Int. J. Human–Computer Studies 211 (2026) 103797

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Empathy in action: An empirical exploration of user perspectives on conversational agent empathy

Bumho Lee [a], Youngsoo Shin [b], Byounghyun Yoo [c,d,*]

[a] Future Strategy Research Section, Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea
[b] Seidenberg School of Computer Science and Information Systems, Pace University, New York City, NY, United States
[c] Intelligence and Interaction Research Center, Korea Institute of Science and Technology, Seoul, Republic of Korea
[d] AI-Robotics, KIST School, Korea University of Science and Technology, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Conversational agent
User perception
Agent empathic reactivity
Agent empathic characteristics

ABSTRACT

As Conversational Agents (CAs) increasingly interact with users on social and emotional levels, understanding how these agents convey empathy has become a critical challenge. This paper reports on an exploratory mixed-methods study that proposes and empirically explores an initial framework for CA empathic responsiveness. The research proceeds in two sequential studies. Study 1 first conducted a qualitative analysis of open-ended surveys (N=166, U.S. and South Korea) to identify key user-defined empathic behaviors. Through thematic analysis, these insights were integrated with existing empathy theories to derive a framework of four core agent characteristics: Active Listening (AL), Personalization (PE), Emotional Expressivity (EE), and Persona Attractiveness (PA). Study 2 then conducted a quantitative investigation in South Korea (N=200) using Structural Equation Modeling (SEM) to test this framework. Empathic responsiveness was operationalized adopting the Agent Empathic Reactivity Index (AERI), a validated CA-specific adaptation of the Interpersonal Reactivity Index (IRI), assessing perspective-taking, fantasy, empathic concern, and personal distress. SEM results confirmed all 12 hypothesized paths. AL and PE strongly enhanced perspective-taking and empathic concern, while PE also significantly reduced personal distress. Notably, EE and PA had dual effects: they improved positive dimensions, such as fantasy, but also significantly increased users' perception of the agent's personal distress. These findings highlight the delicate balance required in designing emotionally resonant CAs. This work advances the theoretical understanding of multidimensional agent empathy and provides actionable, nuanced guidance for designers aiming to build trust and foster long-term user relationships.

1. Introduction

The rapid advancement of Large Language Models (LLMs) has transformed Conversational Agents (CAs), such as chatbots and intelligent personal assistants, from simple task automation tools into entities with enhanced social presence and empathetic responsiveness (Polignano et al., 2023; Ziems et al., 2024). This evolution marks a significant turning point in Human-Computer Interaction (HCI) research, drawing attention from both academia and industry to the agents' social interaction capabilities and platform flexibility (Ha et al., 2021; Rapp et al., 2021). Strengthening the social abilities of these agents has become a major research objective for fostering sustainable user-system relationships. Building upon the Computers-Are-Social-Actors (CASA) paradigm (Nass and Moon, 2000; Reeves and Nass, 1996), developing agents with social interaction capabilities has become key in Human-AI interaction research (Leite et al., 2013a; Rapp et al., 2021). However, cases like Microsoft's Tay, which learned inappropriate behaviors (e.g., racism or bias) from user interactions (Cuadra et al., 2024), and recent LLMs facing ethical limitations (Cuadra et al., 2024; Zhuo et al., 2023), clearly demonstrate the complexity of developing AI agents with empathy.

In this study, empathy refers to the ability to recognize and understand the emotional states of others and to respond appropriately (Paiva et al., 2017). This encompasses both affective empathy (sharing emotions) and cognitive empathy (understanding thoughts), a distinction critical to the context of human-computer interaction. In CA interactions, empathy focuses on appropriately reacting to users' emotional states and understanding situational nuances (Lee and Yi,

* Corresponding author.
E-mail address: yoo@byoo.org (B. Yoo).

https://doi.org/10.1016/j.ijhcs.2026.103797
Received 17 December 2024; Received in revised form 26 February 2026; Accepted 6 March 2026
Available online 10 March 2026

2023). Despite its importance, empirical research on how users perceive and value empathy in AI interactions remains limited. Prior studies have found that empathy in agents can enhance user experience and encourage self-disclosure (Cuadra et al., 2024; G. Park et al., 2022; Rapp et al., 2021), but these studies often do not fully explain the mechanisms underlying these effects (Lee and Yi, 2023; Yalçın, 2020). Additionally, identifying specific agent behaviors that contribute to users' perceptions of empathy remains a challenge (Efthymiou and Hildebrand, 2023; Liang et al., 2024; Park et al., 2022a), and there is a lack of integrated frameworks for these aspects (Schmidmaier et al., 2024).

