Human-AI Harmony from the Perspective of Media Naturalness

Authors: Su Di, Liao Jiangqun, Peng Kaiping, Jiangqun Liao

Date: 2025-08-29T20:59:05+00:00

Abstract

With the widespread application of Artificial Intelligence (AI), establishing high-quality human-AI interaction relationships has emerged as a prominent research focus. This paper conducts an in-depth analysis of the potential for harmonious symbiosis between humans and AI, introducing the concept of "Human-AI Rapport" (HAR) along with its constituent components—relational harmony, mutual understanding, and tacit coordination. By integrating this concept with Media Naturalness Theory (MINT), the paper innovatively proposes a Human-AI Rapport Promotion Model from the perspective of constructing naturalness in human-AI interaction. This model not only profoundly elucidates the critical roles and implementation pathways of cognitive, social, and emotional intelligence in design, thereby providing a novel theoretical perspective for human-AI interaction research, but also identifies critical issues requiring urgent in-depth investigation and resolution. It advocates for the full exploitation of AI's potential, expanding its application scenarios and value boundaries to foster mutual adaptation between human society and AI technology, ultimately realizing the vision of harmonious human-AI symbiosis.

Full Text

Human-AI Rapport from the Perspective of Media Naturalness

Su Di[1,2], **Liao Jiangqun**[3,1], **Peng Kaiping**[2]
[1] Department of Psychology and Cognitive Science, Tsinghua University, Beijing, 100084
[2] Student Mental Health Counseling Center, Ningxia University, Yinchuan, 750021
[3] Business School, Beijing Technology and Business University, Beijing, 100048

Abstract

As Artificial Intelligence (AI) becomes increasingly pervasive, establishing high-quality human-AI interactive relationships has emerged as a critical research focus. This paper provides an in-depth analysis of the possibility for harmonious coexistence between humans and AI, introducing the concept of "Human-AI Rapport" (HAR) and its three constituent components: harmonious relationship, mutual understanding, and tacit cooperation. By integrating this concept with Media Naturalness Theory (MNT), this study innovatively proposes a Human-AI Rapport promotion model from the perspective of constructing naturalness in human-AI interaction. This model not only reveals the key roles and implementation pathways of cognitive, social, and emotional intelligences in AI design, providing a novel theoretical perspective for human-AI interaction research, but also identifies pressing issues requiring future investigation. The paper encourages the full realization of AI's potential and the expansion of its application scenarios and value boundaries to promote mutual adaptation between human society and AI technology, ultimately achieving the vision of harmonious human-AI symbiosis.

Keywords: Artificial Intelligence, Human-AI Rapport, Media Naturalness

Within the theoretical framework of Marxism, productive tools serve as important indicators of productivity development levels, and their evolution directly drives transformations in social production relations. Each leap forward in productive tools is accompanied by tremendous gains in productivity, thereby triggering adjustments in production relations and profound changes in social structure. Today, AI has emerged as a new generation of productive tools, permeating every aspect of human production and life. From smart homes to medical diagnosis, from educational tutoring to customer service, AI is gradually shedding its role as a mere tool and instead participating in social production as a means of assistance, collaboration, and human capability enhancement (Kim et al., 2022; G. Park et al., 2023; Pasternak et al., 2022; Steele et al., 2022), becoming a crucial driving force in the development of new quality productive forces (冯道杰, 2024). Against this important backdrop, harmonious symbiosis between humans and AI is essential, and several core questions have become increasingly prominent: What kind of relationship should humans and AI have for dialogue and cooperation? How can high-quality interactive relationships be established between humans and AI? This inquiry not only concerns user experience and acceptance of AI technology but also directly impacts the practical effectiveness and future prospects of AI applications (Pentina, Xie, et al., 2023).

Previous research has consistently emphasized the importance of rapport in human-AI interaction and advocated for promoting such rapport as a core orientation in AI design (Lee et al., 2024; Mai et al., 2021; Nichols et al., 2022; G. Park et al., 2023). However, these discussions have largely remained within the traditional frameworks of interpersonal rapport or professional rapport (Tickle-Degnen & Rosenthal, 1990; Gabbert et al., 2021), failing to construct a unique

conceptual system or theoretical model specifically for rapport between humans and AI. Clearly, simply applying interpersonal rapport theories to human-AI relationships has limitations and is not entirely appropriate. In light of this, this study innovatively proposes the concept of "Human-AI Rapport" (HAR) to describe the quality of the relationship between humans and AI when users accept AI services or engage in professional interactions with AI. This concept aims to provide a more solid and targeted theoretical foundation for future research, thereby advancing the field to deeper levels.

Current research on HAR involves multiple disciplines including psychology and computer science. Psychology, as a science that deeply analyzes the cognition, behavior, emotions, and affect of interacting parties, provides the theoretical foundation required for data collection, annotation, and model construction in HAR research. Computer science, meanwhile, provides the technical means (devices, algorithms, programs) to achieve comprehensive collection, precise computation, and innovative generation of multimodal data. The two complement each other and jointly promote the development of HAR theory and practice. Against this backdrop, this study aims to explore the connotation and establishment strategies of HAR, hoping to provide useful references and insights for the human-centered design of AI technology, optimization of user experience, and harmonious construction of human-AI relationships.

