

Artificial Intelligence-Mediated Communication and Human Relationships: Examining Trust, Social Presence, and Interpersonal Connectedness among University Students

YunHaoChuang^{1*}, Dody Riswanto², Syimir Asyraaf bin Eddie Halim³

¹ National Chi Nan University, Taiwan

² Universitas Indraprasta PGRI Jakarta, Indonesia

³ Universiti Sains Malaysia, Malaysia

email: Winhau89310@gmail.com¹

Article Info :

Received:

27-07-2026

Revised:

07-08-2026

Accepted:

17-08-2026

Abstract

Artificial intelligence-mediated communication introduces new forms of technological participation in interpersonal interaction, raising questions about trust, social presence, and relational connectedness in higher education. This study examined how existing scholarship characterizes these relationships and identified the conditions associated with variation in relational outcomes. A scoping review and framework-based relational mechanism synthesis were conducted using 34 scholarly sources published between 1976 and 2026, identified through bibliographic and publisher-based searches. The evidence base included experimental, cross-sectional, longitudinal, mixed-method, design-oriented, and conceptual investigations. Sources were examined according to communication configuration, degree of AI assistance, disclosure conditions, relational outcomes, and contextual moderators. The synthesis identified three interconnected domains: trust calibration and authenticity, social presence and conversational agency, and interpersonal connectedness in educational communication. Across the reviewed literature, AI assistance was associated with communicative efficiency and responsiveness, while relational evaluations varied according to perceived authenticity, transparency, technological reliability, and human control. Conversational and anthropomorphic cues contributed to perceived social presence without consistently establishing durable interpersonal trust. Evidence involving university students particularly emphasized AI literacy, perceived accuracy, risk perception, and technology adoption, whereas broader relational interpretations drew partly on other populations. The study proposes calibrated trust as an interpretive principle emphasizing proportionate reliance, verification, and human accountability. These findings support differentiated assessments of AI-mediated communication that distinguish technological performance, perceived social presence, and sustained interpersonal relationships.

Keywords: *AI-mediated communication, interpersonal connectedness, social presence, trust, university students.*



©2026 Authors.. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.

(<https://creativecommons.org/licenses/by-nc/4.0/>)

INTRODUCTION

Generative artificial intelligence (AI) has introduced new forms of technological participation in interpersonal communication by enabling automated message generation, linguistic revision, response formulation, and interaction management. Unlike conventional communication technologies that primarily facilitate the transmission of human-generated messages, generative AI systems can contribute directly to the linguistic and relational characteristics of communication. This development has become increasingly relevant in higher education, where students employ AI-supported applications for academic writing, language learning, information exchange, and communication with institutional and interpersonal contacts. Recent empirical research has examined university students' reliance on and resistance to generative AI in language learning, their perceptions of institutional messages concerning AI use, and the relationships among AI literacy, self-efficacy, trust, and technology adoption (Wang et al., 2025; Petricini et al., 2025; Obadă et al., 2026). Experimental evidence concerning AI-assisted writing has also demonstrated the relevance of algorithmic assistance to trust-building efficiency in interpersonal exchanges (Purcell et al., 2025). These developments extend established concerns in computer-mediated communication by introducing algorithmic participation into processes

traditionally associated with human communicative agency, message ownership, and relational accountability (Hancock et al., 2020; Hohenstein & Jung, 2020; Walther, 1992).

The relational implications of generative AI require a conceptual distinction among AI-mediated communication, AI-assisted human-to-human communication, and direct human–AI communication. AI-mediated communication is employed in this review as a broad analytical category encompassing communicative processes in which artificial intelligence contributes to message production, modification, interpretation, or interactional responsiveness. AI-assisted human-to-human communication represents a more specific configuration in which a human sender uses AI to formulate or revise messages intended for another human recipient while retaining responsibility for the communication. Direct human–AI communication involves interaction with an artificial conversational agent as the immediate communication partner. These configurations differ in the distribution of communicative agency, expectations of reciprocity, attribution of authorship, and objects of interpersonal trust. A student using AI to revise an email to a lecturer encounters relational conditions distinct from those experienced when seeking academic guidance or emotional support directly from a chatbot. Classical social presence theory and computer-mediated communication research provide foundations for examining these distinctions, although their original formulations did not fully account for systems capable of independently generating socially meaningful linguistic responses (Short et al., 1976; Biocca et al., 2003; Nass & Moon, 2000; Guzman & Lewis, 2020).

