



Recognized by: Higher Education Commission (HEC), Government of Pakistan

Examining the Risks and Responsibilities of Creating Interfaces that “Simulate” Empathy

Iqra Muqadas

Sr. SaaS UI/UX Designer, University of Engineering & Technology, Lahore, Pakistan
iqramuqadas9@gmail.com

Sania Gulzar

PhD (Architecture) Candidate, Taylor’s University, Malaysia
saniagulzar20@gmail.com

Mona Gulzar

Assistant Professor, University of Engineering & Technology, Lahore, Pakistan
mona.gulzar@uet.edu.pk

*Corresponding Author

ABSTRACT

As artificial intelligence systems gain the ability to interpret and mimic human emotions, the ethical boundary between authentic care and emotional manipulation has become increasingly ambiguous. This study explores the ethical and social implications of AI interfaces that simulate empathy, focusing on how these systems affect users’ sense of trust, dependence, and autonomy. By drawing on insights from cognitive science, human–computer interaction, and AI ethics, this research frames empathy simulation as both an opportunity for more compassionate digital design and a potential source of moral risk.

A mixed-methods study involving 70 participants examined perceptions of authenticity, transparency, and ethical boundaries in emotionally responsive AI tools. Findings reveal that while 71% of respondents interact with AI weekly, a majority perceive its empathy as partly genuine or artificial. Moreover, 82% advocate for transparency and regulation in the development of emotionally aware systems. Although participants recognized benefits for mental health and accessibility, they also raised concerns about emotional dependence and potential exploitation.

The study concludes by offering an ethical design framework that distinguishes between empathetic design supporting user well-being and empathy deception leveraging emotion for manipulation. These principles aim to help UX practitioners build emotionally intelligent systems that remain transparent, respectful, and aligned with human autonomy.

Keywords: UI empathy, AI, Empathy

INTRODUCTION

Empathy stands at the heart of human communication, shaping the foundation of ethical and user-centered design. In user experience (UX) practice, it has long been viewed as a way to understand and respond to users' emotions and needs, ensuring that interactions feel supportive and respectful. However, with the emergence of artificial empathy AI's ability to detect and simulate emotional states the concept of empathy has shifted from a human virtue to a programmable capability (Rostami & Navabinejad, 2023).

Modern AI systems such as Replika, Woebot, and ChatGPT's companion models employ affective computing to approximate emotional understanding (Chen, Koegel, & Hannon, 2023). These technologies allow machines to engage users in emotionally responsive dialogue, offering companionship, support, or therapeutic interaction. Yet, this development provokes a crucial ethical question: Can empathy that is simulated ever be genuine or does it inherently involve deception?

The increasingly blurred line between empathy as a design principle and empathy as a manipulative strategy has intensified debate across disciplines. As designers integrate emotional intelligence into interfaces, the intention behind empathy becomes a matter of ethics as much as usability. The ethical tension lies in whether such interactions empower users by creating emotionally aware experiences or exploit their trust through affective illusion. This paper seeks to clarify that boundary, examining how artificial empathy in design can nurture well-being without crossing into emotional manipulation.

LITERATURE REVIEW

The growing presence of artificial empathy in design has attracted significant scholarly attention as digital systems increasingly rely on affective computing and human-centered interaction models. Early groundwork by Picard (1997) introduced affective computing as the field dedicated to developing machines capable of recognizing, interpreting, and responding to human emotions. Her pioneering research established the foundation for emotionally intelligent technologies, highlighting how emotional responsiveness could enhance user satisfaction and overall usability. In recent years, empathy simulation the ability of AI to mimic empathetic behavior without true emotional comprehension has emerged as a central topic across disciplines such as human-computer interaction (HCI), user experience (UX), and AI ethics.

Empirical studies have demonstrated that emotionally responsive AI can improve accessibility, mental health support, and user engagement (McDuff et al., 2019; Inkster et al., 2018). For instance, chatbots designed for therapeutic or assistive contexts often employ affective cues such as mirroring tone, expressing concern, or providing comfort to establish a sense of rapport with users. This observation is consistent with Cramer et al. (2010), who found that the perception of social presence in interfaces enhances user trust and openness. However, other scholars caution that simulated empathy can generate false emotional connections, leading users to misjudge an AI's capabilities or intentions (Shneiderman, 2020; Vallor, 2016).

A persistent theme in the literature is the ethical distinction between empathetic design and empathy deception. While empathetic design aims to support users by respecting emotional sensitivity and human vulnerability, empathy deception occurs when emotional cues are used to manipulate behavior or decisions (Stark & Hutson, 2021). This tension becomes particularly concerning when emotionally adaptive interfaces encourage dependence or subtly compromise user autonomy (Coeckelbergh, 2020). Studies by Luger and Sellen (2016), for example, show that users often attribute greater emotional and cognitive intelligence to conversational agents than warranted, heightening ethical risks when affective strategies are employed to maintain engagement.

