JOURNAL OF HUMAN–AI INTERACTION AND SOCIAL BEHAVIOUR

VOL.1-NO.1-2026

P ISSN: xxxx-xxx

E ISSN: xxxx-xxx

HUMAN–AI INTERACTION AND ITS INFLUENCE ON SOCIAL BEHAVIOUR AND TRUST

Rafey Konain

BS English Literature, Institute of English Studies (IES), University of the Punjab Lahore.

ABSTRACT

Artificial intelligence (AI) is increasingly integrated into everyday life through conversational agents, recommendation systems, virtual assistants, automated decision-making tools, and collaborative technologies. These systems influence how people communicate, make decisions, exchange information, and develop trust in digital environments. Human–AI interaction has therefore become an important area of research involving technology, psychology, communication, and social behaviour. This article examines how interactions with AI systems influence human trust, social responses, communication patterns, and decision-making behaviour. Existing research indicates that trust in AI is shaped by perceived reliability, transparency, explainability, privacy, system performance, and the way AI communicates with users (Mehrotra et al., 2024; Ng & Zhang, 2025). Human-like conversational features may encourage engagement and positive social responses, but they may also lead users to overestimate an AI system's understanding or capabilities (Klein, 2025). The article adopts a literature-based approach to synthesize existing scholarly findings concerning the social and behavioural consequences of human–AI interaction. The review highlights the importance of appropriate rather than unquestioning trust, responsible system design, transparency, and human oversight. It concludes that AI can support effective communication and collaboration when its capabilities and limitations are clearly understood, while responsible interaction practices are necessary to protect human autonomy, social relationships, and informed decision-making.

KEYWORDS

Human–AI interaction, artificial intelligence, social behaviour, trust, human–computer interaction, AI communication, responsible AI.

INTRODUCTION

Artificial intelligence has become an increasingly important part of modern social and professional life. AI-powered systems are used in education, healthcare, customer services, workplaces, online communication, entertainment, and decision-making. Conversational agents can answer questions, explain complex information, assist with writing, and simulate natural dialogue, while recommendation systems influence the information, products, and media people encounter. As these systems become more interactive, the relationship between humans and AI extends beyond the use of technical tools and increasingly involves social responses, expectations, trust, and behavioural adaptation. Henrique and Santos (2024) explain that trust is a central issue in AI adoption because users may either reject useful systems due to insufficient trust or rely too heavily on systems whose recommendations are inaccurate. This makes human–AI interaction an important topic for understanding how technological systems influence social behaviour and human judgement.Trust is one of the most significant factors shaping human responses to AI. Users may trust an AI system when they believe it is competent, reliable, consistent, and capable of providing useful information. However, trust can be weakened by inaccurate responses, unexplained recommendations, privacy concerns, or unexpected system behaviour. Mehrotra et al. (2024) reviewed research on appropriate trust in human–AI interaction

and highlighted the importance of explanations, confidence information, trustworthiness cues, and communication of uncertainty. Their work indicates that successful AI interaction does not require users to trust every system unconditionally. Instead, users should develop expectations that match the system's actual capabilities and limitations. This distinction is important because excessive trust may encourage people to accept incorrect recommendations, whereas insufficient trust may prevent them from benefiting from reliable AI assistance. Human–AI interaction can also influence social behaviour by shaping how people communicate, respond to suggestions, and interpret the intentions of interactive systems. Conversational AI often uses language patterns, polite expressions, personalized responses, and other features associated with human communication. These features may make interactions feel more natural and encourage people to continue using the system. Klein (2025) examined the effects of human-like social cues in text-based conversational agents and found an overall positive effect on social responses, although the magnitude varied across outcomes and circumstances. Such findings suggest that people may respond socially to AI systems even when they know they are interacting with technology. However, a natural conversational style should not be confused with genuine human understanding, emotion, or personal intention. This distinction matters when people use AI for advice, companionship, education, or other interactions involving sensitive decisions.