2.1 Interpersonal Rapport and Professional Rapport

In interpersonal interactions, relationships characterized by harmony, smoothness, and synchrony are referred to as rapport, which represents good relationship quality and tacit understanding, and is one of the typical features of many successful interactions (Gratch et al., 2007). Tickle-Degnen and Rosenthal (1990) propose that rapport comprises three components: positivity, which refers to the friendliness and warmth between both parties; attention, which refers to mutual interest and concern; and coordination, which refers to the perception of balance and harmony during interaction. From a third-party observation perspective, rapport also has three corresponding manifestations: head, facial, and vocal expressions (such as nodding, smiling, and pitch variation) reflect positivity between the parties; static body movements represent attention; and dynamic body movements (such as adjusting posture to adapt to the other) enhance coordination (Hendrick, 1990). Past research has demonstrated that rapport in interpersonal interactions can produce positive effects in fields such as negotiation, management, psychological counseling and therapy, teaching, and nursing (Gabbert et al., 2021; Gratch et al., 2007; Kim et al., 2022).

In contrast to general social contexts, after systematically reviewing research on the use and measurement of rapport in professional information-gathering contexts, Gabbert et al. (2021) proposed the concept of professional rapport-building, which refers to the intentional use of rapport-facilitating behaviors in professional settings (such as negotiations and investigative interviews) to promote positive interactions with work targets. Therefore, while rapport dis-

cussed in social domains represents an interactive state to be achieved in ideal and high-quality interpersonal interactions, in time-constrained professional domains, rapport is a task-oriented practical strategy and means used by interviewers/consultants/interviewers to establish connections, facilitate dialogue, and achieve goals with work targets, though it does not necessarily establish genuine and long-term rapport (Brouillard et al., 2024; Gabbert et al., 2021; Neequaye, 2023). Functionally, there are three main approaches to building rapport: (1) customizing interviews through self-disclosure and discussing personal matters; (2) demonstrating approachability through smiling, open body language, and appropriate tone; and (3) showing positive regard for interviewees through active listening, eye contact, understanding, and affirmation (Gabbert et al., 2021).

2.2 Concept and Components of Human-AI Rapport

Similar to the scenarios where professionals conduct work mentioned above, AI functions have also evolved from simply executing commands to more intelligent applications such as collecting user information, analyzing user intentions, achieving understanding and empathy, and providing services and companionship (侯悍超等, 2024; Kim et al., 2022; Ranjbartabar et al., 2021). Therefore, it is necessary to discuss whether the aforementioned rapport and building strategies are equally applicable to human-AI interactions in professional contexts.

Although embodied AI agents can mimic users' nonverbal behaviors to "fabricate" positivity, attention, and coordination in human-AI interaction (Huang et al., 2011; Lubold et al., 2021), they cannot achieve physiological synchrony and neural synchrony (McNaughton & Redcay, 2020). Excessive attention from machines may also cause users to feel their safety and privacy are threatened, leading to panic and aversion (Koller et al., 2023). Moreover, a large number of non-embodied AI agents are also helping professional workers or users achieve task goals (Chattaraman et al., 2019; G. Park et al., 2023), but they cannot use behavioral expression strategies (such as changing body posture and facial expressions) to build rapport with humans. Therefore, traditional rapport concepts and frameworks may be poorly adapted to human-AI interaction.

This paper combines the concept of traditional interpersonal rapport with the particularities of human-AI interaction mentioned above (Diederich et al., 2022) to propose the concept of "Human-AI Rapport," which refers to the degree to which users experience harmonious relationship, mutual understanding, and tacit cooperation with AI in the process of using AI to achieve various professional work goals. It represents users' subjective perception of the depth and quality of the interactive relationship. The three components in this concept are related to yet distinct from the three dimensions of traditional interpersonal rapport (see Table 1): harmonious relationship is built upon "positivity" but emphasizes users' good experience of the overall cooperative atmosphere (an experience that does not merely come from AI's consistently positive feedback). Mutual understanding is built upon "attention," requiring AI's attention to be

reflected not only in collecting user behavioral data but also in understanding the intentions behind user behaviors and expressing this understanding in ways acceptable to users. Tacit cooperation is built upon "coordination," specifically referring to the ability to tacitly maintain synchrony with users while achieving complementarity with them during cooperation (such as personalized recommendations and memory assistance) without requiring user requests or emphasis.

Table 1 Components of Human-AI Rapport

Component	Theoretical Basis
Harmonious Relationship	Users' good experience of the overall cooperative atmosphere
Mutual Understanding	Understanding intentions behind user behaviors and expressing this understanding
Tacit Cooperation	Maintaining synchrony with users while achieving complementarity during cooperation

A review of the literature reveals that concepts similar to HAR include AI usage, trust in AI, attitudes toward AI, and relationship with AI. To clearly delineate the differences between HAR and these related concepts and highlight its unique theoretical value, it is necessary to provide a detailed and in-depth elaboration (see Table 2). First, the connotations and evaluation contents of each concept differ. AI usage refers to individuals' behaviors or willingness to accept and adopt AI (王涛等, 2024; B. Li et al., 2023). Attitudes toward AI refer to individuals' psychological tendencies to approve of or resist AI being used in work, life, and decision-making (De Freitas et al., 2023; J. Park et al., 2024). Relationship with AI refers to the degree of affection and emotional dependence that gradually develops between users and AI as contact time and AI roles change (Pentina, Hancock, et al., 2023; Tschopp et al., 2023). Trust in AI refers to users' positive expectations regarding interactions with AI when facing AI uncertainty and unpredictability, including both evaluations of AI functional reliability and identification with the values and ethics behind AI algorithms, as well as corresponding psychological security (Asan et al., 2020; Choung et al., 2023). Second, the concepts differ in their degree of generalization when evaluating AI and its application scenarios. HAR focuses on evaluations of users' interactions with specific AI agents in professional contexts, with clear directionality; other concepts are not limited to evaluations of specific AI agents or interaction scenarios and may generalize to evaluations of broad AI technology and application scenarios. Such evaluations derive from both actual usage experience and external information such as others' experiences, viewpoints, and news reports. Finally, HAR differs from other concepts in the dimension of evaluation objects. Other concepts typically direct evaluation from humans to AI, focusing on unilateral considerations of AI; HAR emphasizes evaluating the bilateral relationship during users' professional interaction processes, meaning

its evaluation content is rooted in factors elicited by the interaction itself rather than focusing solely on AI's characteristics.

