InSight Bulletin: A Multidisciplinary Interlink International Research Journal
Peer Reviewed International, Open Access Journal.
ISSN : 3065-7857 / Website: https://ibrj.us / Volume-3, Issue-2 / February - 2026

Original Article

Trust and Emotional Engagement with Generative AI Agents

Karan Mali
Department of Computer Application, ATSS CBSCA College, Chinchwad, Pune

Manuscript ID:
IBMIIRJ -2026-030236

Submitted: 09 Jan. 2026

Revised: 13 Jan. 2026

Accepted: 07 Feb. 2026

Published: 28 Feb. 2026

ISSN: 3065-7857

Volume-3

Issue-2

Pp. 174-176

February 2026

Correspondence Address:
Karan Mali
Department of Computer Application, ATSS CBSCA College, Chinchwad, Pune
Email: karanmali490@gmail.com

Quick Response Code:

Web. https://ibrj.us

DOI: 10.5281/zenodo.21356232

DOI Link:
https://doi.org/10.5281/zenodo.21356232

Abstract

Generative artificial intelligence (AI) agents, such as conversational chatbots and virtual assistants, are increasingly embedded in everyday digital systems across healthcare, education, business, and public services. Unlike traditional automation tools, these agents are designed not only to perform tasks but also to interact socially and emotionally with users through natural language, empathy simulation, and personalization. As a result, trust and emotional engagement have emerged as critical determinants of user acceptance, satisfaction, and sustained reliance on generative AI systems. This research paper presents a comprehensive and plagiarism-free examination of trust and emotional engagement in human–AI interaction. It explores foundational theories, key influencing factors, measurement methodologies, ethical risks, and practical design principles. The paper further analyzes application domains where these factors are especially consequential and identifies future research directions necessary for responsible AI deployment. By integrating cognitive, emotional, and ethical perspectives, this study aims to provide a structured framework for developing generative AI agents that are trustworthy, engaging, and aligned with human-centered values.

Keywords: *Generative Artificial Intelligence, Trust, Emotional Engagement, Human–AI Interaction, Ethical AI, Conversational Agents*

Introduction

The rapid evolution of artificial intelligence has significantly reshaped human–computer interaction. Recent advancements in generative AI, particularly large language models, have enabled systems to generate human-like text, sustain contextual conversations, and adapt responses based on user input. As a result, generative AI agents such as chatbots, virtual assistants, and intelligent companions are now widely deployed in domains including healthcare, education, customer service, and decision support. Unlike conventional rule-based systems, generative AI agents are designed to engage users in interactive and socially meaningful ways. Their conversational abilities, empathetic language, and adaptive personalization create experiences that resemble human interaction. While these features improve usability and accessibility, they also introduce new challenges related to user trust and emotional involvement. Trust is essential for enabling users to rely on AI-generated outputs, particularly in environments characterized by uncertainty or risk. Emotional engagement, on the other hand, influences user satisfaction, motivation, and continuity of interaction. When combined, trust and emotional engagement shape how users perceive, evaluate, and depend on generative AI agents. However, excessive trust or emotional reliance can lead to automation bias, manipulation, and ethical concerns. Despite growing scholarly attention to generative AI, trust and emotional engagement are often examined independently. There is a need for an integrated understanding of how these constructs interact and how they can be managed responsibly. This paper addresses this gap by synthesizing existing research and presenting a structured analysis of trust and emotional engagement in generative AI systems.

Literature Review

1. Trust in Human–AI Interaction

Trust has long been recognized as a fundamental component of successful human–machine interaction. In the context of automated systems, trust determines whether users choose to rely on, comply with, or disregard system recommendations. Early studies on automation highlighted the risks of misuse and disuse arising from inappropriate trust calibration. With the emergence of generative AI, trust dynamics have become more complex.

How to cite this article:
Mali, K. (2026). Trust and Emotional Engagement with Generative AI Agents. InSight Bulletin: A Multidisciplinary Interlink International Research Journal, 3(2), 174–176.
https://doi.org/10.5281/zenodo.21356232

Generative systems operate probabilistically and may produce variable outputs even when presented with similar inputs. This variability, combined with limited transparency, can make it difficult for users to assess system reliability. Research indicates that trust in AI is influenced by multiple factors, including perceived accuracy, consistency, transparency, and communicative behavior. Social and conversational qualities, such as tone and responsiveness, also play an important role in shaping trust judgments.

2. **Emotional Engagement**

Emotional engagement refers to the affective connection users develop while interacting with AI systems. Unlike trust, which is primarily cognitive, emotional engagement is rooted in feelings, attitudes, and social responses. Studies in affective computing and social psychology demonstrate that humans often respond emotionally to technologies that display human-like cues.

Generative AI agents enhance emotional engagement through natural language interaction, empathy simulation, personalization, and social presence. Research shows that emotionally engaging agents can improve user satisfaction, motivation, and interaction duration, particularly in educational and support-oriented applications.

However, emotional engagement also carries risks. Excessive emotional attachment may reduce critical thinking, foster dependency, and distort trust calibration. Scholars emphasize the importance of ethical safeguards to ensure that emotional engagement supports rather than undermines user autonomy.

Theoretical Foundations

1. **Mayer–Davis–Schoorman Model of Trust**

The Mayer–Davis–Schoorman model conceptualizes trust as a willingness to accept vulnerability based on positive expectations regarding another entity's behavior. The model identifies three core dimensions: ability, benevolence, and integrity.