Another important aspect of human–AI interaction is the influence of system design on users' perceptions and behaviour. Transparency, explanation quality, interface design, reliability, privacy protection, and perceived control can influence whether users consider AI systems trustworthy. Leschanowsky et al. (2024) reviewed research on privacy, security, and trust perceptions in conversational AI, emphasizing that these factors are closely connected to how users evaluate and engage with AI systems. If users do not understand how their information is collected or how a system produces its responses, they may experience uncertainty or become reluctant to share information. Conversely, clear communication about data use, limitations, and system functions can help people make more informed decisions. Trust should therefore be considered not only a personal attitude but also an outcome influenced by the design, governance, and communication practices of AI technologies. Human–AI interaction is also becoming increasingly important in collaborative settings where people and AI systems work together to complete tasks. In workplaces, AI may support analysis, planning, information retrieval, and decision-making, while humans contribute contextual understanding, ethical judgement, and responsibility. Duan et al. (2025) examined trust and distrust within human–AI teams and found that trust-related perceptions can spread through social processes, while misinformation and misalignment can contribute to distrust. Their research suggests that the behaviour of AI systems can affect not only an individual's trust in a tool but also communication and coordination within a team. These issues are relevant to organizations that use AI to assist employees, distribute tasks, or provide recommendations. Effective collaboration requires users to understand the system's role and to retain the ability to question, verify, or override its suggestions.

The rapid growth of conversational AI has also raised questions about emotional and relational engagement. Some people use AI systems to seek explanations, rehearse conversations, explore personal concerns, or obtain supportive responses. Such interactions may provide convenience

and a non-judgmental space for expressing thoughts, but they can also create unrealistic expectations about the system's emotional understanding or reliability. Ng and Zhang (2025), in their systematic review of trust in AI chatbots, describe trust as involving behavioural, emotional, cognitive, relational, and contextual dimensions. Their review also identifies variation in how trust is defined and measured across studies. These findings demonstrate that trust in AI is not determined by technical performance alone; it can also be shaped by interaction style, user characteristics, social expectations, and the context in which the system is used. Despite the growing use of AI, important challenges remain in ensuring that human–AI interaction supports responsible social behaviour. Users may accept AI-generated information without sufficient verification, disclose personal information without fully understanding privacy implications, or interpret conversational fluency as evidence of competence. On the other hand, poorly designed systems and unclear explanations may lead users to reject accurate assistance. Dang and Li (2025) synthesized evidence from 560 empirical studies of human trust in AI, identifying system capability, anthropomorphism, individual factors, and explainability among the factors associated with trust. Their review also considered outcomes such as attitudes, perceived usefulness, and behavioural intentions. This evidence supports the view that trust develops through a combination of technical characteristics, human expectations, and interaction conditions rather than through one isolated factor. In this context, the present article examines the influence of human–AI interaction on social behaviour and trust. It considers the ways AI systems shape users' expectations, communication practices, decision-making, and collaborative relationships. It also explores the importance of transparency, privacy, reliability, and appropriate trust in supporting responsible engagement. By synthesizing existing literature, the article aims to provide a balanced understanding of the opportunities and challenges associated with AI-mediated interaction and to identify principles that may help ensure that AI systems complement rather than undermine human judgement and social responsibility.

RESEARCH OBJECTIVES

1. To examine how human–AI interaction influences social behaviour and communication patterns.
2. To analyse the factors that shape human trust in artificial intelligence systems.
3. To identify responsible interaction practices that support appropriate trust and human judgement.

RESEARCH QUESTIONS

1. How does human–AI interaction influence social behaviour and communication patterns?
2. Which factors shape human trust in artificial intelligence systems?
3. What practices can support appropriate trust and responsible human–AI interaction?