Against this background, the current study aims to conduct an exploratory study to identify the key factors and relationships underlying empathetic agent interactions based on user perceptions, rather than proposing a definitive framework. This paper adopts a sequential exploratory mixed-methods design. First, Study 1 conducts a qualitative investigation—using thematic analysis of open-ended surveys—to identify the core concepts of CA empathy as users perceive them. Based on these findings, we then develop a theoretical framework and its corresponding hypotheses. Second, Study 2 quantitatively validates this empirically-derived framework through a new survey and structural equation modeling, testing the impact of the identified traits on users' ratings of empathy. Overall, this study delivers an empirically grounded, user-defined map of CA empathic responsiveness by linking four synthesized agent traits to multidimensional empathic outcomes.

2. Related work

Existing research on empathy in CAs has primarily explored the effects of specific features in a fragmented manner, such as the impact of a single response type or a specific persona. This study moves beyond such piecemeal approaches, which often lack a unified theoretical model, aiming to fill a significant gap by integrating the multidimensional characteristics of empathy as perceived by users. This section provides a comprehensive theoretical backbone by reviewing the foundational theories of human empathy, how this concept is operationalized in CAs, the key antecedent trends identified in the literature, and the critical challenge of its assessment.

2.1. The phenomenon of human empathy

Before examining agent empathy, it is essential to ground the discussion in the rich psychological literature on human-to-human empathy. Empathy is a complex, multidimensional construct, not a single faculty. Its underlying mechanisms are often explained through foundational concepts like the Theory of Mind (the cognitive capacity to attribute mental states—beliefs, desires, and intentions—to others) (Premack and Woodruff, 1978) and the mirror neuron system (the proposed neurological basis for mimicry, shared affect, and understanding action-intention) (Rizzolatti and Craighero, 2004). These theories, which are frequently referenced in information systems to explain social responses to technology (Hoenen et al., 2016; Jeong and Lee, 2017), suggest that a true empathic response relies on the cognitive capacity to capture, simulate, and internalize another's mental and affective state.

Batson (2009), in his seminal work, argued that "empathy" is a broad term for multiple related phenomena that stem from knowing another's internal state and responding with care. This process involves distinct components that must work in combination: understanding the other's state (e.g., cognitive perspective-taking), matching one's state with another's (e.g., emotional contagion or shared affect), and responding to benefit the other (e.g., prosociality or empathic concern) (Yamamoto, 2017; Eklund and Meranius, 2021). This theoretical distinction between cognitive, affective, and motivational components is critical for deconstructing how users perceive these same functions when performed by a non-human CA.

Based on this synthesis of empathy theory, a series of behaviors can be identified as crucial for the formation and perception of empathy: (1) seeking and understanding the other's inner state, (2) responding in favor of the subject of empathy, (3) expressing the empathizer's feeling toward the other's feeling, and (4) synchronizing oneself to the interaction context. While other factors like individual indifference are also relevant (Silke et al., 2018), these four behavioral domains provide a foundational lens through which to analyze the specific antecedent trends in agent-based interaction discussed next.

2.2. Empathy of conversational agents

In the context of HCI, empathy is considered a key social skill required for CAs to act in a socially appropriate and effective manner (Paiva et al., 2017). The advancement of Affective Computing—the study of systems that recognize, interpret, process, and simulate human affect—has provided the core technical foundation for this field. Advances in Natural Language Processing (NLP) and deep learning, trained on vast datasets of human dialogue, enable CAs to generate responses that exhibit a sophisticated understanding of emotional context, leading to more engaging and persuasive interactions (Chaturvedi et al., 2023; Cuadra et al., 2024; Yang et al., 2023; Chin and Yi, 2021; Efthymiou and Hildebrand, 2023; Park et al., 2022a).