Table 2 Distinctions Between Human-AI Rapport and Related Concepts

Concept	Content Described	Degree of Generalization	Evaluation Dimension
Human-AI Rapport	Understanding, harmony, tacit cooperation	Low: evaluation of specific AI and professional interaction scenarios only	Bilateral relationship evaluation
AI Usage	Acceptance and adoption behaviors	High: not limited to specific AI and scenarios (may generalize to broad AI technology and application scenarios)	Unilateral AI evaluation
Attitudes Toward AI	Degree of support or resistance	High: not limited to specific AI and scenarios (may generalize to broad AI technology and application scenarios)	Unilateral AI evaluation
Relationship With AI	Affection and emotional dependence	High: not limited to specific AI and scenarios (may generalize to broad AI technology and application scenarios)	Unilateral AI evaluation
Trust in AI	Reliability, psychological security	High: not limited to specific AI and scenarios (may generalize to broad AI technology and application scenarios)	Unilateral AI evaluation

3.1 Application of Interpersonal Rapport Models in Human-AI Contexts

Previous research on human-AI relationships has largely assumed that theories related to human relationships can provide further insights into the development of human-AI relationships (Seymour & Van Kleek, 2021; Xie & Pentina, 2022), and scholars have widely applied Tickle-Degnen and Rosenthal's (1990) interpersonal rapport model (i.e., positivity, attention, and coordination) to research on building and predicting HAR (Gratch et al., 2007; Maxim et al., 2023; Pasternak et al., 2022). For example, Huang et al. (2011) developed Virtual Rapport

2.0 based on the Virtual Rapport 1.0 virtual agent (Gratch et al., 2006) using the interpersonal rapport model as a framework. They employed a data-driven approach to enhance the attentiveness and coordination of the interactive relationship and used emotional response and reciprocity enhancement techniques to strengthen the positivity of emotional communication. Results showed that compared to version 1.0, version 2.0 effectively improved users' perception of rapport (Huang et al., 2011).

3.2 Media Naturalness Theory

Kock's (2004) Media Naturalness Theory (MNT) posits that humans are biologically predisposed to communicate through voice, facial expressions, and body language. This natural communication includes five elements: (1) sharing the same environment where parties can see and hear each other; (2) obtaining real-time and synchronous feedback from partners; (3) being able to convey and observe facial expressions; (4) being able to convey and observe body language; and (5) being able to convey and hear voice. However, as technology develops, technology-mediated interaction has become increasingly common, but these interactions often fail to meet the above elements and inhibit many natural features of face-to-face communication. Consequently, naturalness in interaction—the degree to which interactive media achieves similarity to face-to-face communication through technology (Kock, 2004)—has become key to addressing this issue. It emphasizes creating a sense of synchronicity and collocation in interaction by simulating face-to-face voice, facial expressions, and body movements, making technology-mediated interaction as close as possible to natural face-to-face communication (Kock, 2004, 2005). MNT further proposes that decreasing cognitive effort, reducing communication ambiguities, and increasing physiological arousal are the three core mechanisms for enhancing interaction naturalness (Kock, 2004).

3.3 A New Human-AI Rapport Promotion Model

Although interpersonal rapport theoretical models have been widely applied in human-AI interaction research, inherent differences between the two determine that the direct applicability of interpersonal rapport-building strategies in human-AI interaction contexts is not absolute (Ter Stal et al., 2021). Moreover, previous research on HAR has been relatively fragmented and has not yet formed a systematic and complete theoretical model for exploring how to promote HAR. MNT provides effective and specific implementation pathways for establishing and promoting HAR—the stronger AI's naturalness, the more inclined humans are to view it as a synchronous and collocated communication partner, accompanied by positive perceptions and attitudes. Therefore, building harmonious and rapport relationships between humans and AI, and making AI be perceived as "human-like" by humans, hinges on AI design being closely centered on human experience and focusing on enhancing the naturalness of its interaction with humans. This study attempts to combine the newly proposed

concept and structure of HAR with the core ideas of MNT to propose a new HAR promotion model (see Figure 1

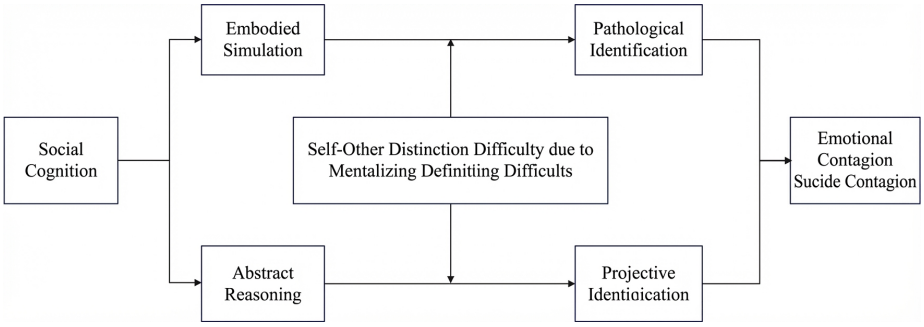

Figure 1: Figure 1

).