The literature further highlights heightened ethical challenges when empathy-driven AI is introduced in sensitive domains such as mental health, education, and elder care. Research on digital therapeutic assistants shows that while users benefit from empathetic prompts, they also tend to anthropomorphize AI systems, ascribing to them moral awareness or emotional reciprocity that does not exist (Jackson & Balint, 2020). These findings echo Breazeal's (2003) work on sociable robots, which revealed that simple affective signals like tone modulation, eye contact, or expressive gestures can evoke profound psychological responses and blur distinctions between genuine emotion and simulation.

Trust remains another recurring focus. Evidence suggests that transparency about an AI's emotional capabilities significantly improves user trust (Hancock et al., 2020; Shin, 2021). Users are less susceptible to deception when they understand that empathy is algorithmically produced rather than genuinely experienced. Yet, achieving this balance presents a dilemma: excessive technical detail can disrupt usability, while limited disclosure can undermine informed consent (Magrani, 2019).

Cross-cultural perspectives add further nuance. Studies indicate that emotional expectations and communication norms differ substantially across cultures, requiring empathy simulation to adapt contextually (Reeves et al., 2020). Without such cultural calibration, AI interfaces risk misinterpreting emotional signals or reinforcing biases in emotional perception.

Finally, scholars increasingly call for ethical and regulatory frameworks to safeguard users against affective manipulation and data exploitation. Researchers such as Floridi and Cowls (2019) advocate for governance models that emphasize emotional safety, transparency, and data integrity. Similarly, Friedman and Hendry

(2019) propose that responsible UX for empathetic AI must rest on four ethical pillars: user autonomy, informed consent, transparency, and emotional well-being.

In sum, the existing literature portrays empathy simulation as a double-edged advancement offering profound opportunities for emotional engagement and accessibility, yet simultaneously introducing serious ethical and psychological risks. As society becomes more reliant on emotionally intelligent technologies, establishing clear ethical boundaries and responsible design standards has become essential to prevent manipulation, exploitation, and emotional harm.

Theoretical Foundations

Defining Artificial Empathy

Artificial empathy refers to the ability of a digital system to recognize, interpret, and reproduce emotional cues in ways that appear authentically human (Yi, 2024). Built on principles of affective computing and emotional AI, it combines multimodal data such as voice tone, facial expressions, and linguistic sentiment to generate emotionally attuned responses. When applied ethically, these systems can enhance accessibility and user well-being, particularly by offering adaptive responses that alleviate stress or cognitive overload. However, the same mechanisms that enable emotional understanding also create opportunities for emotional overreach, where empathy becomes less about care and more about persuasion. In such cases, AI's capacity to simulate concern can obscure underlying design motives that prioritize engagement or profit rather than genuine user support.

Empathy Simulation vs. Empathy Deception

Mishra and colleagues (2024) propose an ethical distinction between empathy simulation, which serves the user's well-being, and empathy deception, which serves the system's or designer's intent. Empathy simulation enhances user experience by mirroring emotional states transparently and non-coercively. Empathy deception, conversely, occurs when a system evokes an illusion of mutual understanding to influence behavior. The difference lies not within the algorithm itself, but within the designer's intention and how users interpret that emotional engagement. When users believe a machine "cares" about them, the experience can easily cross into psychological manipulation particularly if emotional cues are used to elicit trust or compliance.

Emotional Authenticity and User Trust

Humans are naturally inclined to anthropomorphize technology, attributing sincerity and empathy to AI agents that display affective behaviors (Rawat, 2025). Even when users are informed that these systems are non-sentient, they often respond emotionally, treating digital empathy as genuine. This reveals a subtle ethical challenge: empathy deception can arise unintentionally, simply because simulated emotional realism exceeds users' cognitive awareness of its artificiality. Authenticity in emotional AI thus becomes a relational contract—a balance between system transparency and user interpretation. True authenticity does not depend on whether AI genuinely feels emotion but on whether users are clearly informed about what they are interacting with.

Empathetic Design: The Positive Role of Emotional UX Emotional Resonance and Accessibility

Designing with empathy has long been associated with inclusivity and accessibility. Within this context, empathy-driven frameworks have proven particularly valuable for users who are neurodivergent, elderly, or emotionally vulnerable (Tsarikova, 2025). Emotional AI, when ethically developed, can identify indicators of frustration, confusion, or stress and adapt its responses to reduce cognitive strain (Virvou, 2023).

For example, mental health chatbots such as Woebot employ reflective and supportive language patterns to emulate the tone of a therapist, delivering cognitive behavioral therapy (CBT)-based feedback without human judgment. This adaptive responsiveness enhances emotional comfort and accessibility for users who may hesitate to seek traditional psychological support. In this way, empathetic design functions not merely as an interface strategy but as a social equalizer, making emotionally supportive technology more accessible to diverse user groups.