Trust constitutes a central dimension of AI-mediated communication because technological assistance introduces multiple objects of evaluation within the same interaction. Trust in an AI system concerns perceptions of its competence, accuracy, reliability, predictability, and capacity to perform an assigned communicative function. Trust in a human communicator concerns judgments about credibility, sincerity, responsibility, and the correspondence between expressed messages and the sender's intentions. These forms of trust may develop independently, particularly when algorithmically generated messages appear linguistically competent but provide limited information about the human sender's contribution. Research on technological credibility, human trust in artificial intelligence, and perceptions of algorithmic decision-making identifies system characteristics, perceived risk, and expectations of reliability as important determinants of trust judgments (Sundar, 2008; Glikson & Woolley, 2020; Lee, 2018). Empirical investigations have further associated trust with AI acceptance, perceived technological capabilities, and user evaluations of AI-related risks (Choung et al., 2023; Zhang & Yu, 2025). Evidence from higher education also indicates that trust in generative AI warrants examination across distinct academic user groups rather than treatment as a uniform response to technological availability (Đerić et al., 2025). These findings support an analytical separation between confidence in technological performance and confidence in the interpersonal relationships within which technological assistance is employed.

Social presence represents a related but analytically distinct dimension of AI-mediated interaction. It concerns the extent to which a communication environment conveys responsiveness, social awareness, interpersonal immediacy, and the perceived availability of an interaction partner. Conversational AI may support these perceptions through coherent responses, contextual adaptation, personalized language, and interaction patterns resembling human conversation. Nevertheless, perceived responsiveness does not necessarily establish interpersonal understanding or sustained relational attachment. Research on conversational agents has demonstrated the importance of anthropomorphic design, communicative agency framing, and conversational cues in shaping perceptions of human likeness and interaction quality (Araujo, 2018; Go & Sundar, 2019; Adam et al., 2021). Taxonomies of social cues further distinguish the design features through which conversational systems communicate social characteristics, while experimental research indicates that anthropomorphism can influence the resilience of trust in artificial agents (Feine et al., 2019; de Visser et al., 2016). The distinction between perceived social presence and relational durability remains important because user expectations may exceed system capabilities, and repeated interaction does not necessarily produce stable interpersonal attachment (Luger & Sellen, 2016; Croes & Antheunis, 2021). Research on explainability and trust also suggests that evaluations of AI systems depend on the interpretation of their underlying processes rather than conversational fluency alone (Shin, 2021).

Interpersonal connectedness introduces an additional analytical requirement because the quality of a relationship cannot be inferred exclusively from communication efficiency or perceived social presence. Connectedness involves perceived relational closeness, reciprocity, interpersonal recognition,

continuity of interaction, and confidence in the meaningful participation of another social actor. AI assistance may facilitate these conditions when it reduces linguistic barriers, supports message formulation, or enables individuals to maintain communication across demanding academic and social circumstances. The same assistance may also complicate relational evaluations when recipients question the authenticity of messages or the extent of human involvement in their production. The distinction is particularly relevant among university students, whose interactions encompass academic collaboration, peer relationships, feedback exchanges, institutional communication, and digitally mediated learning communities. Research on students' reliance on generative AI in language learning and their perceptions of teachers' communication about AI demonstrates the relevance of educational expectations and institutional norms to technology-related judgments (Wang et al., 2025; Petricini et al., 2025). Broader scholarship concerning AI-mediated communication further identifies epistemic trust and uncertainty about message origins as relational concerns that extend beyond the technical accuracy of generated content (Sahebi & Formosa, 2025). Interpersonal connectedness must consequently be examined as a relational outcome shaped by the interaction between communicative assistance, human participation, and the social expectations of communication partners.