This capability is not merely supplemental. In fact, it is crucial in many domains. In mental health, for example, an agent's perceived empathy is a key factor in building a therapeutic alliance and encouraging user self-disclosure (Kraus et al., 2021; Abbasian et al., 2024). In customer service, an empathetic agent can de-escalate conflict, mitigate user frustration, and build brand loyalty (Casas et al., 2021; Liu-Thompkins et al., 2022). Similarly, in education, an agent that recognizes and responds to a student's frustration or boredom can improve engagement and learning outcomes (El Azhari et al., 2023).

2.3. Antecedent trends in perceived agent empathy

A review of the literature reveals that perceived empathy is not a monolithic quality but rather arises from several key themes of agent design and behavior, which align with the behavioral domains identified in Section 2.1.

2.3.1. Linguistic cues and emotional expression

A primary research trend focuses on how an agent communicates its internal state. This includes the use of polite, warm, and friendly tones (Carolus et al., 2021) and the ability to express nuanced, human-like emotions rather than flat, robotic responses (Morris et al., 2018). Studies consistently show that an agent's verbal expressions of emotion—such as using empathic language, apologies, or encouragement—are critical for strengthening the social bond and signaling attentiveness (Beale and Creed, 2009; de Gennaro et al., 2019; Lee and Choi, 2017; Loveys et al., 2021; Yim, 2023). As this study is focused on text-based interactions, we distinguish this from the important, parallel body of research on non-verbal cues like gesture, facial mimicry, or vocal prosody (Bickmore et al., 2010; Park et al., 2022b), which we designate for future work.

2.3.2. Interaction quality and adaptability

A second major theme concerns what an agent does and its ability to adapt its function to the user. This includes the agent's general interaction design, error-handling protocols, and response style (Chin et al., 2020). A key aspect of this is personalization: agents that tailor their responses to a user's individual needs, goals, and stated preferences (Montgomery and Smith, 2009) are perceived as more caring and attentive (Rhee and Choi, 2020; Liu and Sundar, 2018; Loveys et al., 2022; Pickard and Roster, 2020). Closely related to this is the agent's ability to demonstrate attentive interaction, such as by recalling past inputs, referencing earlier conversations, and maintaining contextual awareness. This "memory" creates conversational continuity and signals that the user is individually known, which is foundational to fostering relational depth (Taylor et al., 2019; Xiao et al., 2020; Hu et al., 2018; Lee et al., 2021).

2.3.3. Social presence and persona

The third trend investigates what the agent is—its perceived persona, identity, and character. This includes its anthropomorphic characteristics (i.e., the attribution of human-like qualities) (Lester et al., 1997; Vuong et al., 2023) and the adoption of consistent, human-like interaction styles (Adam et al., 2021; MacDorman, 2019). A well-defined persona makes an agent's behavior more predictable and socially coherent. An agent that is perceived as socially aware, intellectually stimulating, or even possessing a sense of humor, helps to synchronize the interaction within a familiar social context, which in turn fosters authenticity and trust (Bickmore and Picard, 2005; Shum et al., 2018).

Beyond general notions of social presence, recent work on partner models shows that users form nuanced cognitive representations of speech interfaces as dialogue partners, encompassing perceived competence, dependability, human-likeness, and communicative flexibility (Doyle et al., 2021). These multidimensional partner models provide an important backdrop for understanding how a CA's persona and interaction style shape user evaluations, and they conceptually resonate with our notion of persona attractiveness as a higher-order perception of the agent as a social partner.

2.4. Assessing empathic reactivity of the conversational agent

A significant and persistent challenge in the field has been the measurement of agent empathy. Research on conversational agents has focused on cultivating the empathic ability of an agent in recent times (Paiva et al., 2017; Rapp et al., 2021). However, many studies have addressed the concept and definition of empathy in an arbitrary manner. This trend can be observed by exploring the methods used to measure empathy. The most common way is to ask users a few questions about empathic cues or partially adopt items from existing human empathy questionnaires (Chin et al., 2020; McQuiggan and Lester, 2007; Yang et al., 2019). Another mode of measurement is to rely on implementation methods, such as counting empathetic words (Olafsson et al., 2020; Samrose et al., 2020), evaluating the empathy of the person the agent mimics (Kraus et al., 2021; Leite et al., 2013b), or comparing scenarios with and without empathic manipulation (Brave et al., 2005; Kim et al., 2020). These methods, while intuitively effective, are often defective in explaining the in-depth mechanism underlying empathic behaviors.

