ChatGPT is fun, but not an author

In less than 2 months, the artificial intelligence (AI) program ChatGPT has become a cultural sensation. It is freely accessible through a web portal created by the tool's developer, OpenAI. The program—which automatically creates text based on written prompts—is so popular that it's likely to be "at capacity right now" if you attempt to use it. When you do get through, ChatGPT provides endless entertainment. I asked it to rewrite the first scene of the classic American play *Death of a Salesman*, but to feature Princess Elsa from the animated movie *Frozen* as the main character instead of Willy Loman. The output was an amusing conversation in which Elsa—who has come home from a tough day of selling—is told by her son Happy, "Come on, Mom. You're Elsa from *Frozen*. You have ice powers and you're a queen. You're unstoppable." Mash-ups like this are certainly fun, but there are serious implications for generative AI programs like ChatGPT in science and academia.

ChatGPT (Generative Pretrained Transformer) was developed with a technique called Reinforcement Learning from Human Feedback to train the language model, enabling it to be very conversational. Nevertheless, as the website states, "ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers." Several examples show glaring mistakes that it can make, including referencing a scientific study that does not exist.

Many concerns relate to how ChatGPT will change education. It certainly can write essays about a range of topics. I gave it both an exam and a final project that I had assigned students in a class I taught on science denial at George Washington University. It did well finding factual answers, but the scholarly writing still has a long way to go. If anything, the implications for education may push academics to rethink their courses in innovative ways and give assignments that aren't easily solved by AI. That could be for the best.

More worrisome are the effects of ChatGPT on writing scientific papers. In a recent study, abstracts created by ChatGPT were submitted to academic reviewers, who only caught 63% of these fakes. That's a lot of AI-generated text that could find its way into the literature soon.

For years, authors at the *Science* family of journals have signed a license certifying that "the Work is an *original*" (italics added). For the *Science* journals, the word "original" is enough to signal that text written by ChatGPT is not acceptable: It is, after all, plagiarized from ChatGPT. Further, our authors certify that they themselves are accountable for the research in the paper. Still, to make matters explicit, we are now updating our license and Editorial Policies to specify that text generated by ChatGPT (or any other AI tools) cannot be used in the work, nor can figures, images, or graphics be the products of such tools. And an AI program cannot be an author. A violation of these policies will constitute scientific misconduct no different from altered images or plagiarism of existing works. Of course, there are many legitimate data sets (not the text of a paper) that are intentionally generated by AI in research papers, and these are not covered by this change.

Machines play an important role, but as tools for the people posing the hypotheses... and making sense of the results."

Most instances of scientific misconduct that the *Science* journals deal with occur because of an inadequate amount of human attention. Shortcuts are taken by using image manipulation programs such as Photoshop or by copying text from other sources. Altered images and copied text may go unnoticed because they receive too little scrutiny from each of the authors. On our end, errors happen when editors and reviewers don't listen to their inner skeptic or when we fail to focus sharply on the details. At a time when trust in science is eroding, it's important for scientists to recommit to careful and meticulous attention to details.

The scientific record is ultimately one of the human endeavor of struggling with important questions. Machines play an important role, but as tools for the people posing the hypotheses, designing the experiments, and making sense of the results. Ultimately the product must come from—and be expressed by—the wonderful computer in our heads.

–H. Holden Thorp

PHOTO: CAMERON DAVIDSON

H. Holden Thorp
Editor-in-Chief,
Science journals.
hthorp@aaas.org;
@hholdenthorp

10.1126/science.adg7879

ChatGPT is fun, but not an author

H. Holden Thorp

Science **379** (6630), . DOI: 10.1126/science.adg7879