

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

The Illusion of Subjectivity in Conversational AI: Emotional Projection as a interactional Phenomenon

Permalink

<https://escholarship.org/uc/item/11z6q9q5>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Barsand de Leucas, David Leonardo

Caseli, Helena de Medeiros

Iasulaitis, Sylvia

Publication Date

2026

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

The Illusion of Subjectivity in Conversational AI: Emotional Projection as a interactional Phenomenon

Authors: David Leonardo Barsand de Leucas; PPCTS - UFSCar

Helena de Medeiros Caseli; PPGCC - UFSCar

Sylvia Iasulaitis; PPCTS - UFSCar

Abstract

Large language models, particularly in their conversational deployments, invite users to relate to them as if they possessed inner perspectives. This paper conceptualizes "perceived subjectivity" as a socially consequential interactional effect co-constituted in human-AI encounters rather than an intrinsic property of AI or the user's attribution error bias. Through interdisciplinary synthesis, we distinguish perceived subjectivity from phenomenal consciousness, showing how interface features, prompts, histories and user dispositions jointly stabilize impressions of mindedness. This conceptualization shifts the focus from metaphysical debates about machine consciousness and sentience to the interactional dynamics that reorganize trust, emotional dependence, and epistemic authority. By treating perceived subjectivity as a design-mediated effect, we provide pragmatic implications for implementing "productive friction", deliberate safeguard mechanisms that restore critical distance and preserve user autonomy. This approach aims to enable designers and regulators to target interactional conditions to reduce manipulation, deception and epistemic deference risks in affective conversational systems.