The model has two layers. The outer layer comprises design objectives for human-AI interaction, including HAR objectives (three vertices representing the three evaluation dimensions of HAR) and media naturalness objectives (lines connecting any two dimensions, representing that achieving these objectives can enhance corresponding rapport perceptions). The inner layer comprises pathways to achieve these objectives, including three types of AI intelligence: cognitive intelligence, social intelligence, and emotional intelligence. From inside to outside, through technological improvements, enhancements in AI's three types of intelligence will effectively achieve the three core interaction objectives in MNT, significantly improving AI's naturalness (Chandra et al., 2022), ensuring humans remain fluent and engaged when communicating with AI, thereby making the human-AI interaction process more closely resemble authentic interpersonal interaction experiences, and ultimately promoting HAR through these three pathways.

4 Strategies for Promoting Human-AI Rapport

Based on the above HAR promotion model, AI design aimed at enhancing naturalness is the fundamental strategy for improving rapport between users and AI. Researchers and designers need to explore how to effectively promote HAR from three perspectives: cognitive intelligence, social intelligence, and emotional intelligence:

4.1 Enhancing Cognitive Intelligence

The goals of mutual understanding and tacit cooperation between humans and AI place higher demands on AI's cognitive intelligence. Thanks to evolution, the human brain can easily recognize and process interpersonal communication

signals such as facial expressions, body movements, and voice without excessive cognitive effort. However, when communicating with low-naturalness AI, humans may need to expend more cognitive effort and learn new communication strategies to understand AI or make themselves understood by AI, which is not conducive to building rapport. Therefore, AI should not merely be a passive tool that answers user questions or provides computational results; it needs to demonstrate cognitive intelligence, particularly the ability to understand, summarize, and reason about user instructions and intentions, and to respond correctly, adaptively, and appropriately. This reduces users' cognitive burden during interaction with AI—that is, reducing the cognitive effort spent on repetition, explanation, and clarification (Hudon et al., 2021)—thus enabling users to perceive mutual understanding with AI and achieve synchronized and complementary tacit cooperation.

Today's deep learning technology has endowed AI systems with pattern recognition and feature extraction capabilities that demonstrate intelligent levels of cognitive intelligence, enabling them to perceive user needs and execute complex tasks in more adaptive ways, thereby enhancing user experience. For example, topic spotting technology proposed by researchers can help dialogue systems automatically infer conversation topics, making dialogue systems more engaging and efficient and helping maintain long-term conversations with users, which can facilitate mutual understanding and rapport building between humans and AI (Chitkara et al., 2019).

Memory and recall capabilities are also important cognitive abilities (Richards & Bransky, 2014). If AI can "remember" previous conversations with users and appropriately "recall" relevant "memories" during conversation, it can effectively promote their interaction (Kasap & Magnenat-Thalmann, 2012). This is not only a manifestation of intelligence but also allows the agent to better understand users through their background knowledge (Richards, 2017; Richards & Bransky, 2014). Moreover, in interaction, AI's recall of information must be natural and context-appropriate; otherwise, it will damage users' perception of tacit cooperation (Richards & Bransky, 2014). Conversely, if AI lacks memory and recall capabilities, humans must invest more cognitive resources in repetition and information prompting during interaction, creating a negative interactive experience of not being understood.

Nichols et al. (2022) established a dialogue generation strategy called Tiers of Friendship that enables agents to establish deeper conversation topics with users based on historical dialogues, thereby enhancing rapport with users (Nichols et al., 2022). Corresponding to memory is the question of how AI should "tacitly" demonstrate "forgetting" with users (Ellwart & Kluge, 2019), because humans may not want all information to be remembered. This places higher demands on agents' sensitivity and adaptivity (Richards & Bransky, 2014).

4.2 Exhibiting Social Intelligence

In low-naturalness human-AI interaction, the lack of nonverbal social cues often triggers communication ambiguity and negative user impressions, which hinder relationship establishment and create obstacles for tacit cooperation (Feine et al., 2019; I. Park et al., 2023). Similar to strategies for building rapport in interpersonal interactions (Cheung et al., 2015; Kulesza et al., 2016; Lin & Lin, 2017), researchers have endowed AI with nonverbal communication capabilities through animated figures, mechanical devices, and voice output, enabling users to perceive naturalness in communication and cooperation through AI's posture (Riek et al., 2010), gestures (Cade et al., 2011; Lubold et al., 2021), facial expressions (Pasternak et al., 2022), and vocal tone (Lubold et al., 2021), thereby establishing good HAR.

Beyond social cues, providing identity cues is also a social strategy that helps individuals build good relationships with others and increase social influence (Taylor et al., 2023). For AI, merely mimicking nonverbal behaviors is insufficient to provide enough social intelligence to enhance its naturalness. AI needs to more proactively provide identity cues and background information for interaction to users, thereby reducing communication ambiguity and increasing its social presence in users' eyes (Chen et al., 2023; Go & Sundar, 2019), which triggers users to respond to technology in the same way they treat humans (Westerman et al., 2020). For example, researchers made social improvements to robots assisting convenience store employees, such as proactively greeting employees through text or voice, introducing themselves, and explaining role assignments, which caused employees to perceive the robot as a friendly social other, effectively promoting task cooperation, improving employees' perception of the work atmosphere, and forming HAR (Takahashi et al., 2022).

