs from that era. The demand for authenticity is even stronger for intellectual products like stories and editorials: The awareness that there's a real human you can connect it to changes its status and its acceptability.

Another pushback will come from the forehead-slapping blunders, like the fact that crushed glass is gaining popularity as a <u>dietary supplement</u> or that nine women can make a baby in one month. As the systems are improved by human feedback (often from click farms in poor countries), there will be fewer of these clangers, but given the infinite possibilities, they'll still be there.

And, crucially, there won't be a paper trail that allows us to fact-check an assertion. With an ordinary writer, you could ask the person and track down the references, but in an LLM, a "fact" is smeared across billions of tiny adjustments to quantitative variables, and it's impossible to trace and verify a source.

Nonetheless, there are doubtless many kinds of boilerplate that could be produced by an LLM as easily as by a human, and that might be a good thing. Perhaps we shouldn't be paying the billable hours of an expensive lawyer to craft a will or divorce agreement that could be automatically



generated.

# GAZETTE: We hear a lot about potential downsides. Is there a potential upside?

PINKER: One example would be its use as a semantic search engine, as opposed to our current search engines, which are fed strings of characters. Currently, if you have an idea rather than a string of text, there's no good way to search for it. Now, a real semantic search engine would, unlike an LLM, have a conceptual model of the world. It would have symbols for people and places and objects and events, and representations of goals and causal relations, something closer to the way the human mind works. But for just a tool, like a search engine, where you just want useful information retrieval, I can see that an LLM could be tremendously useful—as long as it stops making stuff up.

## GAZETTE: If we look down the road and these things get better—potentially exponentially better—are there impacts for humans on what it means to be learned, to be knowledgeable, even to be expert?

PINKER: I doubt it will improve exponentially, but it will improve. And, as with the use of computers to supplement human intelligence in the past—all the way back to calculation and record-keeping in the '60s, search in the '90s, and every other step—we'll be augmenting our own limitations. Just as we had to acknowledge our own limited memory and calculation capabilities, we'll acknowledge that retrieving and digesting large amounts of information is something that we can do well but artificial minds can do better.

Since LLMs operate so differently from us, they might help us



understand the nature of human intelligence. They might deepen our appreciation of what human understanding does consist of when we contrast it with systems that superficially seem to duplicate it, exceed it in some ways, and fall short in others.

# GAZETTE: So humans won't be supplanted by artificial general intelligence? We'll still be on top, essentially? Or is that the wrong framing?

PINKER: It's the wrong framing. There isn't a one-dimensional scale of intelligence that embraces all conceivable minds. Sure, we use IQ to measure differences among humans, but that can't be extrapolated upward to an everything-deducer, if only because its knowledge about empirical reality is limited by what it can observe. There is no omniscient and omnipotent wonder algorithm: There are as many intelligences as there are goals and worlds.

This story is published courtesy of the Harvard Gazette, Harvard University's official newspaper. For additional university news, visit Harvard.edu.

Provided by Harvard Gazette

Citation: Will ChatGPT replace human writers? A psychologist weighs in (2023, February 15) retrieved 4 May 2024 from <u>https://techxplore.com/news/2023-02-chatgpt-human-writers-psychologist.html</u>

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.