The Limits and Potentials of AI in Culturally Mediated Psychological Counseling: Insights From the PASSI@Unito Project

Lei Huang, University of Turin, Italy

The Barcelona Conference on Education 2025
Official Conference Proceedings

Abstract

Artificial intelligence (AI) has become increasingly central to mental health services, offering scalable, accessible, and data-driven forms of psychological support. Yet, as algorithms enter spaces of human vulnerability and care, fundamental questions arise about their cultural adequacy, emotional authenticity, and ethical implications. This working paper outlines the framework for an ongoing clinical ethnographic study developed within the PASSI@Unito project at the University of Turin, a program providing intercultural psychological counselling for international students. The study investigates how AI-based counselling tools, such as chatbots and mood-tracking applications, shape experiences of distress, empathy, and belonging, and how culturally mediated human counselling restores meaning where algorithmic understanding fails. Drawing on preliminary case material and established theoretical traditions in anthropology, psychology, and ethics of care, the project proposes a hybrid model of counselling in which digital systems facilitate access while human mediation sustains interpretation and relational depth. The research aims to contribute both conceptually and practically to the development of culturally responsive, ethically grounded digital mental-health infrastructures.

Keywords: artificial intelligence, intercultural counseling, international students, clinical ethnography

iafor
The International Academic Forum
www.iafor.org

Prefatory Note and Purpose of This Paper

This paper serves as a working document accompanying the BCE 2025 conference presentation. It introduces the conceptual foundation, empirical context, and methodological design of an ongoing research project investigating the role of artificial intelligence (AI) in intercultural psychological counselling for international students. Rather than reporting final results, the paper presents a research framework through which the emerging fieldwork within the University of Turin's PASSI@Unito initiative can be interpreted and extended. It clarifies the problem space, delineates the theoretical and ethical stakes, and synthesizes preliminary insights from early observations in culturally mediated counselling contexts. By situating the study within current debates on digital mental health, empathy, and transcultural psychology, the paper outlines how the project will be conducted and what it aims to achieve.

Introduction

Over the past decade, AI technologies have become deeply embedded in psychological and counselling services. Conversational agents, sentiment-analysis applications, and automated cognitive–behavioural therapy (CBT) platforms have been promoted as low-cost and scalable solutions to address the global mental-health gap (Thakkar et al., 2024). In university contexts, where demand for support continues to outstrip human capacity, such systems have been welcomed for their immediacy, privacy, and accessibility. For international students, who often navigate stressors linked to migration, language, and cultural adaptation, AI-enabled tools appear to offer a form of neutral companionship unbounded by time, stigma, or geography (Fu & Liu, 2024).

However, optimism surrounding AI-mediated counselling conceals significant epistemic and ethical challenges. Many AI systems operate on datasets derived primarily from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) populations, embedding within their design culturally partial assumptions about emotion, selfhood, and mental well-being (Shen et al., 2024). As Kleinman (1988) and Good (1994) remind us, expressions of suffering are not universal but rooted in local moral and linguistic worlds. When such expressions are translated into algorithmic categories such as "anxiety" or "sadness," their cultural logic—and with it, their moral and relational substance—tends to disappear.

Recent empirical research underscores this dissonance. Adolescents in India and Latin American users of AI chatbots report that automated responses often feel "generic" and "culturally irrelevant," suggesting that the algorithm's structure of understanding does not align with their lived idioms of distress (Lanfer et al., 2024; Sehgal et al., 2025). Such misrecognition raises questions not only about representational accuracy but about the redefinition of empathy itself. As Halpern (2024) and Ferrario et al. (2024) argue, machines are increasingly able to mimic the language of care without participating in the affective labour that makes care meaningful. The danger, as Crawford (2021) cautions, is that artificial empathy risks mimicking understanding while erasing the moral labor that gives care its human meaning—producing the illusion of connection without its ethical substance.

These issues are not simply technical limitations but moral and cultural transformations in the meaning of care. They challenge the assumption that mental health can be universally standardized and instead call for methods capable of capturing the subtle interplay between technology, culture, and emotion. The current project responds to this challenge by using clinical ethnography—a methodology situated at the intersection of anthropology and mental

health—to interpret what AI can and cannot apprehend in human experience. The study situates its fieldwork within the PASSI@Unito program at the University of Turin, an initiative combining psychological counselling with intercultural mediation for international students.

The first author's dual role as cultural mediator and counsellor offers a unique position from which to examine these dynamics. This professional proximity enables a reflexive exploration of how digital systems intersect with the embodied, moral, and linguistic dimensions of counselling practice. The project seeks to understand both the potentials and the limits of AI in cross-cultural psychological support and to articulate how technology and human mediation can coexist without diminishing the ethical depth of care.