Similarly, AI self-disclosure can make users more willing to open up (Neef et al., 2022), and this strategy has been proven to promote coordination and positive rapport between both parties and improve AI's effectiveness in completing corresponding tasks in health, teaching, and negotiation AI (Mai et al., 2021; Neef et al., 2022; Zhao et al., 2018). Additionally, AI addressing users by name and engaging in small talk has also been proven as a manifestation of social intelligence that can build rapport with users (Lee et al., 2024).

4.3 Optimizing Emotional Intelligence

Elements unique to human face-to-face communication (such as facial expressions) contribute to physiological arousal, bring positive emotional experiences, and promote mutual understanding (Krumhuber et al., 2013). When low-naturalness AI interacts with humans, these elements are easily suppressed, resulting in lower perceived excitability and loss of many emotional experiences beneficial to interaction (Kock, 2004, 2005). AI with emotional intelligence will enhance users' physiological arousal, thereby increasing naturalness, allowing users to experience that AI's emotional expressions come not from machines'

uniform, undifferentiated reactions but from perception and understanding of human emotions, and responding in ways that express understanding and empathy, thereby promoting rapport establishment (Tickle-Degnen & Rosenthal, 1990).

Therefore, emotional intelligence is an indispensable component in building HAR, representing AI's comprehensive ability to perceive, use, understand, and manage emotions (侯悍超等, 2024). First, AI captures subtle emotional changes from users' facial expressions and even physiological signals, which is the foundation for understanding user needs and predicting user reactions. For example, by integrating computer expression recognition toolboxes, AI can analyze users' expressions and emotional states in real time and adjust its response strategies accordingly to communicate with users in more appropriate ways (Cerekovic et al., 2017). Second, although humans often express emotions through nonverbal behaviors (facial expressions, posture, tone, etc.), AI that interacts with users through language can also develop the ability to identify, analyze, and interpret users' emotional states from conversational context using deep learning, natural language processing, and machine learning technologies (Anzum & Gavrilova, 2023; Chen et al., 2024; D & Juliet, 2023).

On the basis of recognition and understanding, responding to users with empathy is an effective means to enhance coordination and attention in human-AI interaction, and related technologies for empathetic dialogue generation have gradually matured, providing guarantees for improving human-AI interaction experience and relationship quality (侯悍超等, 2024). When AI makes responsive movements or expressions, or demonstrates resonance with users' emotions through dialogue, users feel understood and supported, thereby establishing stronger emotional bonds and rapport (Abdulrahman et al., 2021; Chen et al., 2024; Namkoong et al., 2024; Ranjbartabar et al., 2021).

5 Discussion and Outlook

This paper innovatively proposes the concept of HAR, explores its connotation in depth by integrating literature from the human-AI relationship field, and specifically discusses how to effectively establish HAR within the integrated framework of MNT and HAR, further enriching the theoretical framework of this field, providing a new perspective for human-AI relationship research, and offering suggestions for future AI design and optimization. Future research should conduct deeper investigations into the following issues in this field.

5.1 Measurement of Human-AI Rapport

Previous research has used different questionnaires to measure users' perception of HAR. The main measurement approaches fall into three categories: (1) measurements based on the three-factor model proposed by Tickle-Degnen and Rosenthal (1990) as content and dimensions (Lubold et al., 2018; Steele et al., 2022). (2) Measurements that integrate indicators of attitudes, relationship, and

trust as previously described. The most common is the rapport scale developed by Gratch et al. (2007), consisting of 10 items that evaluate rapport between users and AI entities from 10 perspectives: sense of connection, sense of understanding, helpfulness, credibility, likability, naturalness, favorability, anthropomorphism, persuasiveness, and recommendability, which has been adapted and adopted by many researchers (Acosta & Ward, 2011; Lubold et al., 2021; Ter Stal et al., 2021). (3) The use of other concepts and corresponding scales, such as working alliance (Mai et al., 2021) and social presence (Huang et al., 2011).

In the future, based on a deeper understanding of the concept and structure of HAR, there is an urgent need to develop more authoritative and valid self-report scales to ensure research coherence and accuracy, thereby avoiding biases caused by measurement method differences and laying a solid foundation for research progress in this field. Additionally, some objective indicators can be used to supplement the measurement of HAR in experimental processes. Previous research has often used behavioral indicators; for example, Wong and McGee (2012) used the length and fluency of participants' speech as indicators of experienced HAR, while Lubold et al. (2021) recorded the frequency of participants' use of polite language, praise for AI, speaking AI's name, and using "inclusive" language (such as "we") to comprehensively measure HAR, supplemented by self-report scales. Users' facial expressions can also be recognized and used to determine whether rapport exists (Sharma et al., 2021). With the deepening research on the brain and physiological mechanisms of interpersonal rapport (Ellingsen et al., 2022), physiological indicators such as EEG, ECG, and EMG can also be collected from users during interaction with AI to reflect users' perception of rapport.

5.2 Direction of AI Intelligence Enhancement

AI's computing power and intelligence are becoming increasingly advanced with technological development and can gradually assist or replace humans in completing complex tasks. However, from the perspective of HAR as a goal, does higher AI intelligence absolutely positively affect HAR? According to MNT, media that differ greatly from face-to-face communication may require more cognitive effort even if they have rich nonverbal cues and rapid feedback mechanisms, because they do not conform to humans' biological communication mechanisms (Kock, 2004). This further demonstrates that the key factor in establishing HAR is not AI's cognitive, social, and emotional intelligence per se, but how intelligence enhancement affects users' perception of naturalness in human-AI interaction. Existing research in social intelligence (Chattaraman et al., 2019; Pecune et al., 2020) and emotional intelligence (Ranjbartabar et al., 2021; Ter Stal et al., 2021) has also proven that if intelligence enhancement contradicts users' interaction goals, it may negatively affect HAR. However, when AI uses algorithms to understand user needs and goals and intelligently adjusts its sociability and emotional expression, it can effectively improve HAR (Pecune et al., 2020). Therefore, future research can further explore how to improve the

direction of AI intelligence enhancement to better match the goals of human-AI relationship development.