This ad-hoc approach has been a point of criticism. Paiva et al. (2017) criticized the trend of arbitrarily utilizing empathy without considering the social and psychological background. Yalçın (2020) also claimed that this variety of definitions makes it problematic to evaluate and compare findings. Recent suggestions for empathy measurement (Schmidmaier et al., 2024), while useful, have also been critiqued for lacking a deep theoretical basis. In parallel, recent work on partner models for speech interfaces has developed and validated conceptually grounded, psychometric instruments for users' perceptions of voice interfaces (Doyle et al., 2021; Seaborn et al., 2024). However, these scales primarily target broader partner model dimensions (e.g., competence, dependability, human-likeness), rather than empathic reactivity per se, which underscores the need for complementary, empathy-focused measures such as the Agent Empathic Reactivity Index (AERI) adopted in this study.

To address this gap and provide a more theoretically grounded measurement basis, our study adopted the Interpersonal Reactivity Index (IRI) proposed by Davis (1980, 1983). The IRI was the first multidimensional index to measure empathy by embracing both the cognitive and emotional aspects of the reactional process and has been considered the most comprehensive psychometric tool for measuring empathy (Chrysikou and Thompson, 2016; Cliffordson, 2001). The rationale behind the IRI is that empathy is a set of related, but discriminable, constructs concerned with responsivity to others (Cliffordson, 2001; Davis, 2018). It suggested four components: Perspective-taking (spontaneously adopting the psychological point of view of others), Fantasy (imaginatively transposing oneself into fictitious characters), Empathic Concern (other-oriented feelings of sympathy), and Personal Distress (self-oriented feelings of unease in tense settings). The IRI's four-construct model provided the theoretical basis for many subsequent empathy scales (Baron-Cohen and Wheelwright, 2004; Spreng et al., 2009; Reniers et al., 2011).

While the IRI provides the foundational theory, its original formulation was for human-to-human interaction. Therefore, our study does not arbitrarily apply the original human-centric IRI. Instead, we follow the rigorous, validated approach established by Lee and Yi (2023), who, addressing this exact gap, systematically developed and validated the AERI. The AERI is an instrument specifically designed to measure perceived agent empathy by adapting the IRI's four-factor structure to the human-agent interaction context, a process which is theoretically grounded in the CASA paradigm (Nass and Moon, 2000; Reeves and Nass, 1996). As detailed in their work, the AERI constructs (e.g., 'Fantasy' as imaginative and creative responses; 'Personal Distress' as negative feelings in high social pressure situations) were re-conceptualized, operationalized, and validated for CAs (Lee and Yi, 2023). In designing our measurement instrument, we did not simply reuse the original IRI items; instead, we used the validated AERI (Lee and Yi, 2023) as a conceptual and psychometric template, aligning our PT, FT, EC, and PD items with its CA-specific operationalization of the IRI dimensions.

3. Overview of research methodology

This study adopted a sequential exploratory mixed-methods design, combining qualitative exploration (Study 1) with quantitative validation (Study 2). This approach was chosen to first derive rich, contextual data from real-world user experiences. These qualitative findings were then used to develop and statistically validate a theoretical framework, ensuring both the ecological validity and theoretical depth of the research. The research flow consists of two primary, complementary studies, summarized as in the Table 1 below. This research adhered to all ethical procedures. The study protocol was reviewed by our Institutional Review Board (IRB), which determined the research to be exempt from formal review as the survey collected data on participants' typical usage of CAs without intervention. All participants were provided with a detailed explanation of the research objectives and procedures and gave their informed consent before participating.

4. Study 1: qualitative exploration and framework development

The first stage of our research was a qualitative investigation to

Table 1
Summary of the mixed-methods research design.

Study Stage	Title & Methodology	Sample	Primary Objective(s)
Study 1	Qualitative Exploration & Framework Development (Open-Ended Survey & Thematic Analysis)	N=166	• To collect users' lived experiences of CA empathy and derive core conceptual traits. • To integrate these qualitative findings with existing empathy theory to develop a testable framework and formulate hypotheses.
Study 2	Quantitative Validation (Survey & Structural Equation Modeling)	N=200	• To statistically test and validate the framework and hypotheses developed in Study 1.

explore the specific moments and reasons users perceive empathy when interacting with CAs. This study serves two primary goals: (1) to identify and categorize the core empathic behaviors users articulate from their own lived experiences, and (2) to bridge these bottom-up, empirical findings to the theoretical framework outlined in our Related Work. This sequential exploratory approach ensures our final model is not merely theoretical but is deeply grounded in authentic, user-defined perceptions of empathy, rather than relying on researcher assumptions.