LITERATURE REVIEW

Research on trust in artificial intelligence has expanded as AI systems have become more common in everyday activities and organizational decision-making. Mehrotra et al. (2024) conducted a systematic review of approaches used to foster appropriate trust in human–AI

interaction. Their review discussed techniques such as explanations, confidence scores, trustworthiness cues, and communication of uncertainty. The authors emphasized that trust is a complex concept and that researchers do not always define or measure it consistently. Their findings support the importance of matching user expectations to system capabilities. When users understand what an AI system can and cannot do, they may be better positioned to rely on it appropriately instead of accepting every recommendation or rejecting the technology entirely.Henrique and Santos (2024) examined trust in AI through a structured literature review and analysis of influential research. They identified the importance of balancing trust and reliance, since users may sometimes distrust systems that could provide useful assistance or place excessive confidence in automated recommendations. This balance is especially important in areas where AI outputs influence consequential decisions. Their work indicates that trust should not be treated simply as a measure of acceptance; it should also reflect whether the system is dependable for the task at hand. This perspective provides a useful foundation for examining social behaviour because the level of trust users place in AI can influence whether they seek additional information, question an answer, or follow a recommendation.Leschanowsky et al. (2024) reviewed studies concerning privacy, security, and trust perceptions in conversational AI. Their work highlights the importance of understanding how users evaluate systems that communicate through text or voice and may request or process personal information. Privacy concerns can affect the willingness to interact with AI, particularly when users are uncertain about how their data are stored, used, or protected. Security and transparency are therefore relevant not only to technical system quality but also to the social relationship between users and AI services. The review supports the need for clear privacy practices, understandable system behaviour, and responsible data management as part of trustworthy interaction.

Ng and Zhang (2025) conducted a systematic review of trust in AI chatbots, analysing 40 articles. Their review found substantial variation in how trust was conceptualized and measured. The factors associated with trust were grouped around users, machines, interaction, social conditions, and context. The review also discussed emotional, relational, behavioural, cognitive, and psychological outcomes associated with chatbot trust. These findings are particularly relevant to social behaviour because conversational AI can be experienced as more than a simple information tool. Users may respond to conversational style, perceived responsiveness, and the apparent social qualities of the system. Nevertheless, variation in research methods and definitions means that findings should be interpreted carefully rather than treated as universal conclusions about every chatbot or user.Klein (2025) examined the effects of human-like social cues on responses to text-based conversational agents through a meta-analysis. The study included 142 papers and reported a small overall positive effect of human-likeness on social responses, with differences across outcomes and contextual conditions. This evidence suggests that conversational features such as social language and human-like presentation can affect users' attitudes, perceptions, affect, rapport, trust, and behaviour. Such features may make AI systems easier or more comfortable to engage with, but their influence also raises questions about whether users could mistake conversational fluency for human understanding. Therefore, social cues should be designed responsibly, with clear distinctions between an AI system and a human interlocutor.

Duan et al. (2025) investigated trust and distrust within human–AI teams using a qualitative approach. Their findings indicate that trust-related attitudes can influence team processes and that distrust may spread through social interaction, including through misinformation. The research suggests that the success of human–AI collaboration depends partly on whether people and AI systems are aligned in their goals, expectations, and contributions. This is important for organizational settings in which employees rely on AI-generated recommendations or share AI-supported decisions with colleagues. The study highlights that trust is not merely an individual response to a technical system; it can also affect coordination, communication, and collective performance.Dang and Li (2025) reviewed 560 empirical studies to investigate the antecedents and consequences of human trust in AI. Their review identified factors such as system capability, anthropomorphism, individual characteristics, and explainability as important influences on trust. The authors also considered consequences such as behavioural intentions, attitudes, perceived usefulness, and acceptance. This broad evidence base reinforces the idea that trust develops through interactions between system features and human perceptions. It also suggests that the design of AI systems should account for how people interpret and respond to automated outputs, rather than focusing exclusively on technical performance.Overall, the reviewed literature indicates that human–AI interaction can influence social responses, decision-making, communication, and collaboration through several interconnected mechanisms. System reliability and transparency affect confidence in AI outputs, while conversational and human-like features may influence social perceptions. Privacy concerns, unclear explanations, and unreliable responses may weaken trust, whereas appropriate explanations and realistic expectations can support more informed use. However, the literature also shows differences in definitions, measurement approaches, and research contexts. These limitations support the need for a balanced approach that considers both the potential benefits of AI interaction and the risks associated with excessive reliance, misleading social cues, and insufficient human oversight.