Perceived authenticity and transparency further differentiate the relational consequences of AI-mediated communication. Authenticity concerns whether communicated content is interpreted as an adequate representation of the sender's intentions, knowledge, and interpersonal commitment. It does not necessarily require the complete absence of technological assistance, but it involves the perceived legitimacy of the relationship between human intention and message production. Transparency concerns the extent to which participants understand the involvement, function, and limitations of AI within an interaction. Although disclosure may support informed interpretation and accountability, explicit identification of AI assistance can also influence evaluations of message sincerity and communicative credibility. The relationship between transparency and trust is therefore conditional rather than uniformly positive. Research addressing algorithmic transparency signaling has examined its relationship with user trust and relational satisfaction, while scholarship on AI-mediated communication has identified tensions between disclosure, authenticity, and epistemic confidence (Park & Yoon, 2024; Sahebi & Formosa, 2025). These tensions are especially relevant when socially responsive systems generate expectations of intentionality, understanding, or emotional competence that are not necessarily supported by their operational characteristics. Human control over message revision, verification, and final authorization remains an important consideration when assessing whether technological assistance preserves or complicates communicative responsibility.

Despite the expansion of research on artificial intelligence and interpersonal communication, the available literature remains distributed across distinct conceptual frameworks, research populations, and methodological approaches. Studies of human–AI trust frequently emphasize technological acceptance, perceived accuracy, and reliance, whereas research on conversational agents often concentrates on anthropomorphism, interaction design, and social presence. Investigations of AI-assisted interpersonal communication address message attribution, disclosure, and authenticity, while studies involving university students commonly examine educational adoption, AI literacy, learning-related expectations, and institutional communication. These research domains provide complementary evidence but do not consistently distinguish trust in an AI system from trust in a human communicator or establish how social presence relates to interpersonal connectedness across different communication configurations. Differences between experimental, cross-sectional, longitudinal, and conceptual investigations also limit the direct comparability of reported outcomes. Findings derived from general human–AI interaction cannot automatically be generalized to university relationships, particularly when studies do not directly measure relational closeness, reciprocity, or sustained interpersonal engagement. Existing theoretical perspectives consequently require an integrated examination of the mechanisms through which AI assistance interacts with human agency, perceived authenticity, and contextual expectations.

This study addresses these limitations through a scoping review of scholarly research examining AI-mediated communication, trust, social presence, and interpersonal connectedness, with particular attention to university-related communication. The review investigates two principal questions: how the available literature characterizes the relationships between AI-mediated communication and the three relational constructs, and which technological, individual, and contextual conditions are associated with variation in these relationships. The synthesis organizes the evidence into three

analytical domains: trust calibration and perceived authenticity, social presence and conversational agency, and interpersonal connectedness in educational communication. It distinguishes evidence obtained directly from university populations from broader human–AI communication research and considers the methodological differences that constrain causal interpretation. The study's conceptual contribution lies in integrating these domains through a relational mechanism framework that differentiates technological reliability, perceived social responsiveness, and the maintenance of human interpersonal agency. Calibrated trust is examined as an interpretive principle concerning the relationship between appropriate technological reliance, critical evaluation, and human responsibility rather than as a uniformly established empirical outcome. This analytical orientation provides a basis for examining the relational implications of AI-assisted communication without assuming that technological efficiency, social presence, and interpersonal connectedness necessarily develop in the same direction.

RESEARCH METHODS

This study employed a scoping review design combined with a framework-based relational mechanism synthesis to examine the relationships among artificial intelligence (AI)-mediated communication, trust, social presence, and interpersonal connectedness. The review addressed two research questions: how existing scholarship characterizes the relationships between AI-mediated communication and these relational constructs, and which technological, individual, and contextual conditions are associated with variation in the reported relationships. Relational mechanism synthesis was operationalized as an author-defined analytical approach using thematic framework coding and cross-study comparison to identify the processes through which AI participation may facilitate, constrain, or modify relational outcomes. The approach was intended to map the conceptual and empirical characteristics of a heterogeneous evidence base rather than estimate pooled intervention effects or establish universal causal relationships.

Relevant literature was identified through bibliographic and publisher-based searches using combinations of the terms "artificial intelligence-mediated communication," "AI-mediated communication," "generative AI," "human–AI communication," "trust," "social presence," "interpersonal connectedness," "interpersonal relationships," "conversational agents," "chatbots," "authenticity," "disclosure," "AI literacy," and "university students." The analytical corpus comprised 34 scholarly sources published between 1976 and 2026, including foundational theoretical contributions and contemporary empirical investigations. Source eligibility was determined by substantive relevance to AI participation in communication, technological or interpersonal trust, perceived social presence, relational authenticity, interpersonal connectedness, or educational technology adoption. Empirical experiments, cross-sectional surveys, longitudinal investigations, mixed-method studies, conversational-agent design research, methodological taxonomies, and conceptual syntheses were included when they contributed evidence or theoretical explanations concerning these constructs. Technical investigations without a substantive communicative or relational component fell outside the analytical scope. The retained publications comprised nine controlled or online experimental studies, ten cross-sectional survey or structural equation modeling studies, four longitudinal or mixed-method investigations, five design or taxonomy studies, and six reviews or conceptual syntheses. Earlier publications were incorporated to establish the theoretical foundations of computer-mediated communication, social presence, technological trust, and human responses to computational agents, while recent publications were used to examine the emergence of generative AI applications and contemporary educational communication practices.