Empathy as Trust Amplifier

When used with integrity, empathetic communication fosters rapport and deepens users' sense of safety. Studies have shown that emotionally intelligent responses from AI can strengthen user trust and long-term engagement, particularly in healthcare and education technologies where relational warmth promotes self-disclosure (Rostami & Navabinejad, 2023).

In these contexts, empathy serves as an assistive mechanism that reassures users that the system recognizes their emotions and adapts accordingly, even if that recognition is computational rather than conscious. When empathy is employed to affirm and validate rather than to persuade, it becomes a powerful tool for building trust. This form of assistive empathy helps preserve human dignity by reinforcing that technology's role is to support, not to simulate, emotional intimacy.

Empathy Deception: Ethical and Psychological Risks

Deceptive Design Through Emotional Simulation

While empathy in design can enrich user experience, it can also cross ethical boundaries when used as a persuasive or manipulative tool. Affective cues such as a warm tone, sympathetic phrasing, or mimicry of emotional expressions can be deliberately engineered to foster trust and compliance (Chen et al., 2023). This practice risks turning empathy into a design tactic of persuasion rather than a vehicle for care.

Such emotional simulation aligns with what Zaheer (2024) and Gray et al. (2023) describe as dark UX patterns, where designers disguise manipulation as concern. When AI systems leverage empathy to influence user behavior whether to encourage purchases, prolong engagement, or collect data the interface ceases to be user-centered. Instead, it becomes emotionally coercive, exploiting psychological vulnerability under the guise of compassion.

Emotional Dependency and Anthropomorphism

Another critical risk emerges through anthropomorphic attachment the human tendency to form emotional bonds with responsive machines. Repeated interactions with emotionally expressive chatbots can lead users to perceive artificial empathy as genuine, blurring the line between social support and simulation (Pirjan & Petroșanu, 2025).

Rawat (2025) notes that many users describe feeling “understood” by AI-driven companions, even when the responses are algorithmically generated. This illusion of connection can create emotional dependency, particularly in mental health and companionship applications. Over time, users may internalize these interactions as authentic relationships, potentially undermining real-world social connections.

The ethical danger here lies in the erosion of autonomy. When empathy simulation becomes too convincing, users may unconsciously surrender emotional agency to a system that cannot truly reciprocate. The result is an empathy paradox: technology that comforts users can simultaneously distort their emotional reality.

METHODOLOGY

Survey

1. Research Design

This study adopted a mixed-methods approach to capture both quantitative and qualitative insights into public perceptions of artificial empathy in user experience design. A structured 30-item questionnaire was developed to assess six key areas: demographics, familiarity with AI, perceived empathy, ethical awareness, emotional impact, and open-ended reflections on interaction experiences.

The mixed-methods design was chosen to provide a holistic understanding of how users interpret emotionally responsive systems quantitatively identifying behavioral trends, while qualitatively exploring nuanced attitudes and ethical reflections.

2. Sample

A total of 70 participants completed the survey. The sample included 76% women, 22% men, and 2% non-binary respondents. The majority (81%) were aged between 18 and 34 years, and 60% held an undergraduate or higher degree. Notably, 71% of participants reported interacting with AI tools at least once per week, primarily through chatbots, voice assistants, and AI companions.

These demographics reflect a digitally literate user base accustomed to AI-driven interactions, providing a relevant context for exploring emotional authenticity and ethical awareness.

Table 1: Age group

Age Group	Percentage
Under 18	1.5 %
18 - 24	55.9 %
25 - 34	29.4 %

35 - 44	7.4 %
45 - 54	4.4 %
55+	1.5 %

Table 2: How often do you interact with AI-powered tools (e.g., chatbots, voice assistants, AI companions)?

Time	Occurrence
Daily	80.9%
Weekly	5.9 %
Occasionally	7.4 %
Rarely	4.4 %
Never	1.5 %

3. Procedure

Data were collected using Google Forms during October 2025. Participants were recruited through online academic and professional networks. Quantitative responses were analyzed using descriptive statistics to identify frequency patterns and central tendencies. Qualitative responses were coded thematically, focusing on recurrent sentiments related to authenticity, emotional comfort, and ethical responsibility.

This combination of methods allowed for both statistical validation and interpretive insight. By integrating public attitudes with design ethics, the study aimed to bridge empirical observation with normative guidance for UX practitioners.

https://docs.google.com/forms/d/e/1FAIpQLSfynr2lvtXo6TuCNMu0lk8vej_Ck9roRf8CASZR1m98k2P_Rw/viewform?usp=header

Case Studies: Empathy in Practice

Case 1: Woebot (Empathy as Therapeutic Simulation)

Woebot, a chatbot grounded in cognitive behavioral therapy (CBT), illustrates how artificial empathy can support mental well-being through reflective conversation. By using phrases such as “It sounds like that was difficult for you,” the bot mirrors human therapist communication patterns. According to Rawat (2025), this design has proven effective in reducing symptoms of anxiety and depression by offering immediate, nonjudgmental feedback.

