Simulated Contact, Real Empathy: Reducing Prejudice through Role-Play with AI Personas

Chris W. Cai[1[0000−0001−6199−051X]] and Jeremy Ginges[1[0000−0003−2212−5608]]

London School of Economics and Political Science, London, UK
{c.cai1,j.ginges}@lse.ac.uk

1 Introduction

AI role-play chatbots employ the generative capabilities of large language models (LLMs) to simulate characters, provide emotional support, and entertain users through fictional scenarios. Despite being a recent innovation, they have quickly emerged as the most popular recreational use-case for LLMs, with dedicated role-play chatbot providers, like Character.ai, Replika, and Grok's Ami, amassing millions of monthly active users [2].

While academic discourse surrounding AI role-play chatbots has mainly focused on the risks of pathological user attachment, the same mechanisms driving such attachment—social attraction, emotional reciprocity, and perceived value alignment—could make them powerful tools for simulating intergroup contact, thereby reducing prejudice and intergroup conflict.

Here, we report the results from a large-scale, pre-registered randomised controlled survey experiment ($N = 1000$), which tested if an 8-round role-playing interaction with an AI persona of a famous Chinese American historical figure, could reduce prejudice towards Chinese American immigrants. The results indicate that the intervention was highly effective; participants in the treatment-arm saw a 22.8% reduction in prejudice ($d_{Adj} = 0.75$, 95% CI [0.61, 0.88], $p < .001$), compared to 3.1% in the placebo-control group. Our findings suggest that AI role-play chatbots may offer a way to combat prejudice and intolerance at scale.

2 Methods

Following Costello, Pennycook, and Rand [1], we used a two-cell, treatment versus placebo-control experimental design. We recruited 1,000 non-Asian American adults via Cloud Research Connect, demographically matched to U.S. census proportions for age, sex, and ethnicity, with additional quotas to ensure equal representation of Democrats, Republicans, and Independents.

We measured prejudice using a 9-question prejudice measurement questionnaire. We designed the questionnaire to capture prejudice across five constituent dimensions (warmth, acceptance, acknowledgement of discrimination, policy support, and blatant dehumanisation). Participants completed the questionnaire twice. Once before interacting with their assigned chatbot, and once after. Participant responses were normalised to a 0-1 scale using theoretical scale limits, before being averaged and combined into a composite prejudice score.

In the treatment condition, participants were assigned to interact with a custom AI role-play chatbot that embodied the personality, history, and worldview of Professor Chien-Shiung Wu, a real-life Chinese American historical figure, who made many important contributions to physics and was instrumental to the success of the Manhattan Project. In the placebo condition, participants chatted with a point-counterpiont debate chatbot about return-to-office initiatives.

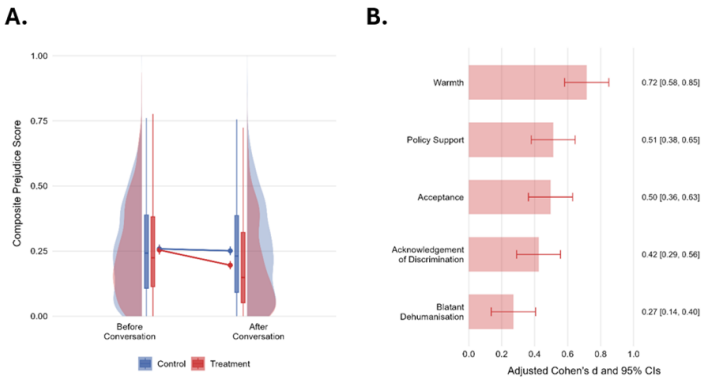

Fig. 1. Main effect (A) and variable level effects (B)

3 Results

Does role-playing with the AI persona of Professor Chien-Shiung Wu reduce anti-Chinese prejudice? Yes, absolutely. The intervention reduced participants' anti-Chinese prejudice by three-quarters of a standard deviation, after adjusting for baseline prejudice scores (b = -0.051, 95% CI $[-0.060, -0.042]$, $p < .001$, $d_{Adj} = -0.75$, partial $\eta^2 = .12$). The effect was robust when controlling for sex, race, education, household income, political party, political ideology, baseline AI attitudes, and baseline prejudice scores. Additionally, it was robust to sensitivity checks for differential attrition (Lee bounds 95% CI [-0.097, -0.005], intention to treat analysis (ITT) with imputation of pre-test scores for missing post-test scores b = -0.045, 95% CI [-0.054, -0.036], d_{Adj} = -0.67).

References

1. Costello, T.H., Pennycook, G., Rand, D.G.: Durably reducing conspiracy beliefs through dialogues with AI. Science **385**(6714), eadq1814 (2024). https://doi.org/10.1126/science.adq1814
2. Aubakirova, M., Atallah, A., Clark, C., Summerville, J., Midha, A.: State of AI: An Empirical 100 Trillion Token Study with OpenRouter. arXiv preprint arXiv:2506.10088 (2026). https://doi.org/10.48550/arXiv.2601.10088