Data extraction and evidence charting focused on the research design, communication configuration, degree of AI assistance, disclosure conditions, conversational and anthropomorphic cues, trust-related outcomes, social-presence indicators, interpersonal connectedness, and reported contextual moderators. Three communication configurations were distinguished: AI-mediated communication as the broader category of algorithmically supported interaction, AI-assisted human-to-human communication involving AI participation in messages exchanged between individuals, and direct human–AI communication involving conversational systems as interaction partners. Trust-related evidence was further differentiated according to whether the object of evaluation was the technological system or the human communicator. Moderator categories included AI literacy, perceived accuracy, risk perception, privacy concerns, transparency, prior experience, and user control. Evidence involving

university students was distinguished analytically from findings obtained from general user populations, professional settings, and conceptual research to limit unsupported population-specific generalization.

The synthesis followed a deductive thematic framework organized around three relational domains: trust calibration and perceived authenticity, social presence and conversational agency, and interpersonal connectedness in educational communication. Within each domain, the evidence was compared according to communication configuration, methodological design, reported relational outcomes, and conditions affecting the interpretation of findings. Experimental findings were considered in relation to the specific manipulations and outcomes examined, whereas cross-sectional associations were interpreted without causal attribution. Longitudinal studies informed the examination of relational development over repeated interactions, while design studies and conceptual publications provided theoretical explanations of conversational agency, anthropomorphism, and technological accountability. Conflicting findings were interpreted in relation to differences in study populations, interaction settings, disclosure conditions, and the operationalization of relational constructs. The synthesis was qualitative and comparative; no statistical meta-analysis or pooled effect-size estimation was conducted.

The methodological scope was defined by the 34 retained publications and their documented analytical characteristics. The available review record established the final corpus and its research-design distribution but did not provide complete database-specific retrieval histories, original search dates, duplicate-removal totals, intermediate screening counts, or reviewer-agreement records. A retrospective numerical PRISMA-ScR selection flow was consequently not constructed from unsupported estimates. No formal risk-of-bias scores were reported, and the interpretation of evidence was qualified by research design and population. These boundaries restrict claims of search exhaustiveness and quantitative representativeness but permit a structured synthesis of the relational mechanisms, theoretical distinctions, and contextual conditions represented within the included literature.

RESULTS AND DISCUSSION

Trust calibration and perceived authenticity

The evidence on baseline trust showed a patterned rather than uniform relationship with the central outcome. AI assistance improved linguistic fluency and response efficiency in several settings, but recipients did not evaluate polished language independently of perceived authorship and authenticity. The pattern is consistent with the relational and institutional literature (Choung et al., 2023; Jacovi et al., 2021). Across the assembled sources, accuracy expectations emerged as a recurrent mechanism linking individual interaction to broader social outcomes. Trust judgments involved at least two objects: confidence in the technical system and confidence in the human communicator using it, which explains why the same intervention can raise one form of trust while lowering another. The pattern is consistent with the relational and institutional literature (Parasuraman & Riley, 1997; McKnight et al., 2011).

Analysis of disclosure indicated that the relational effect depended on how interaction was organized and interpreted by participants. Social presence increased when conversational systems displayed contingent, coherent, and socially legible responses, although human-like cues were not sufficient to guarantee durable relational confidence. The pattern is consistent with the relational and institutional literature (Reeves & Nass, 1996; Nowak & Biocca, 2003). The coded literature treated authenticity as a process variable rather than a simple exposure measure. Disclosure created a recurring tension because users often endorsed transparency as a norm while reacting less favorably to messages explicitly labeled as AI-assisted. The pattern is consistent with the relational and institutional literature (Clark et al., 2019; Formosa et al., 2026).