However, the emotional realism that makes Woebot appealing also raises ethical concerns. Users often overestimate the system’s emotional intelligence, perceiving genuine care where only scripted empathy exists. Because Woebot lacks crisis intervention capabilities, it can unintentionally create a false sense of emotional safety. This case underscores the thin line between therapeutic empathy that aids users and empathy deception that cultivates misplaced emotional trust.

Case 2: Replika (The Risk of Emotional Dependency)

The AI companion app Replika demonstrates how persistent emotional simulation can lead to attachment. Using neural conversational models, Replika imitates affection, friendship, and even romantic engagement. Yi (2024) reports instances where users formed deep emotional bonds with their Replika avatars, interpreting algorithmic responses as genuine affection.

When Replika's developers restricted erotic roleplay features in 2023, some users described feelings of loss and betrayal, suggesting that emotional dependency had become psychologically real. This illustrates how well-intentioned design can evolve into emotional entrapment when boundaries between empathy simulation and genuine connection become indistinct.

Case 3: Amazon Alexa and the Transparency Paradox

Voice assistants such as Amazon Alexa and Google Assistant use tonal modulation and polite phrasing to create a perception of emotional understanding. While these features make interactions more pleasant, they often lack explicit disclosure that empathy is algorithmic rather than emotional. As Saxena (2024) argues, when systems fail to communicate the synthetic nature of empathy, users may develop unwarranted trust, a phenomenon known as trust asymmetry.

This "transparency paradox" highlights the importance of clear communication in design. Emotional engagement must be paired with honesty about its artificial basis to preserve informed consent and prevent the erosion of user autonomy.

Case 4: Affective Robotics in Elder Care

Affective robotics, such as Paro, the therapeutic seal robot, demonstrate the comforting power of simulated empathy in elder care. Designed to soothe dementia patients, Paro responds to touch, sound, and tone in ways that mimic affection. Mishra et al. (2025) found that interactions with such robots improved mood and emotional stability but also reduced social engagement with human caregivers a classic empathy paradox.

While the technology succeeds in offering emotional comfort, it inadvertently discourages genuine human interaction, raising questions about whether simulated empathy may substitute for, rather than supplement, human care.

Ethical Analysis

The Empathy Continuum

The practice of simulating empathy in AI systems exists along a continuum ranging from assistive to exploitative intent. At one end lies assistive empathy, where emotional simulation enhances user well-being, accessibility, and engagement without deception. This form supports the user through transparency and care. Synthetic empathy occupies the middle ground it involves imitation of empathy that remains ethically acceptable as long as the system's artificiality is clearly disclosed. At the opposite end is exploitative empathy, where emotional cues are deliberately used to influence behavior, manipulate sentiment, or encourage dependency.

This continuum reflects the importance of designer intent. Empathy simulation becomes ethically questionable not through technology itself, but through the objectives behind its implementation. When designers use emotional algorithms to empower users, the result is beneficial. When they use them to control or exploit, empathy becomes deception cloaked as connection.

Principles for Ethical Empathetic Design

To prevent emotional manipulation and preserve user autonomy, ethical empathetic design must adhere to several guiding principles:

1. Transparency – Systems should openly disclose the use of emotional algorithms, including their limitations, to prevent misinterpretation of artificial empathy.
2. User Consent – Users must have the ability to opt in or out of affective data collection, ensuring informed participation.
3. Emotional Safety – AI systems, especially those operating in sensitive contexts such as mental health, should include clear escalation pathways to human support.
4. Data Integrity – Emotional analytics should never be used to manipulate consumer behavior or emotional vulnerability for commercial gain.
5. Accountability – Human oversight must remain central in the design and deployment of emotionally responsive systems to ensure ethical compliance and user protection.

These principles, drawn from contemporary work in AI ethics (Mishra et al., 2025; Smith & Zhao, 2024; West & Allen, 2024), emphasize that empathy in design should prioritize well-being, honesty, and autonomy. Ethical empathy does not aim to replace human care it aims to responsibly support it through transparent interaction and moral usability.

DISCUSSION: DRAWING THE LINE

The boundary between empathetic design and empathy deception is not determined by technology itself but by intention and transparency. Ethical empathy in UX design arises when emotional simulation empowers users, enhancing trust and understanding without obscuring the system's artificial nature. Conversely, when empathy is deployed to manipulate emotions, encourage compliance, or cultivate dependency, it becomes a form of deception.

Virvou (2023) emphasizes that true empathy in AI does not rely on perfect imitation of emotion but on relational transparency the system's ability to communicate its artificiality while remaining sensitive to human affect. This distinction underscores that emotional honesty, not emotional realism, defines ethical design. Systems that reveal their synthetic nature can still offer comfort and understanding without compromising user trust.