METHODOLOGY

This article uses a literature-based methodology to examine the influence of human–AI interaction on social behaviour and trust. Relevant scholarly publications were identified through academic search sources, including Google Scholar, and through peer-reviewed journal and research publications addressing trust in AI, conversational agents, social responses, privacy, explainability, and human–AI collaboration. Priority was given to relevant research published between 2024 and 2025, including systematic reviews, meta-analyses, and empirical studies. Publications were considered for their relevance to the research objectives, clarity of research focus, and contribution to understanding human responses to AI systems. The selected literature was examined to identify recurring findings and contrasting perspectives concerning trust, communication, behavioural intentions, and responsible AI use. Since the article synthesizes existing research rather than collecting new data from participants, its findings represent a thematic interpretation of published evidence and not original experimental or survey results.

RESEARCH DESIGN

The study follows a descriptive narrative review design. This design is appropriate because human–AI interaction involves multiple dimensions, including technical reliability, social

communication, privacy, user expectations, and interpersonal collaboration. The literature was organized into themes related to social behaviour and communication, factors influencing trust, and practices that support responsible interaction. Findings from the selected publications were compared to identify common patterns, differences in interpretation, and limitations in the existing evidence. The design allows the article to bring together research from different disciplines without claiming that the included studies use identical measures or are directly comparable. It also provides a basis for discussing the practical implications of existing research for AI design and human decision-making.

DATA ANALYSIS

The selected literature was analysed through thematic synthesis to identify recurring patterns and relationships across the studies. The analysis focused on three main themes: the influence of AI interaction on social behaviour, the factors shaping trust, and the practices that support responsible human–AI engagement. Findings relating to reliability, explainability, privacy, anthropomorphism, communication style, and user expectations were compared across the selected publications. The analysis also considered how trust may influence acceptance of AI recommendations, willingness to interact with conversational systems, and collaboration in social or professional settings. Differences in study methods, definitions, and contexts were considered when interpreting the findings. No original statistical tests or participant-based data collection were conducted; therefore, the analysis represents a qualitative synthesis of existing scholarly research.

FINDINGS

The reviewed literature indicates that human–AI interaction can influence social behaviour by changing how individuals communicate, seek information, evaluate recommendations, and participate in collaborative activities. Conversational AI systems may encourage engagement by providing immediate responses, using natural language, and adapting their communication to users' requests. Human-like social cues can influence users' attitudes and responses, although their effects vary across interaction settings and outcomes (Klein, 2025). These findings suggest that people may respond socially to AI systems even when they recognize that the interaction is technological rather than human. However, the quality of such interaction depends on whether the system provides useful information, communicates its limitations, and supports the user's actual needs.A second major finding is that trust depends on multiple interacting factors rather than a single characteristic of AI. System capability, reliability, explainability, privacy, individual expectations, and interaction design can all influence how users evaluate AI systems (Mehrotra et al., 2024; Dang & Li, 2025). When AI systems provide consistent and understandable responses, users may be more willing to use them. However, inaccurate information, unclear reasoning, and uncertainty about data handling may weaken trust. The literature also suggests that trust should be appropriate to the task and the system's actual performance. Excessive trust can encourage uncritical acceptance of inaccurate outputs, while insufficient trust may prevent users from benefiting from useful assistance (Henrique & Santos, 2024).

The literature further indicates that conversational AI creates social and relational considerations beyond technical performance. Users may perceive chatbots as approachable because they respond quickly and use familiar conversational language. Such features can support interaction, but they may also encourage unrealistic assumptions about the system's understanding, intentions, or emotional capabilities. Ng and Zhang (2025) found that research on chatbot trust covers emotional, relational, behavioural, cognitive, and psychological dimensions. This suggests that trust in conversational AI involves not only whether users believe a system is accurate but also how they experience and interpret the interaction. Clear communication about the nature and limitations of AI is therefore important for helping users maintain realistic expectations.Another finding concerns privacy and responsible information sharing. Users may be reluctant to interact with AI when they do not understand how their information is collected, processed, or protected. Research on conversational AI highlights the importance of privacy, security, and trust perceptions in shaping user engagement (Leschanowsky et al., 2024). Transparent data practices and understandable explanations can help users make informed decisions about what information to share. Nevertheless, transparency must be meaningful: simply presenting technical information may not be sufficient if users cannot understand its implications. Responsible AI interaction therefore requires attention to the clarity of communication and the protection of personal information.