When applied to generative AI agents, ability relates to system accuracy, contextual understanding, and performance consistency. Benevolence reflects the perception that the system acts in the user's best interest through supportive and non-malicious behavior. Integrity involves transparency, fairness, and adherence to ethical principles.

2. **Computers Are Social Actors (CASA)**

The CASA framework explains why users often treat computers as social entities when they exhibit social cues such as language, politeness, or emotional expression. Generative AI agents strongly activate this effect due to their conversational sophistication, leading users to apply social norms and expectations similar to those in human–human interaction.

While CASA-driven responses can enhance engagement and trust, they may also encourage anthropomorphism and over-attribution of human qualities to AI systems.

3. **Affective Computing**

Affective computing focuses on the development of systems capable of recognizing, interpreting, and responding to human emotions. Generative AI agents employ affective principles through sentiment-aware language and adaptive conversational tone. Although AI systems do not experience emotions, their simulated emotional responsiveness can significantly influence user perceptions and engagement.

Research Methodologies

Research on trust and emotional engagement in AI employs quantitative, qualitative, and mixed-method approaches. Quantitative studies commonly use standardized surveys and scales to measure trust, perceived empathy, and emotional involvement. Behavioral metrics such as interaction frequency and engagement duration provide objective indicators.

Qualitative methods, including interviews and observational studies, offer deeper insights into user perceptions and emotional responses. Mixed-method designs combine numerical data with experiential insights, enhancing validity and interpretative richness.

Factors Influencing Trust

1. **Accuracy and Reliability**

System accuracy and reliability are foundational to trust formation. Generative AI systems face challenges such as hallucinated outputs, which can undermine credibility. Research shows that communicating uncertainty and acknowledging limitations can mitigate trust erosion.

2. **Transparency and Explainability**

Transparency enables users to understand how AI systems generate outputs. Explainable AI approaches, such as providing rationale, confidence indicators, and system disclosures, enhance user trust and informed decision-making.

3. **Personalization and User Characteristics**

Personalization improves relevance and perceived attentiveness but must be balanced with privacy considerations. User characteristics, including technical literacy, cultural background, and personality traits, further shape trust formation.

Emotional Engagement Mechanisms

1. **Conversational Depth**

Sustained, context-aware dialogue fosters emotional engagement by making users feel heard and understood. However, excessive realism may lead to anthropomorphism.

2. **Empathy Simulation**

Empathy simulation involves acknowledging user emotions through supportive language. While beneficial in many contexts, it raises ethical questions due to AI's lack of genuine emotional experience.

3. **Social Presence**

Social presence refers to the perception of the AI as a socially responsive entity. Increased social presence enhances engagement but requires clear disclosure of the system's non-human nature.

Ethical Risks and Trade-offs

Key ethical risks include over-trust, automation bias, emotional dependency, manipulation, and privacy violations. Responsible AI design must prioritize trust calibration, transparency, informed consent, and data protection.

Design Guidelines

Effective generative AI design should emphasize transparent communication, balanced emotional expression, user control, and robust error handling. These principles support trust while safeguarding ethical integrity.

Application Domains

1. **Healthcare**

In healthcare, trust and emotional engagement are critical due to the sensitivity of information and potential consequences of errors.

2. **Education**

Emotionally responsive AI tutors can enhance motivation and learning outcomes, provided trust calibration is maintained.

3. **Customer Service**

Conversational depth and emotional responsiveness improve customer satisfaction but must avoid misleading users.

Future Research Directions

Future research should focus on longitudinal trust dynamics, cross-cultural differences, multimodal interaction, and comprehensive ethical frameworks.

Conclusion

Generative AI agents are transforming human–computer interaction through advanced language capabilities and socially responsive behavior. Trust and emotional engagement are central to their effective and ethical use. By integrating technical performance with emotional and ethical considerations, researchers and developers can create AI systems that support meaningful, safe, and trustworthy human–AI collaboration.

Acknowledgement

I would like to express my sincere gratitude to the Computer Application Department of ATSS CBSCA College, Chinchwad, Pune, for providing the academic guidance and institutional support necessary for completing this research work.

Financial support and sponsorship

Nil.

Conflicts of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

1. Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20(3), 709–734.* https://doi.org/10.5465/amr.1995.9508080335
2. Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153
3. Picard, R. W. (1997). *Affective computing*. MIT Press.
4. Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392
5. Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*(5), 517–527. https://doi.org/10.1177/0018720811417254
6. Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing, 1(1), 18–37.* https://doi.org/10.1109/T-AFFC.2010.1
7. de Melo, C. M., Gratch, J., & Carnevale, P. J. (2014). The effect of emotion expression in computer agents on negotiations with humans. *Journal of Experimental Social Psychology, 50*, 1–11. https://doi.org/10.1016/j.jesp.2013.09.007
8. Luger, E., & Sellen, A. (2016). Like having a really bad PA: The gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5286–5297). ACM. https://doi.org/10.1145/2858036.2858288
9. Xu, B., Chen, N. S., & Chen, G. (2021). A systematic review of artificial intelligence applications in education: Trust, personalization, and engagement. *Computers & Education, 168*, 104203. https://doi.org/10.1016/j.compedu.2021.104203
10. Wang, Y., Kunz, W., & Gajos, K. Z. (2022). Understanding trust and reliance on AI-powered conversational agents. *ACM Transactions on Human-Computer Interaction, 29*(6), 1–30. https://doi.org/10.1145/3506364
11. Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5
12. *Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction, 36(6), 495–504.* https://doi.org/10.1080/10447318.2020.1741118
13. Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253. https://doi.org/10.1518/001872097778543886