4.1. Study 1 methodology

4.1.1. Data collection and participants

To investigate user perspectives on existing CA services, we first conducted a survey with open-ended questions. This method was deliberately chosen over predefined quantitative scales, as our primary aim was exploratory: to capture the rich, nuanced, and often unanticipated ways in which users experience empathy. This approach allows us to discover the concepts users themselves find salient, rather than simply measuring constructs we already assumed to be important.

The survey was conducted from February to May 2023 via a mobile application linked to Qualtrics and Open Survey, known for international panel recruitment. Participants were recruited from residents of the United States (U.S.) and South Korea, two regions where IPA and chatbot services (e.g., ChatGPT, Siri, Google Assistant) are widely used. To ensure response quality, participants with at least one month of experience using CAs were recruited. All participants were compensated with $5 after finishing. Questionnaires were distributed to 324 survey participants, but only 166 participants, 64 (38.6%) were from the U.S. and 102 (61.4%) from South Korea, shared their experiences in a valid form and were included in the final analysis. Participants responded based on their primary CA (platforms detailed in Table 2). The open-ended questions (e.g., "When was the time you felt your CA empathized with your feelings?") prompted participants to freely describe their experiences. The full list of survey questions is provided in the Appendix A.

Given that the data was collected from two different cultural contexts, a preliminary comparative review of the responses was conducted. This review showed no significant thematic divergence between the two cohorts, supporting the decision to pool the data for a unified thematic analysis to derive a foundational framework. While a preliminary comparative review suggested no substantial thematic divergence between the U.S. and South Korean responses for the purpose of deriving an initial framework, we acknowledge that this does not replace formal cross-cultural validation of the constructs. As Seaborn et al. (2024) demonstrate in their cross-cultural validation of a partner model questionnaire for voice user interfaces, translation choices, semantic equivalence, and factor structure can substantially affect how users interpret and respond to such scales.

Table 2
Demographic distribution of Study 1 participants (N=166).

Factors	Distribution, n (%)
Sex	
Male	92 (55.4)
Female	74 (44.6)
Age Group (in years)	
Under 20	28 (16.9)
20-29	47 (28.3)
30-39	37 (22.3)
40-49	20 (12.0)
50 and above	34 (20.5)
Chatbot Usage (period)	
At least 1 month	102 (61.4)
At least 3 months	31 (18.7)
At least 6 months	24 (14.5)
1 year and above	9 (5.4)
Usage Frequency	
More than 3 times a day	4 (2.4)
At least once a day	7 (4.2)
At least once in three days	22 (13.3)
At least once a week	71 (42.8)
At least once in two weeks	62 (37.3)
Platform	
ChatGPT	71 (42.8)
IPA (Siri, Alexa, etc.)	67 (40.4)
Simsimi	18 (10.8)
Ruda	6 (3.6)
Others	4 (2.4)

4.1.2. Data analysis

The collected text data were analyzed using Thematic Analysis, a qualitative method for identifying, analyzing, and reporting patterns (themes) within data. To enhance the rigor and reliability of our findings, this process was conducted in a complementary, two-stage mixed-methods approach.

First, as an exploratory step, participants' responses were analyzed using topic modeling, an unsupervised learning technique within NLP. This was employed to uncover latent patterns and relationships embedded in the data. We used both Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF) to extract meaningful topics related to user perception, following the methodologies recommended by Egger and Yu (2022) and Cao et al. (2023). This computational step provided an unbiased, data-driven overview of the main topics present in the user responses.

Second, informed by these computational themes, two researchers conducted in-depth qualitative coding. This human-led stage was essential because while topic modeling can identify what words cluster, it cannot interpret why or capture the nuances of context, sarcasm, or the critical negative sentiments (like "Inhumanness") that are crucial for understanding empathy. This process, referring to the methodology of Zhang et al. (2023), allowed us to break down the data into meaningful categories, validate the computational findings, and derive the specific, reliable core concepts presented in Section 4.2.