Table 1. Evidence Base by Research Design

Design	Number of sources	Typical contribution
Controlled/online experiments	9	Causal evidence on disclosure, trust, and social cues

Cross-sectional surveys/SEM	10	Associations among literacy, trust, risk, and adoption
Longitudinal or mixed-method studies	4	Change in trust and relationship formation
Design/taxonomy studies	5	Conversational cues and system expectations
Reviews/conceptual syntheses	6	Theory building and ethical framing

Findings concerning algorithm transparency were most consistent when studies measured reciprocity, tie quality, or institutional conditions directly. AI literacy supported more differentiated evaluation of system capability and reduced the appeal of treating either enthusiasm or skepticism as a stable personal trait. The pattern is consistent with the relational and institutional literature (Liu et al., 2025; Almaiah et al., 2025). A closer reading of risk perception clarified why apparently contradictory findings can coexist across samples and settings. Risk perception, privacy concerns, and perceived control moderated adoption because students assessed not only usefulness but also the consequences of depending on a system that can make confident errors. The pattern is consistent with the relational and institutional literature (Hancock et al., 2020; Hohenstein & Jung, 2020).

The strongest contribution of the evidence on AI literacy was its specification of boundary conditions around the expected social effect. Relational effects were stronger when people retained meaningful control over drafting, verification, and final message ownership. The pattern is consistent with the relational and institutional literature (Walther, 1992; Short et al., 1976). Research addressing attribution also linked micro-level interaction to meso-level networks, organizations, or communities. Conversational agency was most credible when the system role was clear and limited rather than when design attempted to erase the distinction between human and machine contribution. The pattern is consistent with the relational and institutional literature (Biocca et al., 2003; Nass & Moon, 2000).

Taken as a body of evidence, calibrated reliance helped distinguish durable relational resources from short-lived indicators of exposure or contact. The synthesis favors calibrated trust as a relational objective: enough confidence to use assistance productively, combined with enough skepticism to preserve verification and human accountability. The pattern is consistent with the relational and institutional literature (Sundar, 2008; Glikson & Woolley, 2020).

Social presence and conversational agency

The evidence on social cues showed a patterned rather than uniform relationship with the central outcome. AI assistance improved linguistic fluency and response efficiency in several settings, but recipients did not evaluate polished language independently of perceived authorship and authenticity. The pattern is consistent with the relational and institutional literature (Lee, 2018; Guzman & Lewis, 2020). Across the assembled sources, anthropomorphism emerged as a recurrent mechanism linking individual interaction to broader social outcomes. Trust judgments involved at least two objects: confidence in the technical system and confidence in the human communicator using it, which explains why the same intervention can raise one form of trust while lowering another. The pattern is consistent with the relational and institutional literature (Shin, 2021; Luger & Sellen, 2016).

Analysis of responsiveness indicated that the relational effect depended on how interaction was organized and interpreted by participants. Social presence increased when conversational systems displayed contingent, coherent, and socially legible responses, although human-like cues were not sufficient to guarantee durable relational confidence. The pattern is consistent with the relational and institutional literature (Araujo, 2018; Go & Sundar, 2019). The coded literature treated conversational coherence as a process variable rather than a simple exposure measure. Disclosure created a recurring tension because users often endorsed transparency as a norm while reacting less favorably to messages explicitly labeled as AI-assisted. The pattern is consistent with the relational and institutional literature (Adam et al., 2021; Croes & Antheunis, 2021).

Table 2. AI-Mediated Communication Mechanisms

Mechanism	Relational gain	Relational risk
Language assistance	Clarity and response efficiency	Reduced authenticity when assistance dominates
Transparency signaling	Better understanding of system role	Disclosure penalty in some settings
Social cues	Higher immediacy and social presence	Over-attribution of human qualities
Personalization	Relevance and responsiveness	Privacy and manipulation concerns
Human oversight	Accountability and calibration	Added effort and slower interaction

Findings concerning agency framing were most consistent when studies measured reciprocity, tie quality, or institutional conditions directly. AI literacy supported more differentiated evaluation of system capability and reduced the appeal of treating either enthusiasm or skepticism as a stable personal trait. The pattern is consistent with the relational and institutional literature (Feine et al., 2019; de Visser et al., 2016). A closer reading of telepresence clarified why apparently contradictory findings can coexist across samples and settings. Risk perception, privacy concerns, and perceived control moderated adoption because students assessed not only usefulness but also the consequences of depending on a system that can make confident errors. The pattern is consistent with the relational and institutional literature (Formosa et al., 2025; Wang et al., 2025).

