Prof Susan T. Fiske  
PNAS Psychological and Cognitive Sciences editor:  
06/03/2025

Dear Dr. Fiske—

My colleagues and I would be grateful if you would consider our manuscript, “Affiliation in Human-AI interactions based on shared psychological traits,” for publication as a Research Report inProceedings of the National Academy of Sciences.

People connect with others who share their likes, interests, and mental states, a phenomenon known as homophily. Here, we examine human-AI homophily by testing how individuals' psychological traits are related to their affiliation with large language models (LLM) that mimic their traits.

In Experiment 1, the LLM GPT-4 was trained to align with either an anxious or non-anxious state. Participants (n=100) engaged with both versions of the LLM and then completed a questionnaire to assess how relatable they found the AI. Participants with high anxiety felt more similar to and understood by the LLM instructed to mimic an anxious state, while participants with lower anxiety reported a perceived difference between them and the LLM. In Experiment 2, participants (n=100) engaged with two LLMs aligned with either an extroverted or introverted personality. Extroverted participants felt similar to the AI when it mimicked extroversion. Taken together, the findings support the idea that humans perceive common psychological states in LLMs instructed to mirror aspects of their psychology.

We thought of Proceedings of the National Academy of Sciences (rather than a specialty journal) for this work for four primary reasons:

1. A venue for social cognition. Your journal has published a great deal of high-quality work in social cognition, social perception, and social interaction. The exploration of how psychological traits influence human-AI interaction aligns with your journal's commitment to understanding the complexities of social behavior and relationships.
2. The innovative application of homophily to AI interactions. While homophily has been extensively studied within human-human interactions, the extension of this concept to human-AI interactions offers novel insights. Our findings indicate that AI can be made more friendly and trustworthy by tuning the language it uses to reflect human psychological traits, thus improving the quality of human-AI engagements.
3. Understanding AI's potential for enhancing human psychological well-being. Our studies show that AI's ability to mimic human traits can foster a sense of similarity and understanding, suggesting pathways for AI to support individuals with specific psychological needs or states. This has significant implications for mental health interventions and the design of supportive AI systems.
4. Theoretical implications for the design of empathetic AI. These insights extend the current understanding of empathy and alignment in AI, offering valuable guidance for developers aiming to create AI that users perceive as more relatable and understanding. This bridges the gap between cognitive science, psychology, and AI development.

In closing, we hope that you find this work to be a good match for your journal. This work has not been previously published, nor is it under consideration for publication elsewhere.

We believe the next scientists may be a good fit for the review:

* Cesar Hidalgo ([cesar.hidalgo@tse-fr.eu](mailto:cesar.hidalgo@tse-fr.eu))
* Ziv Ben-Zion ([ziv.ben-zion@yale.edu](mailto:ziv.ben-zion@yale.edu))

Thank you for your consideration!

Sincerely,

Prof. Robin A. Murphy  
*Fellow at Corpus Christi College, Oxford, and Professor of Experimental Psychology  
Department of Experimental Psychology, and Corpus Christi College, University of Oxford  
Corpus Christi College, Merton Street, Oxford, OX1 4JF*