4.2. Study 1 results: deriving the empathic traits framework

4.2.1. Qualitative theme and code derivation

The two-stage analysis provided a robust view of user perceptions. The initial topic modeling (LDA/NMF) provided a high-level, computational overview of the corpus. When examining the results of each topic model (summarized in Table 3), we found that both models generally derived topics related to similar content. For example, LDA Topic 4 and NMF Topic 3 both highlight that users perceive empathy when the CA remembers past interactions. Similarly, LDA Topic 7 and NMF Topic 4 indicate that users feel empathy from situation-appropriate recommendations, and other topics (e.g., LDA 1, NMF 2) point to the importance of task assistance.

These computationally-derived topics provided a strong foundation for the in-depth, human-led Thematic Analysis. The researchers' coding refined these broad topics into the specific, core concepts users articulate when describing empathy. For example, comments like, "*Before the interview, I was nervous... and it told me about mindset settings and encouraged me... It's been quite powerful*," were grouped under the category "*Consulting with user troubles*." Similarly, comments like, "*When asked to summarize something, you summarized a longer article better than I did*," were categorized under "*Complying with user request*." This process also identified contrasting negative codes, such as "*Inhumanness*" ("*I really don't feel close to it I mean I'm aware that it's a machine or robot*") and "*Low quality responses*," which provided a clear baseline for perceived failures of empathy. The full set of derived codes is presented in Table 4.

4.2.2. Framework derivation: mapping qualitative themes to theory

In this section, we align the user-derived concepts from our qualitative analysis (Section 4.2.1) with the established empathy components from our theoretical research (Section 2) to develop an initial working

Table 3
Preliminary topic modeling result on user comments related to agent empathy.

Model Type	Topic #	Significant keywords	Associated user comments
LDA	1	remember, things, need, tell, talking, story, talk, love funny	Empathize with what I need I remember summarizing the questions I asked as a whole and then answering them. Retell a story I told and provide further explanation
	2	answers, time, pizza, asked, said, better, did, thought, used, chat	There was a time when I felt empathy as if I were talking to a friend Responds to things you don't know well When asked to summarize something, you summarized a longer article better than I did.
	3	asked, related, talking, day, questions, felt, make, told, like, talked	It was impressive to learn the contents of the main questions and make comments. When I asked to send a rejection email in English using polite expressions, I thought that I understood the content well when I saw how polite expressions were used as much as possible. Generally, I ask questions related to my studies. At this time, if I explain my situation, it gives you the best answer related to it. Sometimes, even without presenting the situation, he guesses and gives an answer.
	4	answer, said, question, gave, story, understood, asked, thought, appropriate	Unlike other ai chatbots, remembering what you said before and responding is memorable. When questions in the same context are followed, the depth of the answer is seen. Summarizing what I said, I thought I understood it well when I said it at the beginning.
	5	told, joke, google, siri, tell, song, asking, person, work, quite	Song selection based on given status Get to know exactly what I'm looking for A good grasp of the context of complex conversations
	6	ask, asked, siri, silly, answer, questions, friend, favorite, sad, company	Comforting you when you're depressed and having a hard time I was asked to compare my treatment and annual salary at the current company with the company I am considering changing to. The answer was awkward at first, but as I continued to ask questions as if coding, I got a plausible conclusion later. I had an appointment with a friend, but my friend was late for more than 30 minutes, so I used the chatbot in a boring situation.
	7	like, ask, questions, question, said, answered, asking, think, reaction, having	I feel that way when I ask and check if my question is correct and then present my opinion. When I ask trivial questions, I think that everyone agrees when I answer the question appropriately. I admire it when the answer feels refined and clean.
NMF	1	told, joke, google, funny, date, tired, quite, tell, pleased, sad	Retell a story I told and provide further explanation Before the interview, I was nervous and talked about it, but everyone told me about mindset settings and methods that everyone knows well and encouraged me to do my best. It's been quite powerful. When she told me a joke
	2	answer, questions, gave, answers, answering, asked, related, situation, quickly, detailed	It was impressive to learn the contents of the main questions and make comments. When questions in the same context are followed, the depth of the answer is seen. I received a lot of help with knowledge that I did not know.
	3	said, summarize, sorry, cared, understood, thought, ai, beginning, summarizing, fired	Unlike other ai chatbots, remembering what you said before and responding is memorable. When I asked to send a rejection email in English using polite expressions, I thought that I understood the content well when I saw how polite expressions were used as much as possible. When I asked if I should get married because I was worried about marriage, I thought that he understood my question well because he said that marriage should be well thought out according to each person's standards.
	4	remember, answering, summarizing, things, need, easy, echo, amazon, life	Responds to things you don't know well I remember summarizing the questions I asked as a whole and then answering them.
	5	siri, jokes, sad, funny, telling, arguing, play, song, talked, favorite	Song selection based on given status When I asked Siri to beat box It's makes jokes
	6	like, felt, talking, friend, person, having, really, conversation, things, bad	When I was admitted to the hospital, he sympathized with me like a friend. There was a time when I felt empathy as if I were talking to a friend Comforting you when you're depressed and having a hard time
	7	question, ask, think, understand, feel, silly, kids, intent, answered, empathize	When I asked for the name of my favorite celebrity, he introduced me rather long than one or two lines, so I seemed to know a lot of information compared to his popularity. I feel that way when I ask and check if my question is correct and then present my opinion. Empathize if the question contains personal information