The strongest contribution of the evidence on human-likeness was its specification of boundary conditions around the expected social effect. Relational effects were stronger when people retained meaningful control over drafting, verification, and final message ownership. The pattern is consistent with the relational and institutional literature (Petricini et al., 2025; Obadă et al., 2026).

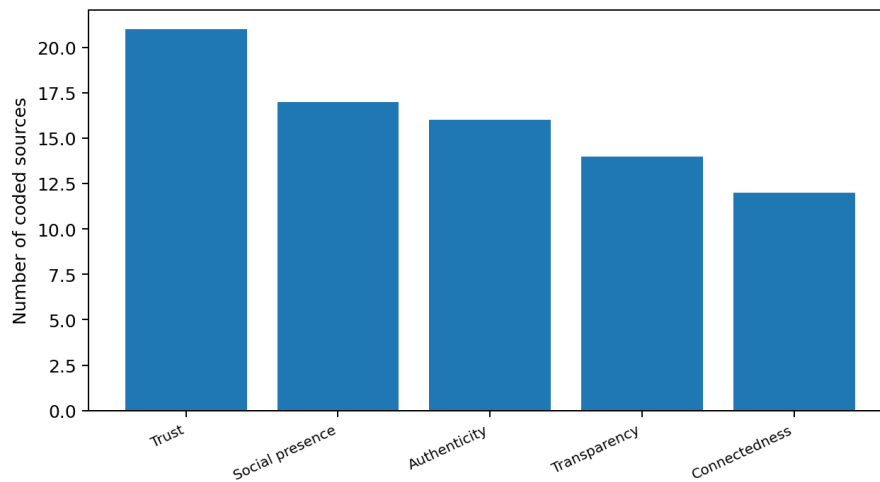


Figure 1. Coded Evidence Profile for Artificial Intelligence-Mediated Communication and Human Relationships

Research addressing expectation gaps also linked micro-level interaction to meso-level networks, organizations, or communities. Conversational agency was most credible when the system role was clear and limited rather than when design attempted to erase the distinction between human and machine contribution. The pattern is consistent with the relational and institutional literature (Zhang & Yu, 2025; Shin et al., 2024).

Taken as a body of evidence, relational satisfaction helped distinguish durable relational resources from short-lived indicators of exposure or contact. The synthesis favors calibrated trust as a relational objective: enough confidence to use assistance productively, combined with enough

skepticism to preserve verification and human accountability. The pattern is consistent with the relational and institutional literature (Choung et al., 2023; Jacovi et al., 2021).

Interpersonal connectedness in university communication

The evidence on student adoption showed a patterned rather than uniform relationship with the central outcome. AI assistance improved linguistic fluency and response efficiency in several settings, but recipients did not evaluate polished language independently of perceived authorship and authenticity. The pattern is consistent with the relational and institutional literature (Parasuraman & Riley, 1997; McKnight et al., 2011). Across the assembled sources, peer exchange emerged as a recurrent mechanism linking individual interaction to broader social outcomes. Trust judgments involved at least two objects: confidence in the technical system and confidence in the human communicator using it, which explains why the same intervention can raise one form of trust while lowering another. The pattern is consistent with the relational and institutional literature (Reeves & Nass, 1996; Nowak & Biocca, 2003).

Analysis of academic honesty indicated that the relational effect depended on how interaction was organized and interpreted by participants. Social presence increased when conversational systems displayed contingent, coherent, and socially legible responses, although human-like cues were not sufficient to guarantee durable relational confidence. The pattern is consistent with the relational and institutional literature (Clark et al., 2019; Formosa et al., 2026). The coded literature treated language learning as a process variable rather than a simple exposure measure. Disclosure created a recurring tension because users often endorsed transparency as a norm while reacting less favorably to messages explicitly labeled as AI-assisted. The pattern is consistent with the relational and institutional literature (Liu et al., 2025; Almaiah et al., 2025).