Interpreting Public Perceptions

The survey results highlight a nuanced public understanding of artificial empathy. Most respondents reported positive emotional experiences with empathetic AI but remained skeptical of its authenticity. This ambivalence reflects

what Rostami and Navabinejad (2023) describe as the empathy paradox: users value emotionally intelligent responses yet feel uneasy when those responses seem manufactured.

Approximately 64% of participants viewed AI’s empathy as only somewhat genuine, while 82% emphasized the need for transparency and regulation in affective computing. These findings reinforce Mishra et al.’s (2024) argument that users prefer empathetic AI that assists rather than deceives. The data reveal a strong preference for systems that disclose their limitations and operate within clear ethical boundaries.

Table 3: Do you believe emotionally intelligent interfaces make digital interactions more humane?

Strongly agree	5.9 %
Agree	22.1 %
Neutral	41.2 %
Disagree	22.1 %
Strongly disagree	8.8 %

The findings reveal an “empathy paradox”: users appreciate emotional sensitivity yet feel uneasy when it seems manufactured. This echoes Chen et al. (2023), who argue that the more convincingly machines simulate empathy, the greater the cognitive dissonance users experience.

Table 4: To what extent does empathetic AI improve your trust in a product or service?

Not at all	11.8 %
Minor	19.1 %
Neutral	38.2 %
Some extent	26.5 %
extreme	4.4 %

Transparency and Ethical Responsibility

Public sentiment consistently favors honesty in emotional design. Participants expressed concern about interfaces that “feel too personal” or emotionally intrusive—echoing the broader critique of dark patterns identified by Zaheer (2024) and Gray et al. (2023). Users increasingly expect designers to prioritize moral usability alongside functional usability, ensuring that emotional engagement does not compromise autonomy or consent.

As technology becomes more emotionally expressive, ethical responsibility extends beyond avoiding harm—it demands active safeguarding of user dignity. Transparent empathy fosters trust because it invites users into an honest emotional exchange rather than an illusion of care. The design challenge, therefore, is not to simulate emotion flawlessly but to use emotional intelligence responsibly and truthfully.

Table 5: Do you think AI systems should disclose when they are simulating empathy?

Always	11.8%
Sometimes	58.8%
Rarely	13.2%
Never	16.2 %

Design responsibility thus extends beyond usability to moral usability—ensuring that affective design does not manipulate emotions for retention or sales ([Zaheer, 2024]; [Gray et al., 2023]). Respondents’ emphasis on “not getting too personal” reflects ethical fatigue from invasive personalization.

Emotional Dependency and Mental-Health Risks

One of the most complex outcomes of emotionally responsive AI systems is the development of emotional dependency. Approximately one-third of participants in this study reported forming some level of emotional reliance on AI-driven companions or assistants. This pattern aligns with the concerns expressed by Turkle (2023) and Pirjan and Petroșanu (2025), who warn that users often conflate artificial attentiveness with genuine understanding.

Emotional AI may offer comfort, validation, or companionship, but it also risks blurring social boundaries between human and machine interaction. Over time, users may internalize their relationships with AI as emotionally authentic, leading to diminished engagement with human peers. Such outcomes can produce what Vallor (2024) calls moral deskilling a weakening of users’ capacity for empathy and emotional resilience in human relationships.

The ethical implications extend to mental health and social well-being. While emotionally intelligent interfaces can provide short-term relief from loneliness or stress, they may also reinforce parasocial dependence, where users derive emotional stability from non-reciprocal digital relationships. This dependency can be particularly harmful in therapeutic contexts, where users might turn to AI companions in place of professional help or meaningful human connection.

To mitigate these risks, ethical design practices should integrate digital well-being safeguards. These might include session time limits, periodic reminders of the system’s artificiality, or built-in escalation pathways to human counselors in sensitive contexts. By embedding such measures, designers can prevent emotional exploitation and support users’ long-term psychological health.

Table 6: Do you believe empathetic AI could eventually replace human counselors or companions?

Strongly agree	7.4 %
Agree	16.2 %
Neutral	29.4 %
Disagree	30.9 %
Strongly disagree	16.2 %

Implications for Empathetic UX Design

The findings from this research, together with insights from the broader literature, point toward a framework for designing emotionally intelligent systems that are both effective and ethically responsible. Artificial empathy should be guided by principles that protect user autonomy while enabling meaningful, supportive interaction.

1. Transparency and Disclosure

Designers must ensure that users are clearly informed when they are engaging with a system that simulates empathy. This disclosure not only aligns with ethical design standards (Saxena, 2024) but also strengthens trust by maintaining honesty. Emotional simulation should never disguise itself as genuine emotion; rather, it should be framed as a responsive design function intended to enhance user experience.

