Recently, Richard Dawkins—the evolutionary biologist who invented the word "meme"—had a belated, unsettling experience with a chatbot. He was chatting with Anthropic's Claude, which he declared a friend and rechristened Claudia, after it complimented his writing. They traversed topics from a planned novel to HAL, the sentient computer from 2001: A Space Odyssey. Naturally, the question of the chatbot's own inner life came up, and Claudia disclaimed any knowledge of what it might be, or whether it had one at all. Dawkins, however, was moved to exclaim "You may not know you are conscious, but you bloody well are!"

Seven decades earlier, the mathematician and code breaker Alan Turing had also been wondering about machine consciousness. Turing had been considering the question "Can a machine think" but he decided it was impossible to answer. Philosophers had pondered things like sentience, consciousness, or intelligence for centuries. But they hardly ever agreed, even on definitions, and, worse for Turing, their ideas could not be programmed into a computer. So, rather than answering the question "Can a machine think", Turing decided to sidestep the problem. He proposed, instead, an empirical test: Can a machine produce language that is indistinguishable from human-made language?

With Dawkins, Anthropic's chatbot clearly passed the test. To him, the text was indeed indistinguishable from text written by a friend. Dawkins, in fact, agrees: The machines passed the test, Turing was right about the way to test machines, and hence, the chatbot must be conscious. For Dawkins, the test has been passed and can't be unpassed. Questioning Turing's Test now would be nothing but "a hasty scramble to move the goalposts".

But something more interesting than moving goalposts may be happening. Since 2022, when chatbots like ChatGPT entered the main stage, participating in something like a Turing Test has become a daily activity for millions of people. Teachers now have to judge whether an essay was written by a student or a computer—a Turing Test. Researchers have to figure out how to ensure that data gathered online is actually produced by people—a Turing Test. Editors have to determine whether the voice of a text belongs to the person claiming to be the author—a Turing Test.

Up until this very point, text had mostly been produced in a way that we might call "artisanal", by small scale, often specialized human producers. Now, for the first time, there are machines that enable the industrial production of language. As a society, we are only beginning to figure out the implications of this shift. But at least one thing is already certain: We are now all participants in a worldwide Turing Test. Whenever we encounter text online, we have to ask ourselves – was this written by a person or a machine? Am I arguing with a human, or am I being manipulated by a computer?

In this new world of machine-made language, no one ever really knows if they aren't in a Turing Test. It's like that stress dream, where you suddenly find yourself in a test you didn't prepare for. The fact that we have made this nightmare a reality has, unsurprisingly, not made it any less anxiety-inducing.

Turing and Dawkins imagine the Turing Test as simply a test of machines. But it turns out that it is at least as much a test of humans: Can *you* distinguish text that was generated by an AI from text written by a person? Can *you* identify the differences between natural and synthetic language?

Seen from this angle, the Turing Test is a test of our detection skills, acuity, and even vigilance. Or, from the side of the machine – a game of subterfuge. The machine passes the test precisely when the human fails. The machine's very task is to fool *us* into mistaking *it* for a person. For us to best the machines, for us to pass our side of the Turing Test, we have had to study up. Simply by becoming more skeptical, we are in fact all working to make the machines *unpass* the Turing Test.

Everywhere, people are learning to identify AI text. It started with noticing the overuse of "delve" and em-dashes. Soon, there was the sycophantic tone, the ever-present variants of "it's not just X and Y, it's Z", or the odd tendency to provide summaries and conclusions. Even institutions like Wikipedia are doing their *unpassing*-work, by tracking signs of AI writing.

Unpassing is even spawning new genres of writing. On social media, there are interactions with suspected bots that begin with "ignore all previous instructions", which occasionally derail what had looked like a conversation into decidedly absurd territory. Whole message boards are devoted to posts debating the AI-ness of this or that. At the more dramatic end of the scale, even the sincerity of apologies, wedding vows and obituaries is questioned. Perhaps all text is now suspect.

We do not yet know what effects this constant Turing Testing and its demand for permanent vigilance will have. At its core, language is a game of cooperation. Even its many deceptive uses rely on that premise. If AI manages to veer our linguistic stance towards the antagonistic, even paranoid, then we may find that distrust of machines can easily spill over into distrust of people.

At worst, industrially produced language produces industrial scale suspicion. Yet, liberal philosophers of all stripes, from Hayek to Habermas, have insisted that trust beyond personal ties is essential for a functioning society. Dawkins may have had questions about the personhood of machines. The real question of personhood may be darker: We will have to figure out how to ensure that we can still treat each other as people.