Theoretical and Conceptual Framework

Empathy, culture, and mediation form the conceptual pillars of this study. Each is being reshaped by the growing influence of algorithmic systems in clinical and educational settings.

Empathy, traditionally considered the moral and affective foundation of counselling, is not merely a cognitive recognition of another's state but a relational process of resonance, vulnerability, and shared temporality (Hollan & Throop, 2008; Ratcliffe, 2022). In digital contexts, empathy is redefined as linguistic pattern recognition. AI-driven counsellors produce textual approximations of compassion without participating in embodied interaction. This distinction between simulation and experience is not trivial: it marks a shift from moral reciprocity to functional mimicry, where emotional responsiveness is treated as a computational output rather than a human commitment.

Culture, likewise, mediates the experience and expression of distress. Anthropological and transcultural psychiatric research has demonstrated that emotional suffering is articulated through idioms, metaphors, and moral vocabularies specific to cultural worlds (Kirmayer, 2020). AI systems, however, tend to universalize affective states through algorithmic models trained on Western linguistic corpora. The result is what Chavarro and Evans (2024) describe as emotional coloniality—a process by which diverse emotional grammars are flattened into standardized categories legible to machines but alien to many users.

Algorithmic mediation thus extends beyond communication to the constitution of care itself. Algorithms define what counts as an appropriate expression of distress and what falls outside the bounds of recognizability (Seaver, 2022). When they fail to respond meaningfully to culturally embedded forms of suffering, responsibility becomes diffuse, shared among designers, clinicians, and institutions (Benjamin, 2019). As Tronto (1993) and Held (2006) argue, care depends on moral attentiveness and accountability—qualities that no system of probabilistic prediction can replace.

In this theoretical light, the project does not seek to oppose technology and humanity but to interrogate their intersection. It asks how AI might function as an entry point to care without eroding the relational and cultural depth that sustains healing.

Research Context and Approach

The study is grounded in the PASSI@Unito program (Psychological Assistance for International Students and Scholars at the University of Turin), founded in 2019 to support the psychological well-being of international students through culturally mediated counselling.

The program's structure pairs professional psychologists with intercultural mediators, enabling a form of dialogue in which linguistic translation becomes an ethical and relational act. This collaboration acknowledges that emotional distress among internationally mobile students is rarely reducible to individual pathology; rather, it often reflects the moral and symbolic dislocations inherent in migration (Huang, 2022; Lagomarsino & Cino, 2023).

Within this institutional framework, the present research adopts clinical ethnography (Good & Good, 2020; Kirmayer, 2020) as its methodological foundation. Clinical ethnography emphasizes participation, reflexivity, and interpretive depth. It involves the reconstruction of lived experiences within their social and cultural contexts, attending to how meaning is performed not only in words but in gestures, silences, and metaphors.

The ongoing fieldwork focuses on international students who have engaged with AI-based counselling tools—such as chatbots, generative dialogue systems, and mood-tracking applications—prior to or alongside human counselling. Through narrative interviews and participant observation, the research examines how these students negotiate emotional expression, cultural belonging, and help-seeking in technologically mediated environments. The analysis attends to both the content of narratives and their affective form: the pacing, hesitation, and tone that reveal how emotion is socially enacted.

The researcher's dual role as intercultural mediator and ethnographer introduces both privilege and risk. It enables an intimate understanding of relational dynamics but demands continuous reflexivity regarding empathy, interpretation, and power. Ethical practice in this study follows the principles of relational accountability and care (Held, 2006; Tronto, 1993), treating informed consent as an ongoing dialogue rather than a single procedural act.

Preliminary Observations and Analytical Orientation

Preliminary case narratives illustrate the ambivalent role of AI in mental health support. Students describe AI tools as simultaneously comforting and alienating. For some, chatbots provided a safe entry point into emotional disclosure, particularly when cultural stigma discouraged open discussion of mental health. Yet these same systems often failed to grasp cultural idioms, producing responses perceived as superficial or misaligned. Users reported feelings of being "heard but not understood," and in some instances, the mechanized consistency of AI dialogue intensified self-doubt and rumination.

By contrast, human counselling sessions, especially those involving cultural mediation, restored a sense of moral and relational recognition. Students found relief not only in being understood linguistically but in having their emotions contextualized within shared cultural logics. Mediators played a critical role in translating metaphors, clarifying cultural references, and reintroducing interpretive ambiguity where algorithms had imposed closure. These early findings confirm that while AI can facilitate access and initial engagement, it cannot substitute the emotional reciprocity that defines therapeutic alliance.

Framework and Future Directions

The research moving forward will deepen and formalize these preliminary observations. It seeks to examine systematically how AI-assisted counselling reshapes emotional self-narration among international students and how intercultural mediation rehumanizes this process. The

analysis will proceed through iterative hermeneutic cycles, connecting narrative interpretation with theoretical reflection on empathy, culture, and technology.