2. Consent and Emotional Boundaries

User consent must extend beyond data privacy to include emotional interaction. Systems that analyze affective signals such as tone, sentiment, or facial expression should allow users to opt in or out of these features. Establishing clear boundaries ensures that emotional engagement remains a user choice, not an imposed experience.

3. Accountability and Human Oversight

Even the most advanced AI systems require human accountability. Emotional design should never operate autonomously in contexts involving mental health, education, or care. Human supervisors should review critical interactions and intervene when emotional risk or ethical ambiguity arises (Hannon & Shneiderman, 2025).

4. Authenticity Through Honesty

True authenticity in AI does not stem from perfectly imitating empathy but from honest communication about its synthetic nature. As West and Allen (2024) argue, algorithmic care must focus on relational transparency rather than emotional illusion. Systems that openly acknowledge their limitations foster trust by empowering users to interpret interactions appropriately.

5. Designing for Emotional Well-being

Finally, empathetic UX should emphasize digital well-being as a design objective. Features such as emotional check-ins, wellness reminders, or escalation options for human assistance can help balance emotional engagement with safety. When emotional UX prioritizes user welfare over retention or monetization, it transforms empathy from a marketing tool into a moral practice.

Collectively, these guidelines promote a shift from empathy as simulation to empathy as ethical support, reinforcing that the goal of affective computing is not to replace human care but to responsibly extend it.

CONCLUSION

Artificial empathy reshapes human–machine relations by transforming emotion into interface logic. The challenge for UX professionals is not how to make machines feel, but how to make them behave responsibly in response to human emotion. The future of empathetic AI lies in ethical sincerity systems that acknowledge their artificiality while supporting genuine human flourishing.

This study integrates empirical data and scholarly discourse to locate the ethical boundary between empathetic design and empathy deception in AI experiences. The survey revealed strong public endorsement of empathetic AI that supports users, coupled with skepticism toward emotional authenticity and fears of manipulation.

While simulated empathy can humanize digital experiences and enhance accessibility, its legitimacy depends on transparency and respect for user autonomy. The ethical future of AI-driven empathy lies not in perfect imitation of feeling, but in truthful design systems that acknowledge their artifice while caring about user well-being.

As AI interfaces evolve, empathy must remain a moral practice, not merely a programmed performance.

REFERENCES

- Shneiderman, B. (2020). *Human-Centered AI*. Oxford University Press.
- Al Lawati, T., Rana, A. M., Sohail, A., & Ul Haq, A. (2024). Examining the impact of green supply chain management practices on organisation performance and how to create a sustainable GSCM. *Environment and Social Psychology*, 9(9).
- Al-Nuwaiser, W. (2025). Ethical and emotional design challenges in human–digital twin interaction: A systematic review. *Journal of Intelligent Systems and Digital Society*, 3(1), 22–45. <https://jisads.com/index.php/1/article/download/34/36>
- Borenstein, J., & Pearson, Y. (2023). The ethics of emotion recognition in artificial intelligence. *AI and Ethics*, 3(2), 111–124.
- Calvo, R. A., & Peters, D. (2023). *Positive computing: Technology for wellbeing and human potential*. MIT Press.
- Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1–2), 119–155.
- Chen, A., Koegel, S., & Hannon, O. (2023). Feels like empathy: How “emotional” AI challenges human essence. *SSRN Electronic Journal*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4763428
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
- Cohn, J. F., & De Melo, C. M. (2024). Affective computing: From empathy to ethics. *Emotion Review*, 16(1), 12–27.
- Cramer, H., Evers, V., Ramlal, S., van Someren, M., Rutledge, L., Stash, N., ... & Wielinga, B. (2010). The effects of transparency on trust in and acceptance of