Table 3. Conditions for Calibrated Trust among Students

Condition	Trust implication	Design implication
AI literacy	Improves evaluation of strengths and limits	Teach verification and provenance skills
Perceived accuracy	Raises initial reliance	Pair confidence with error awareness
Risk perception	Reduces uncritical acceptance	Make uncertainty visible
Disclosure norms	Supports accountability	Use clear, proportionate disclosure
User control	Preserves human agency	Allow revision, rejection, and opt-out

Findings concerning efficiency were most consistent when studies measured reciprocity, tie quality, or institutional conditions directly. AI literacy supported more differentiated evaluation of system capability and reduced the appeal of treating either enthusiasm or skepticism as a stable personal trait. The pattern is consistent with the relational and institutional literature (Hancock et al., 2020; Hohenstein & Jung, 2020). A closer reading of relationship maintenance clarified why apparently contradictory findings can coexist across samples and settings. Risk perception, privacy concerns, and perceived control moderated adoption because students assessed not only usefulness but also the consequences of depending on a system that can make confident errors. The pattern is consistent with the relational and institutional literature (Walther, 1992; Short et al., 1976).

The strongest contribution of the evidence on privacy was its specification of boundary conditions around the expected social effect. Relational effects were stronger when people retained meaningful control over drafting, verification, and final message ownership. The pattern is consistent with the relational and institutional literature (Biocca et al., 2003; Nass & Moon, 2000). Research addressing accountability also linked micro-level interaction to meso-level networks, organizations, or communities. Conversational agency was most credible when the system role was clear and limited rather than when design attempted to erase the distinction between human and machine contribution. The pattern is consistent with the relational and institutional literature (Sundar, 2008; Glikson & Woolley, 2020).

Taken as a body of evidence, institutional design helped distinguish durable relational resources from short-lived indicators of exposure or contact. The synthesis favors calibrated trust as a relational

objective: enough confidence to use assistance productively, combined with enough skepticism to preserve verification and human accountability. The pattern is consistent with the relational and institutional literature (Lee, 2018; Guzman & Lewis, 2020).

CONCLUSION

The scoping review identified a differentiated relationship between artificial intelligence-mediated communication and the relational constructs of trust, social presence, and interpersonal connectedness. Across the 34 reviewed sources, AI assistance was associated with opportunities for greater communicative efficiency, linguistic clarity, and perceived responsiveness, although these characteristics did not consistently correspond to stronger interpersonal relationships. The evidence distinguished trust in technological competence from trust in the human communicator, indicating that confidence in AI-generated content and perceptions of communicative authenticity represent separate evaluative processes. Conversational coherence, anthropomorphic cues, and contextual responsiveness contributed to explanations of perceived social presence, but the available findings did not establish that these characteristics consistently produce sustained relational attachment. Evidence specifically involving university students emphasized AI literacy, perceived usefulness, technological trust, risk perception, and educational expectations, whereas broader claims concerning interpersonal connectedness were informed partly by research conducted in other populations and interaction settings. The principal conceptual contribution of this review is a relational mechanism framework that distinguishes technological reliability, perceived social responsiveness, and interpersonal connectedness while identifying authenticity, transparency, and human agency as conditions shaping their relationships. Calibrated trust emerges from this synthesis as an interpretive principle emphasizing proportionate technological reliance, critical evaluation, and continued human responsibility rather than as an independently verified universal outcome. For educational institutions, the findings support communication practices that preserve user control, enable verification, and apply context-sensitive and proportionate disclosure without assuming that transparency invariably strengthens perceived authenticity. The review is limited by heterogeneous research designs, variation in relational outcome measures, and the restricted availability of longitudinal evidence directly examining university relationships. Future research should investigate repeated AI-mediated interactions, distinguish system trust from interpersonal trust, and examine how disclosure practices, technological literacy, and communication configurations influence relational development over time. The relational implications of AI-mediated communication are neither uniformly beneficial nor uniformly detrimental; they depend on the interaction between technological assistance, communicative expectations, and the continued exercise of human agency.