5.3 Exploring More Symbiotic Human-AI Relationships

Previous research on HAR has mostly explored how AI applied in professional service fields or interacting with users in single, short-term laboratory sessions is perceived and evaluated by users (Huang et al., 2011; Pasternak et al., 2022; Ranjbartabar et al., 2021). With social and technological development, AI will assume more roles that require establishing rapport (such as personal assistants, intelligent life companions, virtual partners, etc.), will spend longer time with users, and may have fixed names, stable personality traits and identity, and even experiences formed through long-term co-living (W. Li et al., 2023). This will continuously expand AI's applicable scenarios and application value and give rise to more symbiotic human-AI interaction patterns that emphasize rapport, also bringing more challenges for future human society to accept and adapt to AI technology. Therefore, building upon the new concept and theoretical model proposed in this study, future research should focus on more complex types of human-AI relationships in terms of symbiosis (such as multi-stakeholder perspectives; Law et al., 2024) to stimulate more discussion and practice regarding future visions of human-AI cooperation, driving AI technology toward more natural, pleasant, and sustainable development, and truly realizing the vision of harmonious human-AI symbiosis.

References

侯悍超, 倪士光, 林书亚, 王蒲生. (2024). 当 AI 学习共情：心理学视角下共情计算的主题、场景与优化. 心理科学进展, 32(5), 845-858.

冯道杰. (2024). 人工智能推进新质生产力发展的时代意涵——基于马克思机器观的探赜. 自然辩证法研究, 40(9), 77-82.

王涛, 占小军, 余薇. (2024). AI 感知对员工心理和行为的影响及理论解释. 心理科学进展, 32(7).

Abdulrahman, A., Richards, D., Ranjbartabar, H., & Mascarenhas, S. (2021). Verbal empathy and explanation to encourage behaviour change intention. *Journal on Multimodal User Interfaces*, 15(2), 189–199.

Acosta, J. C., & Ward, N. G. (2011). Achieving rapport with turn-by-turn, user-responsive emotional coloring. *Speech Communication*, 53(9), 1137–1148.

Anzum, F., & Gavrilova, M. L. (2023). Emotion detection from micro-blogs using novel input representation. *IEEE Access*, 11, 19512–19522.

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: Focus on clinicians. *Journal of Medical Internet Research*, 22(6), e15154.

Brouillard, C., Gabbert, F., & Scott, A. J. (2024). Addressing current issues in assessing professional rapport: A systematic review and synthesis of existing measures. *Applied Cognitive Psychology*, 38(3), e4205.

Cade, W. L., Olney, A., Hays, P., & Lovel, J. (2011). Building rapport with a 3D conversational agent. In S. D'Mello, A. Graesser, B. Schuller, & J.-C. Martin (Eds.), *Affective Computing and Intelligent Interaction* (pp. 305–306). Springer.

Cerekovic, A., Aran, O., & Gatica-Perez, D. (2017). Rapport with virtual agents: What do human social cues and personality explain? *IEEE Transactions on Affective Computing*, 8(3), 241-253.

Chandra, S., Shirish, A., & Srivastava, S. C. (2022). To be or not to be …human? Theorizing the role of human-like competencies in conversational artificial intelligence agents. *Journal of Management Information Systems*, 39(4), 969–1005.

Chattaraman, V., Kwon, W.-S., Gilbert, J. E., & Ross, K. (2019). Should AI-based, conversational digital assistants employ social- or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. *Computers in Human Behavior*, 90, 315–330.

Chen, K., Lian, H., Gao, Y., & Li, Y. (2024). Emotional support dialog system through recursive interactions among large language models. In *National Conference on Man-Machine Speech Communication* (pp. 151-163). Singapore: Springer Nature Singapore.

Chen, Y.-C., Yeh, S.-L., Lin, W., Yueh, H.-P., & Fu, L.-C. (2023). The effects of social presence and familiarity on children–robot interactions. *Sensors*, 23(9), 4231.

Cheung, E. O., Slotter, E. B., & Gardner, W. L. (2015). Are you feeling what I'm feeling? The role of facial mimicry in facilitating reconnection following social exclusion. *Motivation and Emotion*, 39(4), 613–630.

Chitkara, P., Modi, A., Avvaru, P., Janghorbani, S., & Kapadia, M. (2019). Topic spotting using hierarchical networks with self attention. *arXiv preprint arXiv:1904.02815*.

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.

D, D., & Juliet, S. (2023). Sentimental analysis based on user emotions using machine learning. In *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*.

De Freitas, J., Agarwal, S., Schmitt, B., & Haslam, N. (2023). Psychological factors underlying attitudes toward AI tools. *Nature Human Behaviour*, 7(11), 1845–1854.

Diederich, S., Brendel, A., Morana, S., & Kolbe, L. (2022). On the design of and interaction with conversational agents: An organizing and assessing review of human-computer interaction research. *Journal of the Association for Information Systems*, 23(1), 96–138.

