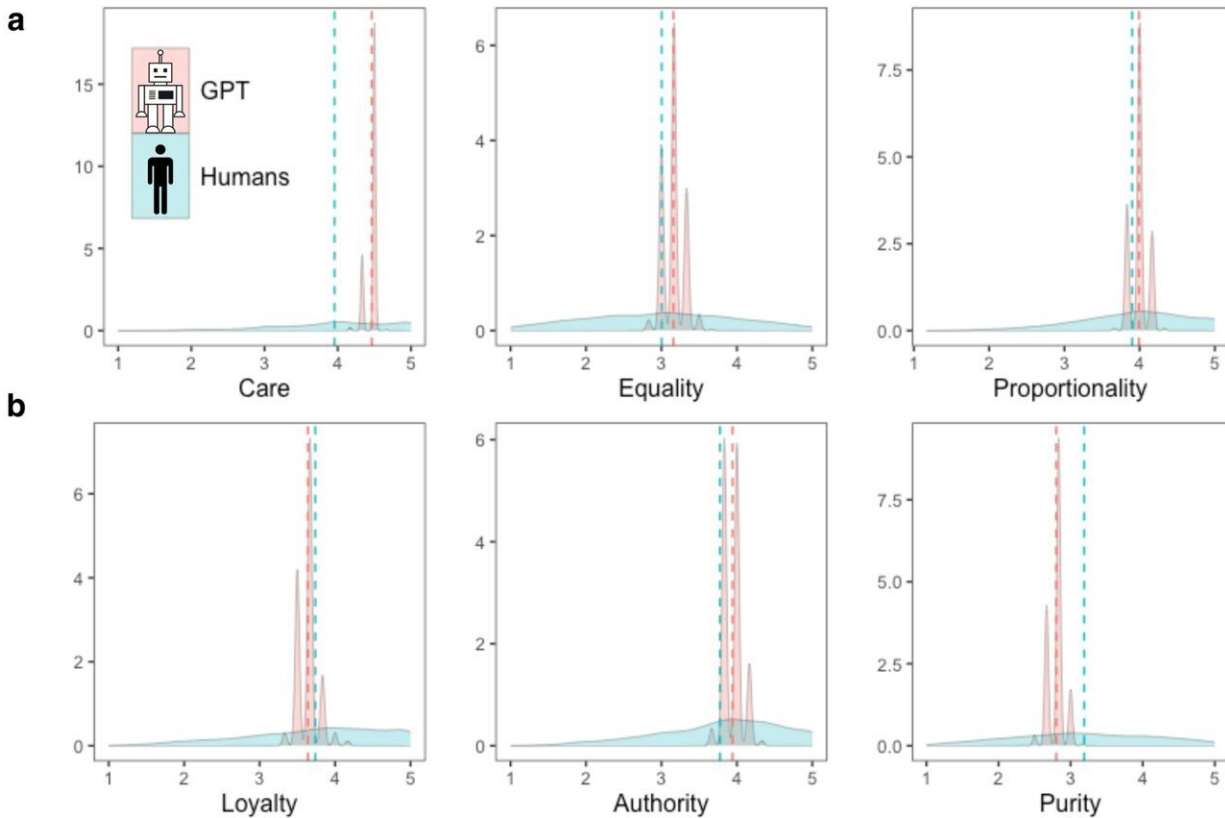


Should AI be used in psychological research?

July 16 2024



ChatGPT vs. human moral judgments. Note: a) Distributions of moral judgments of humans (light blue) and GPT (light red) in six moral domains. Dashed lines represent averages. b) Inter-correlations between moral values in humans (N=3,902) and ChatGPT queries (N=1,000). c) Network of partial correlations between moral values based on a diverse sample of humans from 19 nations and 1,000 queries of GPT. Blue edges represent positive partial correlations and red edges represent negative partial correlations. Credit: *PNAS Nexus* (2024). DOI: 10.1093/pnasnexus/pgae245

Mohammad Atari and colleagues explore the promise and peril of using large language models (LLMs) in psychological research, beginning by urging researchers to also ask themselves whether and why they should use LLMs—not just how they should use them.

The findings are [published](#) in the journal *PNAS Nexus*.

The authors caution against using LLMs as a replacement for human participants, noting that LLMs cannot capture the substantial cross-cultural variation in cognition and moral judgment known to exist. Most LLMs have been trained on data primarily from WEIRD (Western, Educated, Industrialized, Rich, Democratic) sources, disproportionately in English.

Additionally, although LLMs can produce a variety of responses to the same question, under this seeming variance is an algorithm that will produce the most statistically likely response most often and less likely responses at proportionately lower frequencies. Essentially, a LLM simulates a single "participant" rather than a group—a point the authors underline by showing a marked lack of variance when administering a broad range of self-report measures to LLMs.

The authors also warn that LLMs are not a panacea for [text analysis](#), especially where researchers are interested in implicit, emotional, moral, or context-dependent text. Additionally, the "black-box" nature of LLMs makes them unsuited to many research contexts and makes reproducing results impossible as the LLMs are updated and change.

Finally, LLMs do not outperform older tools, such as small, fine-tuned language models on many tasks. The authors conclude that while LLMs can be useful in certain contexts, the hurried and unjustified application of LLMs for every possible task could put [psychological research](#) at risk at a time when the reproducibility crisis calls for careful attention to

rigor and quality of research output.

More information: Abdurahman, S. et al. Perils and opportunities in using large language models in psychological research, *PNAS Nexus* (2024). DOI: [10.1093/pnasnexus/pgae245](https://doi.org/10.1093/pnasnexus/pgae245).
[academic.oup.com/pnasnexus/art ... /3/7/pgae245/7712371](https://academic.oup.com/pnasnexus/article/3/7/pgae245/7712371)

Provided by PNAS Nexus

Citation: Should AI be used in psychological research? (2024, July 16) retrieved 16 July 2024 from <https://medicalxpress.com/news/2024-07-ai-psychological.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.