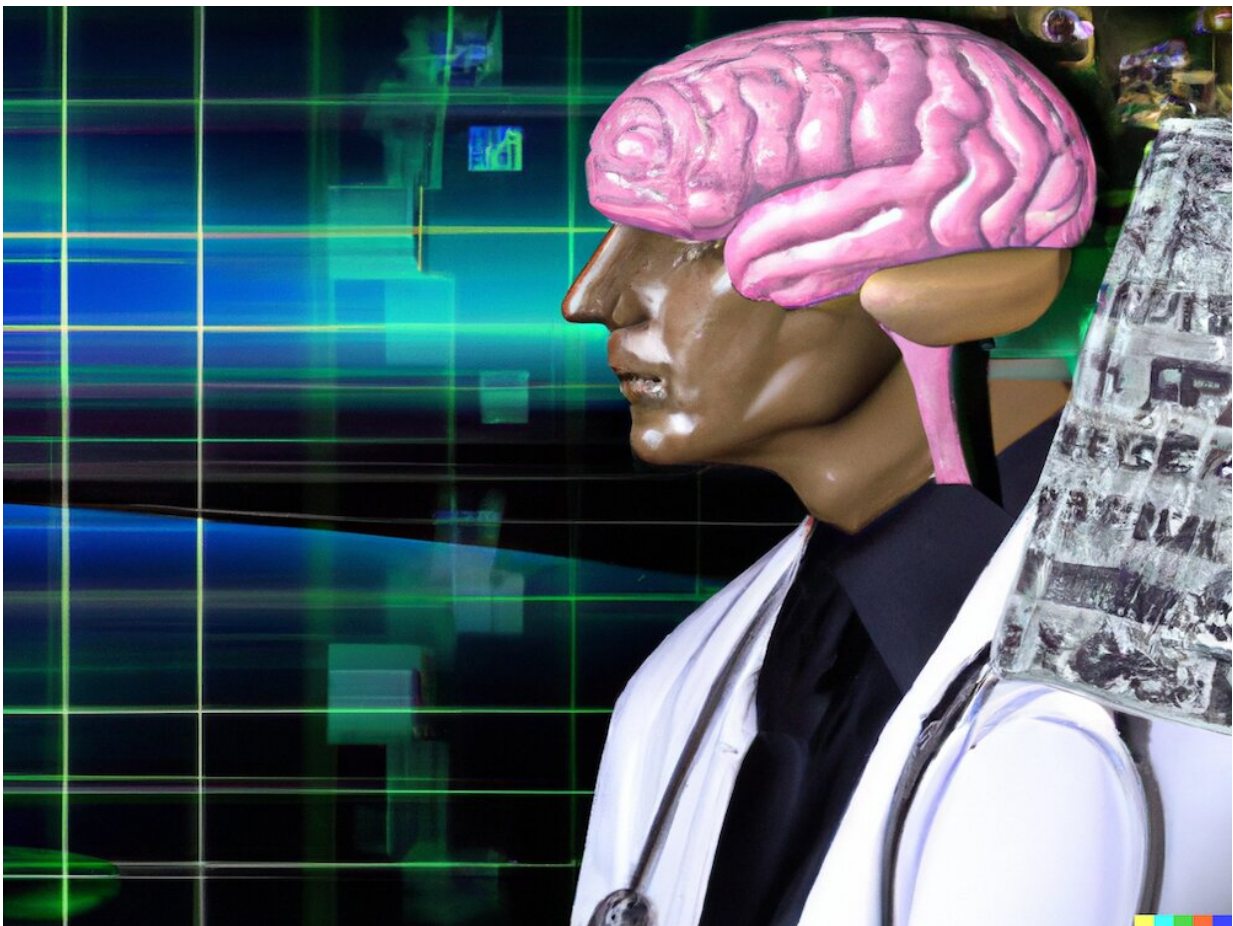


Interview: How does ChatGPT perform on the United States Medical Licensing Examination?

March 30 2023



AI-generated image by DALL-E in response to the request "A futuristic image illustrating the impact of generative artificial intelligence on medical education".
Credit: DALL-E

In a recent interview posted on JMIR TV, JMIR Publications' CEO Dr. Gunther Eysenbach speaks with Dr. Andrew Taylor from Yale University School of Medicine about their paper titled "How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment," published in *JMIR Medical Education*.

The study examined how ChatGPT performed on the United States Medical Licensing Examination (USMLE) compared to other AI language models such as InstructGPT and GPT-3. The researchers found that ChatGPT's performance on the exam was comparable to that of a third-year [medical student](#) in terms of medical knowledge assessment, but more importantly, it outperformed the other two models because its dialogic component enabled it to provide clear rationales for its answers.

ChatGPT's responses were coherent and provided justifiable context. Its accuracy in providing dialogic responses similar to human learners may help create an interactive learning environment for students, supporting problem-solving and reflective practice.

The interview also discusses the limitations of using ChatGPT, such as the need for structured prompts. In their conversation, Dr. Eysenbach remarked how the rapid growth of ChatGPT, which has made AI accessible to end consumers, could be a major disruption and technological shift in the field of [medical education](#). They also cited some concerns with ChatGPT's accuracy in retrieving information such as lack of source identification—a phenomenon called AI hallucination, and the need for additional training or "grounding" of information sources for reliability purposes.

Dr. Eysenbach commented, "There's certainly more work to be done in specifically training ChatGPT on peer-reviewed literature, and perhaps in connecting ChatGPT with more structured databases, which are out

there, like PubMed and CrossRef."

In conclusion, there is interest in exploring how tools like ChatGPT can be used to improve [health care delivery](#), and Dr. Taylor sees potential in using such AI technology in medical education to create a more dynamic learning process for students and practitioners.

"My interest is...how we could potentially use tools like this in the health care system to deliver better and more effective care. And I think we're going to explore potential avenues for that, and I would love to see further development of this in the medical kind of [education](#) space, and I think we will...from a student kind of learning standpoint, it becomes much more dynamic that kind of learning process," added Dr. Taylor.

More information: Aidan Gilson et al, How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment, *JMIR Medical Education* (2023). [DOI: 10.2196/45312](https://doi.org/10.2196/45312)

Provided by JMIR Publications

Citation: Interview: How does ChatGPT perform on the United States Medical Licensing Examination? (2023, March 30) retrieved 6 May 2024 from <https://medicalxpress.com/news/2023-03-chatgpt-states-medical.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.