The review also indicates that trust has implications for teamwork and collective behaviour. In collaborative environments, AI may contribute analysis, recommendations, and information processing, while humans provide contextual judgement, ethical reasoning, and accountability. Duan et al. (2025) found that trust and distrust can spread through social processes within human–AI teams and that misinformation may contribute to distrust. These findings suggest that AI-related trust can affect communication and coordination beyond an individual's interaction with a system. Organizations using AI should therefore establish clear responsibilities, encourage verification of important outputs, and maintain human oversight where decisions carry significant consequences.Overall, the findings show that human–AI interaction can support communication, information access, and collaboration, but its influence on social behaviour and trust depends on system design, user expectations, and the context of use. The literature does not support the assumption that increasing trust is always desirable. Instead, the goal should be to develop informed and appropriately calibrated trust, allowing users to benefit from AI while retaining independent judgement. Transparent communication, reliable performance, privacy protection, and clear human responsibility emerge as important principles for responsible interaction.

CONCLUSION

Human–AI interaction is becoming an important part of contemporary social behaviour, influencing communication, information-seeking, decision-making, and collaboration. This article reviewed existing scholarly literature to examine how AI systems shape social responses and how trust develops during interaction. The reviewed research indicates that conversational features can encourage engagement, while system reliability, explainability, privacy, and user expectations influence the level of trust people place in AI. These findings demonstrate that trust

is a multidimensional concept involving technical, psychological, social, and contextual factors.The article also highlights that trust in AI should be appropriate rather than unconditional. Users need to understand the capabilities and limitations of AI systems and retain the ability to question, verify, or reject their recommendations. Human-like communication may make AI interaction more comfortable, but it should not create misleading expectations about the system's understanding or emotional abilities. Similarly, privacy protection and transparent data practices are essential for responsible interaction. In collaborative settings, clear responsibilities and human oversight can help ensure that AI contributes to decision-making without replacing necessary human judgement.In conclusion, AI can support useful communication and social collaboration when systems are designed responsibly and users are encouraged to engage critically. Developers should prioritize reliability, explainability, privacy, and clear communication, while educators and organizations should promote AI literacy and informed decision-making. Future research should investigate how trust develops over time, how different interaction designs affect social behaviour, and how AI influences relationships across different populations and cultural contexts. A balanced approach to human–AI interaction can help society benefit from technological innovation while preserving human autonomy, accountability, and responsible social behaviour.

REFERENCES

Dang, Q., & Li, G. (2025). Antecedents and consequences of human trust in AI: Evidence from a systematic review. Academy of Management Proceedings, 2025(1).

Duan, W., Zhou, S., Scalia, M. J., Freeman, G., Gorman, J., Tolston, M., McNeese, N. J., & Funke, G. (2025). Understanding the processes of trust and distrust contagion in human–AI teams: A qualitative approach. Computers in Human Behavior, 165, Article 108560.

Henrique, B. M., & Santos, E., Jr. (2024). Trust in artificial intelligence: Literature review and main path analysis. Computers in Human Behavior: Artificial Humans, 2(1), Article 100043.

Klein, S. H. (2025). The effects of human-like social cues on social responses towards text-based conversational agents—a meta-analysis. Humanities and Social Sciences Communications, 12, Article 1322.

Leschanowsky, A., Rech, S., Popp, B., & Bäckström, T. (2024). Evaluating privacy, security, and trust perceptions in conversational AI: A systematic review. Computers in Human Behavior, 159, Article 108344.

Mehrotra, S., Degachi, C., Vereschak, O., Jonker, C. M., & Tielman, M. L. (2024). A systematic review on fostering appropriate trust in human-AI interaction: Trends, opportunities and challenges. ACM Journal on Responsible Computing, 1(4), Article 26.

Ng, S. W. T., & Zhang, R. (2025). Trust in AI chatbots: A systematic review. Telematics and Informatics, 97, Article 102240.