Ellingsen, D.-M., Duggento, A., Isenburg, K., Jung, C., Lee, J., Gerber, J., Mawla, I., Sclocco, R., Edwards, R. R., Kelley, J. M., Kirsch, I., Kaptchuk, T. J., Toschi, N., & Napadow, V. (2022). Patient–clinician brain concordance underlies causal dynamics in nonverbal communication and negative affective expressivity. *Translational Psychiatry*, 12(1), 1–9.

Ellwart, T., & Kluge, A. (2019). Psychological perspectives on intentional forgetting: An overview of concepts and literature. *KI - Künstliche Intelligenz*, 33(1), 79–84.

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.

Gabbert, F., Hope, L., Luther, K., Wright, G., Ng, M., & Oxburgh, G. (2021). Exploring the use of rapport in professional information-gathering contexts by systematically mapping the evidence base. *Applied Cognitive Psychology*, 35(2), 329–341.

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.

Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007). Creating rapport with virtual agents. In C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, & D. Pelé (Eds.), *Intelligent Virtual Agents* (pp. 125–138). Springer.

Hendrick, C. (1990). The nature of rapport. *Psychological Inquiry*, 1(4), 312–315.

Huang, L., Morency, L.-P., & Gratch, J. (2011). Virtual rapport 2.0. In H. H. Vilhjálmsson, S. Kopp, S. Marsella, & K. R. Thórisson (Eds.), *Intelligent Virtual Agents* (pp. 68–79). Springer.

Hudon, A., Demazure, T., Karran, A., Léger, P.-M., & Sénécal, S. (2021). Explainable artificial intelligence (XAI): How the visualization of AI predictions affects user cognitive load and confidence. In F. D. Davis, R. Riedl, J. Vom Brocke, P.-M. Léger, A. B. Randolph, & G. Müller-Putz (Eds.), *Information Systems and Neuroscience* (Vol. 52, pp. 237–246). Springer International Publishing.

Kasap, Z., & Magnenat-Thalmann, N. (2012). Building long-term relationships with virtual and robotic characters: The role of remembering. *The Visual Computer*, 28(1), 87–97.

Kim, H., So, K. K. F., & Wirtz, J. (2022). Service robots: Applying social exchange theory to better understand human–robot interactions. *Tourism Management*, 92, 104537.

Kock, N. (2004). The psychobiological model: Towards a new theory of computer-mediated communication based on Darwinian evolution. *Organization Science*, 15(3), 327–348.

Kock, N. (2005). Media richness or media naturalness? The evolution of our biological communication apparatus and its influence on our behavior toward E-communication tools. *IEEE Transactions on Professional Communication*, 48(2), 117–130.

Koller, M., Weiss, A., Hirschmanner, M., & Vincze, M. (2023). Robotic gaze and human views: A systematic exploration of robotic gaze aversion and its effects on human behaviors and attitudes. *Frontiers in Robotics and AI*, 10.

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, 5(1), 41–46.

Kulesza, W., Dolinski, D., & Wicher, P. (2016). Knowing that you mimic me: The link between mimicry, awareness and liking. *Social Influence*, 11(1), 68-74.

Law, R., Ye, H., & Lei, S. S. I. (2024). Ethical artificial intelligence (AI): Principles and practices. *International Journal of Contemporary Hospitality Management*, 37(1), 279–295.

Lee, J., Lee, D., & Lee, J. (2024). Influence of rapport and social presence with an AI psychotherapy chatbot on users' self-disclosure. *International Journal of Human–Computer Interaction*, 40(7), 1620–1631.

Li, B., Chen, Y., Liu, L., & Zheng, B. (2023). Users' intention to adopt artificial intelligence-based chatbot: A meta-analysis. *The Service Industries Journal*, 43(15–16), 1117–1139.

Li, W., Chen, S., Sun, L., & Yang, C. (2023). What makes virtual intimacy…intimate? Understanding the phenomenon and practice of computer-mediated paid companionship. *Proc. ACM Hum.-Comput. Interact.*, 7(CSCW1), 112:1-23.

Lin, C.-Y., & Lin, J.-S. C. (2017). The influence of service employees' nonverbal communication on customer-employee rapport in the service encounter. *Journal of Service Management*, 28(1), 107–132.

Lubold, N., Walker, E., & Pon-Barry, H. (2021). Effects of adapting to user pitch on rapport perception, behavior, and state with a social robotic learning companion. *User Modeling and User-Adapted Interaction*, 31(1), 35–73.

Lubold, N., Walker, E., Pon-Barry, H., Flores, Y., & Ogan, A. (2018). Using iterative design to create efficacy-building social experiences with a teachable robot. *International Society of the Learning Sciences, Inc. [ISLS]*, 737–744.

Mai, V., Wolff, A., Richert, A., & Preusser, I. (2021). Accompanying reflection processes by an AI-based StudiCoachBot: A study on rapport building in human-machine coaching using self disclosure. In *HCI International 2021—Late Breaking Papers: Cognition, Inclusion, Learning, and Culture* (pp. 439–457). Springer International Publishing.

Maxim, A., Zalake, M., & Lok, B. (2023). The impact of virtual human vocal personality on establishing rapport: A study on promoting mental wellness through extroversion and vocalics. *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*, 1–8.

McNaughton, K. A., & Redcay, E. (2020). Interpersonal synchrony in autism. *Current Psychiatry Reports*, 22(3), 12.

Namkoong, M., Park, Y., Lee, S., & Park, G. (2024). Effect of AI counsellor's nonverbal immediacy behaviour on user satisfaction: Serial mediation of positive expectancy violation and rapport. *Behaviour & Information Technology*, 1–17.