The expected outcomes of the study are threefold. Conceptually, it aims to clarify the distinction between simulated and relational empathy, positioning "algorithmic empathy" as an operational metaphor rather than an ethical reality. Methodologically, it demonstrates the value of clinical ethnography for studying digital care, offering a framework for understanding not only how digital systems function, but how they reshape the human experience of care, empathy, and healing. Practically, it aspires to contribute to the design of hybrid counselling ecosystems, where AI systems facilitate access while human mediators preserve interpretive and moral depth.

Ultimately, the research challenges the assumption that the efficiency of digital tools equates to the adequacy of care. It argues instead that genuine psychological support requires not merely information exchange but recognition—a process rooted in culture, dialogue, and shared vulnerability. The PASSI@Unito project thus provides both a field and a metaphor for this inquiry: a space where machines can assist, but only humans can truly listen.

Conclusion

This working paper outlines the conceptual architecture and methodological design of a study exploring the intersection of AI, empathy, and cultural mediation in mental-health counselling. Positioned within the PASSI@Unito initiative, the project seeks to understand how international students navigate between technological and human forms of support, and what these trajectories reveal about the moral and cultural limits of digital care.

The forthcoming research will continue to elaborate this framework through extended fieldwork, interdisciplinary collaboration, and critical dialogue with practitioners and designers. Its broader objective is to inform both the scholarly discourse on AI and mental health and the institutional practice of counselling in multicultural universities. The paper thus represents a step toward a more reflexive and culturally grounded understanding of how technology can coexist with, rather than replace, the human ethics of care.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that Grammarly, an AI-assisted writing software, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The author further declares that, apart from Grammarly, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

References

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.

Chavarro, D., & Evans, J. (2024). Emotional coloniality and algorithmic universalism in global AI systems. *AI & Society*, *39*(2), 233–248.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence.* Yale University Press.

Ferrario, A., Sedlakova, T., & Trachsel, M. (2024). The illusion of empathy: Ethical reflections on AI in psychotherapy. *Ethics and Information Technology, 26*(1), 12–24.

Fu, J., & Liu, X. (2024). AI-enabled counselling for international students: Potentials and pitfalls. *Journal of Global Higher Education*, *9*(1), 44–59.

Good, B. (1994). *Medicine, rationality, and experience: An anthropological perspective*. Cambridge University Press.

Good, B., & Good, M. J. D. (2020). *Clinical ethnography: Anthropological engagements with illness and suffering*. Cambridge University Press.

Halpern, J. (2024). *The paradox of empathy in automated care.* Yale University Press.

Held, V. (2006). *The ethics of care: Personal, political, and global.* Oxford University Press.

Hollan, D., & Throop, C. (2008). Whatever happened to empathy? Introduction. *Ethos*, *36*(4), 385–401.

Huang, L. (2022). The distance between dreams and reality: Impact of COVID-19 on Chinese students in Turin. *Journal of International Mobility*, *10*(1), 35–54.

Kirmayer, L. J. (2020). Culture and the limits of psychiatric knowledge. *Transcultural Psychiatry*, *57*(6), 829–849.

Kleinman, A. (1988). *The illness narratives: Suffering, healing, and the human condition*. Basic Books.

Lagomarsino, F., & Cino, D. (2023). Culturally mediated counselling in higher education: A relational ethics perspective. *Psychologia dell'Educazione*, *15*(2), 22–40.

Lanfer, H. L., Gratzke, J., & Bürgermeister, E. (2024). Digital clinical empathy in a live chat: Multiple findings from a familial-cancer-focused live chat intervention. *JMIR Mental Health*, *11*, e51598. https://doi.org/10.2196/51598

Ratcliffe, M. (2022). *Real empathy: A philosophical investigation*. Oxford University Press.

Seaver, N. (2022). *Computing taste: Algorithms and the makers of music recommendation*. University of Chicago Press.

Sehgal, R., Kambhamettu, S., Matam, R., Ungar, L., & Guntuku, S. (2025). Cross-cultural evaluation of AI mental health chatbots. *Journal of Affective Computing*, *17*(3), 455–472.

Shen, J., Ben Yehuda, M., & McGinn, M. (2024). Empathy toward artificial intelligence versus human authors: Comparing user responses to AI-written and human-written stories. *JMIR Mental Health*, *11*, e62679. https://doi.org/10.2196/62679

Thakkar, A., Gupta, R., & De Sousa, A. (2024). Artificial intelligence in mental health: Opportunities and ethical dilemmas. *Frontiers in Digital Health*, *6*, 1280235. https://doi.org/10.3389/fdgth.2024.1280235

Tronto, J. C. (1993). *Moral boundaries: A political argument for an ethic of care.* Routledge.

Contact email: lei.huang@unito.it