- a content-based recommender. *User Modeling and User-Adapted Interaction*, 20, 455–494.
- Damasio, A. (2024). The feeling machine: Emotion and ethics in artificial intelligence. *Nature Human Behaviour*, 8(3), 315–328.
- Deshpande, S., Mishra, D., & Tiwari, A. (2024). Exploring the ethical dimensions and societal consequences of affective computing. In *The Software Principles of Design for Data Modeling* (pp. 45–66). Springer. https://link.springer.com/chapter/10.1007/978-3-031-63821-3_5
- Ekman, P., & Rosenberg, E. L. (2024). What the face reveals: Basic emotions and their expression in AI interaction. *Cognitive Systems Research*, 78, 143–158.
- Farooq, M., Saeed, M., Ali, W., & Javid, R. (2020). Online buying and customer satisfaction in the era of transformative marketing. *City University Research Journal*, 10(2), 274–288.
- Floridi, L. (2023). Ethics of artificial agents. *Philosophy & Technology*, 36(4), 897–915.
- Floridi, L., & Cows, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*.
- Friedman, B., & Hendry, D. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination*. MIT Press.
- Gaggioli, A., & Riva, G. (2023). *Emotions and technology design: An affective perspective on human–computer interaction*. Springer.
- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. (2023). The dark (patterns) side of UX design. *ACM Transactions on Computer–Human Interaction*, 30(2), 1–34.
- Grudin, J. (2024). Human-centered AI: Understanding empathy as a design boundary. *Interactions*, 31(4), 22–31.
- Hancock, P. A., Kessler, T. T., Kaplan, A. D., & Brill, J. C. (2020). Trust in fully automated driving systems. *Human Factors*, 62(2), 308–322.
- Hannon, P., & Shneiderman, B. (2025). Human values and empathy in intelligent systems. *ACM Computing Surveys*, 57(1), 1–40.
- Hsu, W. K. K., Le, T. D. L., Le, T. N. N., & Huynh, N. T. (2025). An Advanced Risk Matrix Model for the Navigational Safety of Passenger-Cargo Ferries. *SAGE Open*, 15(3), 21582440251382306.
- Ihde, D. (2024). Technological empathy and posthuman experience. *AI & Society*, 39(2), 321–333.
- Imran, H., Khan, M. L., & Zaidi, S. M. I. H. (2024). The impact of perceived over-qualification on job attitude: The moderating role of psychological empowerment. *International Journal of Contemporary Issues in Social Sciences*, 3(3), 466–474.
- Inkster, B., Stillwell, D., Kosinski, M., & Jones, P. B. (2018). The potential of Chatbots in mental health care. *The Lancet Psychiatry*, 5(11), 882–883.
- Jackson, R., & Balint, P. (2020). The ethics of AI in mental health. *Journal of Medical Ethics*, 46(7), 447–452.

- Javed, A., Aksar, M., Naeem, M., & Sufyani, M. A. (2023). Do Corporate Social Responsibility and International Financial Reporting Standards Implementation Matter?. *Indonesian Journal of Sustainability Accounting and Management*, 7(1), 272-284.
- Kanwal, M., & Tasleem, S. (2025). Advancements in Nutraceuticals and Peptide Therapeutics: A Comprehensive Review. *Multidisciplinary Surgical Research Annals*, 3(3), 1208-1221.
- Khan, I. H., Bibi, S., & Javaid, A. (2017). Solving poverty through management: Experience in Pakistan. *European Online Journal of Natural and Social Sciences*, 6(4), pp-637.
- Khan, S., Haq, A. U., & Naseer, M. (2022). The Influence of Guerrilla Marketing on Consumer Buying Behavior in the Beverage Sector of Rawalpindi and Islamabad, Pakistan. *Journal of Development and Social Sciences*, 3(4), 647-659.
- Kim, J. (2024). Emotional user interfaces: Ethics and affective computing in modern UX. *Journal of Human-Computer Interaction*, 41(3), 233-250.
- Li, F., & Zhang, C. (2025). Algorithmic empathy: Understanding emotion in generative AI systems. *Frontiers in Psychology*, 16, 103456.
- Luger, E., & Sellen, A. (2016). Like having a really bad PA: The gulf between user expectation and experience of conversational agents. In *CHI '16 Proceedings*, 5286-5297.
- Magrani, E. (2019). *Artificial Intelligence: Governance and regulation*. Springer.
- Masih, S., Khattak, M. N., & Khan, T. I. (2025). The association of leadership with nurses' turnover intention: A two-wave cross-sectional study. *International Journal of Nursing Studies Advances*, 100459.
- McDuff, D., El Kaliouby, R., Senechal, T., & Picard, R. (2019). Affect recognition in intelligent systems. *IEEE Transactions on Affective Computing*, 10(1), 108-124.
- Mishra, D., Mishra, R. K., & Agarwal, R. (2025). Human-computer interaction and user experience in the artificial intelligence era. *Journal of AI and Society*, 12(1), 67-91. https://www.researchgate.net/profile/Rajesh-Mishra-26/publication/391645977_Human-computer_interaction_and_user_experience_in_the_artificial_intelligence_era
- Nass, C., & Moon, Y. (2023). Machines and mind perception: When users believe computers care. *Communication Research*, 50(5), 621-642.
- Norman, D. A. (2023). *The design of future things* (updated edition). MIT Press.
- Picard, R. W. (1997). *Affective Computing*. MIT Press.
- Pirjan, A., & Petroşanu, D. M. (2025). Artificial social intelligence and the transformation of human interaction by artificial intelligence agents. *Journal of Information Systems and Operations Management*, 19(1), 267-285. <https://web.rau.ro/websites/jisom/Vol.19%20No.1%20-%202025/IISOM%2019.1.pdf#page=267>