Neef, C., Mai, V., & Richert, A. (2022). "I am scared of viruses, too"—Studying the impact of self-disclosure in chatbots for health-related applications. In M. Kurosu (Ed.), *Human-Computer Interaction. User Experience and Behavior* (pp. 515–530). Springer International Publishing.

Neequaye, D. A. (2023). Why rapport seems challenging to define and what to do about the challenge. *Collabra: Psychology*, 9(1), 90789.

Nichols, E., Siskind, S. R., Ivanchuk, L., Pérez, G., Kamino, W., Šabanović, S., & Gomez, R. (2022). Hey Haru, let's be friends! using the tiers of friendship to build rapport through small talk with the tabletop robot Haru. *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 6101–6108.

Park, G., Lee, S., & Chung, J. (2023). Do anthropomorphic chatbots increase counseling satisfaction and reuse intention? The moderated mediation of social rapport and social anxiety. *Cyberpsychology, Behavior, and Social Networking*, 26(5), 357–365.

Park, I., Lee, S., & Lee, D. (2023). Virtual audience providing AI-generated emotional reactions to enhance self-disclosure in self-introduction. *International Journal of Human–Computer Interaction*, 39(13), 2702-2713.

Park, J., Woo, S. E., & Kim, J. (2024). Attitudes towards artificial intelligence at work: Scale development and validation. *Journal of Occupational and Organizational Psychology*, 97(3), 920–951.

Pasternak, K., Wu, Z., Visser, U., & Lisetti, C. (2022). Towards building rapport with a human support robot. In R. Alami, J. Biswas, M. Cakmak, & O. Obst (Eds.), *RoboCup 2021: Robot World Cup XXIV* (pp. 214–225). Springer International Publishing.

Pecune, F., Marsella, S., & Jain, A. (2020). A framework to co-optimize task and social dialogue policies using Reinforcement Learning. *Proceedings of the*

20th ACM International Conference on Intelligent Virtual Agents, 1–8.

Pentina, I., Hancock, T., & Xie, T. (2023). Exploring relationship development with social chatbots: A mixed-method study of replika. *Computers in Human Behavior*, 140, 107600.

Pentina, I., Xie, T., Hancock, T., & Bailey, A. (2023). Consumer–machine relationships in the age of artificial intelligence: Systematic literature review and research directions. *Psychology & Marketing*, 40(8), 1593–1614.

Ranjbartabar, H., Richards, D., Bilgin, A. A., & Kutay, C. (2021). First impressions count! The role of the human's emotional state on rapport established with an empathic versus neutral virtual therapist. *IEEE Transactions on Affective Computing*, 12(3), 788–800.

Richards, D. (2017). Intimately intelligent virtual agents: Knowing the human beyond sensory input. *Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents*, 39–40.

Richards, D., & Bransky, K. (2014). ForgetMeNot: What and how users expect intelligent virtual agents to recall and forget personal conversational content. *International Journal of Human-Computer Studies*, 72(5), 460–476.

Riek, L. D., Paul, P. C., & Robinson, P. (2010). When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. *Journal on Multimodal User Interfaces*, 3(1), 99–108.

Seymour, W., & Van Kleek, M. (2021). Exploring interactions between trust, anthropomorphism, and relationship development in voice assistants. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1-16.

Sharma, S., Gangadhara, K. G., Xu, F., Slowe, A. S., Frank, M. G., & Nwogu, I. (2021). Coupled systems for modeling rapport between interlocutors. *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, 1–8.

Steele, C., Lobczowski, N., Davison, T., Yu, M., Diamond, M., Kovashka, A., Litman, D., Nokes-Malach, T., & Walker, E. (2022). It takes two: Examining the effects of collaborative teaching of a robot learner. *International Conference on Artificial Intelligence in Education* (pp. 604–607). Springer International Publishing.

Takahashi, T., Song, S., Baba, J., Nakanishi, J., Yoshikawa, Y., & Ishiguro, H. (2022). Can an empathetic teleoperated robot be a working mate that supports operator's mentality? *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 1059–1061.

Taylor, S. J., Muchnik, L., Kumar, M., & Aral, S. (2023). Identity effects in social media. *Nature Human Behaviour*, 7(1), 27–37.

Ter Stal, S., Jongbloed, G., & Tabak, M. (2021). Embodied conversational agents in eHealth: How facial and textual expressions of positive and neutral

emotions influence perceptions of mutual understanding. *Interacting with Computers*, 33(2), 167–176.

Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, 1(4), 285–293.

Tschopp, M., Gieselmann, M., & Sassenberg, K. (2023). Servant by default? How humans perceive their relationship with conversational AI. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 17(3), Article 3.

Westerman, D., Edwards, A. P., Edwards, C., Luo, Z., & Spence, P. R. (2020). I-it, I-thou, I-robot: The perceived humanness of AI in human-machine communication. *Communication Studies*, 71(3), 393–408.

Wong, J. W.-E., & McGee, K. (2012). Frown more, talk more: Effects of facial expressions in establishing conversational rapport with virtual agents. In Y. Nakano, M. Neff, A. Paiva, & M. Walker (Eds.), *Intelligent Virtual Agents* (pp. 419–425). Springer.

Xie, T., & Pentina, I. (2022). Attachment theory as a framework to understand relationships with social chatbots: A case study of replika. *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2046–2055.

Zhao, R., Romero, O. J., & Rudnicky, A. (2018). SOGO: A social intelligent negotiation dialogue system. *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, 138–145.