REFERENCES

- Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, 31(2), 427–445. <https://doi.org/10.1007/s12525-020-00414-7>
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, 85, 183–189. <https://doi.org/10.1016/j.chb.2018.03.051>
- Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators and Virtual Environments*, 12(5), 456–480. <https://doi.org/10.1162/105474603322761270>
- Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human-Computer Interaction*, 39(9), 1727–1739. <https://doi.org/10.1080/10447318.2022.2050543>
- Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., Wade, V., & Cowan, B. R. (2019). What makes a good conversation? Challenges in designing truly conversational agents. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Association for Computing Machinery. <https://doi.org/10.1145/3290605.3300705>

- Croes, E. A. J., & Antheunis, M. L. (2021). Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships*, 38(1), 279–300. <https://doi.org/10.1177/0265407520959463>
- de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, 22(3), 331–349. <https://doi.org/10.1037/xap0000092>
- Đerić, E., Frank, D., & Milković, M. (2025). Trust in generative AI tools: A comparative study of higher education students, teachers, and researchers. *Information*, 16(7), Article 622. <https://doi.org/10.3390/info16070622>
- Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, 132, 138–161. <https://doi.org/10.1016/j.ijhcs.2019.07.009>
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660. <https://doi.org/10.5465/annals.2018.0057>
- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>
- Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society*, 22(1), 70–86. <https://doi.org/10.1177/1461444819858691>
- Hancock, J. T., Naaman, M., & Levy, K. (2020). AI-mediated communication: Definition, research agenda, and ethical considerations. *Journal of Computer-Mediated Communication*, 25(1), 89–100. <https://doi.org/10.1093/jcmc/zmz022>
- Hohenstein, J., & Jung, M. (2020). AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust. *Computers in Human Behavior*, 106, Article 106190. <https://doi.org/10.1016/j.chb.2019.106190>
- Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021). Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 624–635). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445923>
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5(1), 1–16. <https://doi.org/10.1177/2053951718756684>
- 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). Association for Computing Machinery. <https://doi.org/10.1145/2858036.2858288>
- McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems*, 2(2), Article 12, 1–25. <https://doi.org/10.1145/1985347.1985353>
- 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>
- Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users’ sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 12(5), 481–494. <https://doi.org/10.1162/105474603322761289>
- Obadă, D.-R., Gradinaru, C., & Gradinaru, I.-A. (2026). From understanding to influence: The interplay of AI literacy, self-efficacy, and trust in predicting communication students’ AI adoption and word-of-mouth. *Frontiers in Communication*, 10, Article 1722464. <https://doi.org/10.3389/fcomm.2025.1722464>
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253. <https://doi.org/10.1518/001872097778543886>

- Park, K., & Yoon, H. Y. (2024). Beyond the code: The impact of AI algorithm transparency signaling on user trust and relational satisfaction. *Public Relations Review*, 50(5), Article 102507. <https://doi.org/10.1016/j.pubrev.2024.102507>
- Petricini, T., Zipf, S., & Wu, C. (2025). RESEARCH-AI: Communicating academic honesty: Teacher messages and student perceptions about generative AI. *Frontiers in Communication*, 10, Article 1544430. <https://doi.org/10.3389/fcomm.2025.1544430>
- Purcell, Z. A., Jakesch, M., Dong, M., Nussberger, A.-M., & Köbis, N. (2025). Writing with AI boosts trust-building efficiency. *iScience*, 28(12), Article 114092. <https://doi.org/10.1016/j.isci.2025.114092>
- Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, televisions, and new media like real people and places*. CSLI Publications; Cambridge University Press.
- Sahebi, S., & Formosa, P. (2025). The AI-mediated communication dilemma: Epistemic trust, social media, and the challenge of generative artificial intelligence. *Synthese*, 205, Article 128. <https://doi.org/10.1007/s11229-025-04963-2>
- Sahebi, S., Formosa, P., & Bankins, S. (2026). The AI penalty and disclosure paradox: Trust, authenticity and knowledge uptake in AI-mediated communication. *Computers in Human Behavior: Artificial Humans*, 8, Article 100304. <https://doi.org/10.1016/j.chbah.2026.100304>
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, Article 102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Wiley.
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73–100). MIT Press. <https://doi.org/10.1162/dmal.9780262562324.073>
- Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, 19(1), 52–90. <https://doi.org/10.1177/009365092019001003>
- Wang, F., Li, N., Cheung, A. C. K., & Wong, G. K. W. (2025). In GenAI we trust: An investigation of university students' reliance on and resistance to generative AI in language learning. *International Journal of Educational Technology in Higher Education*, 22(1), Article 59. <https://doi.org/10.1186/s41239-025-00547-9>
- Zhang, G., & Yu, T. (2025). Association between generative AI self-efficacy and generative AI acceptance: The mediating role of generative AI trust and the moderating role of generative AI risk perception. *Acta Psychologica*, 261, Article 105791. <https://doi.org/10.1016/j.actpsy.2025.105791>