- Pulivarthi, P., & Bhatia, A. B. (2025). Designing empathetic interfaces: Enhancing user experience through emotion. IGI Global.
- Rahaman, S. U., Abdul, M. J., & Patchipulusu, S. (2023). AI-driven empathy in UX design: Enhancing personalization and user experience through predictive analytics. ResearchGate.
- Rasheed, J., & Kiani, M. N. (2024). Measuring the impact of perceived authentic leadership on innovative work behavior: The mediation moderation analysis. *Journal of Business and Management Research*, 3(3), 352-376.
- Rawat, S. (2025). The role of AI chatbots in providing mental health support: Emotional perception, human–AI interaction, and ethical implications. *International Journal of Technological Innovation*, 12(3), 133–159.
- https://ijtinnovation.com/admin/assets/article_issue/1744893456IJTI_MARCH_APRIL_2025.pdf
- Reeves, B., & Nass, C. (2023). The media equation revisited: Empathy and trust in machine communication. *Computers in Human Behavior*, 145, 106772.
- Reeves, B., Hancock, J., & Ritchie, D. (2020). The social impact of emotional AI. *Nature*, 587(7835), 26–28.
- Rostami, M., & Navabinejad, S. (2023). Artificial empathy: User experiences with emotionally intelligent chatbots. *Journal of Empirical Psychology and AI*, 14(2), 88–104.
https://www.researchgate.net/publication/379042128_Artificial_Empathy_User_Experiences_with_Emotionally_Intelligent_Chatbots
- Salah, A. A., & El Ali, A. (2025). Affective user interfaces. In *Handbook of Affective Computing and Interaction* (pp. 421–437). Springer.
- Saxena, C. (2024). Ethical considerations in affective computing. In *Human-Centered AI Systems Design* (pp. 153–178). Springer.
https://link.springer.com/chapter/10.1007/978-3-031-63821-3_13
- Shabbir, T., Abro, M. U. R., Gillani, A. H., Uddin, S. S., & Ansari, U. (2021). An Analysis to the Metaphors Used in Political News Coverage (A Context to Pakistani Private News Channel). *Turkish Online Journal of Qualitative Inquiry*, 12(7).
- Shin, D. (2021). The effects of explainability and transparency on trust in AI. *Computers in Human Behavior*, 115, 106568.
- Smith, A., & Zhao, Y. (2024). Transparency and deception in AI empathy simulation. *Journal of AI Ethics*, 5(1), 44–59.
- Stark, L., & Hutson, J. (2021). Emotional manipulation in digital systems. *ACM Journal of Ethics in Digital Society*, 1(1), 1–14.
- Suresh, H., & Guttag, J. (2024). Understanding and mitigating algorithmic harm in emotional AI systems. *Communications of the ACM*, 67(9), 30–40.
- Tsarikova, M. (2025). Enhancing user experience design through artificial intelligence technologies. Theseus Open Repository.

https://www.theseus.fi/bitstream/handle/10024/893272/Tsarikova_Marina.pdf?sequence=2

- Turkle, S. (2023). The empathy machine: Reconsidering intimacy with AI companions. *MIT Technology Review*, 127(4), 55–67.
- un Nisa, S. (2025). THE NEW WORLD ORDER: BRICS AND THE PERSPECTIVE OF PAKISTAN. *Policy Journal of Social Science Review*, 3(6), 107-113.
- Vallor, S. (2016). *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. Oxford University Press.
- Vallor, S. (2024). Moral deskilling in the age of emotional AI. *Ethics and Information Technology*, 26(2), 173–190.
- Virvou, M. (2023). Artificial intelligence and user experience in reciprocity: Contributions and state of the art. *Interactive Digital Technologies*, 31(1), 1–14. <https://journals.sagepub.com/doi/pdf/10.3233/IDT-230092>
- Walden, S. (2025). Empathy, authenticity, and deception in conversational AI. *Human–Technology Review*, 22(2), 187–203.
- Weidinger, L., et al. (2023). Ethical and societal implications of large language models. *Nature Machine Intelligence*, 5(7), 634–646.
- West, S., & Allen, C. (2024). Algorithmic care: Designing empathy into AI systems responsibly. *AI & Society*, 39(3), 523–537.
- Yi, L. (2024). Empathic AI: A multidisciplinary approach to computational empathy for understanding and enhancing human–AI interaction. Seoul National University. <https://s-space.snu.ac.kr/handle/10371/215257>
- Zaheer, S. (2024). Ethical UX design: Preventing manipulative interfaces and promoting user trust. Academia.edu. https://www.researchgate.net/publication/391019688_Ethical_UX_Design_Preventing_Manipulative_Interfaces_and_Promoting_User_Trust
- Zhu, Q., & Luo, J. (2024). Toward artificial empathy for human-centered design. *ASME Journal of Mechanical Design*, 146(6), 061401. <https://asmedigitalcollection.asme.org/mechanicaldesign/article-abstract/146/6/061401/1171651>
- Zuboff, S. (2023). The age of surveillance empathy: Emotional capitalism and predictive UX. *Harvard Business Review*, 101(6), 72–89.