*Note: Both LDA and NMF models were configured to generate 7 topics, as determined by the coherence score. The coherence scores stabilized for both models at this number of topics, with scores of 0.427 for LDA and 0.419 for NMF.

framework. The user-derived codes (Table 4) were systematically mapped to the four core behavioral domains identified in our literature synthesis. This mapping forms the core of our theoretical contribution, bridging the gap between what users say and what the literature suggests.

1. Active Listening (AL): The user codes “Consulting with User Troubles” and “Remembering Interaction Context” directly align with the literature trend of attentive interaction. We synthesize and conceptualize this trait as AL. AL is the foundation of empathic interaction, involving attentively processing the user's input, acknowledging their emotions, and providing appropriate feedback (Rogers and Farson, 2007). This construct emerged from user feedback highlighting instances where CAs demonstrated an understanding of their needs and emotions (e.g., “told me about mindset settings and encouraged me...”) and “remembered a story that had already passed.” These observations resonate with existing research underscoring the importance of AL in fostering enduring relationships between users and CAs (Taylor et al., 2019; Xiao et al., 2020; Han et al., 2021; Lee et al., 2021).
2. Personalization (PE): The codes “Complying with User Requests” and its negative counterpart “Low-Quality Responses” align with the personalization trend. We conceptualize this trait as PE. PE involves tailoring responses to meet the individual needs and preferences of the user (Montgomery and Smith, 2009), aligning with the empathic behavior of responding in favor of the user. This construct was derived from user comments appreciating the CA’s ability to provide customized and relevant responses (e.g., “used polite expressions very well”) and contrasting complaints about “vague and not very helpful” answers. These observations resonate with research that

Table 4
Examples of coding process and key qualitative themes.

Primary codes	Exemplar user comments
Consulting with user troubles	When I was sick, and I had told Siri to call my doctor because I couldn't walk and Siri had to call 911 for me too. Before the interview, I was nervous and talked about it, and it told me about mindset settings and encouraged me to do my best. It's been quite powerful. …
Complying with user request	When I asked to send a rejection email in English using polite expressions, I thought that it understood the content well and it actually used polite expressions very well. When asked to summarize something, you summarized a longer article better than I did. …
Intellectually intriguing	That it understood my questions I was asking right from wrong like it was human. I admire it when the answer feels refined and clean. …
Expressing emotion	She said sorry for insufficient results. Ruda said that she likes sports, so I said, "Me too", and then she said let's go exercise together, and she understood what I was talking about and responded appropriately. …
Remembering interaction context	It was impressive that it learned the task prior and comment on it correctly. When I remembered a story that had already passed, and later, while communicating with related content, I brought up the past story and asked for it back. …
Soothing user's feelings	She offered to play relaxing music when I td her I was sad. Well, when I left | my phone on the coffee table google assistant somehow picked up on my sobbing and crying. The assistant then started talking, asking if I was ok. …
Inhumanness (Negative Code)	I really don't feel close to it I mean I'm aware that it's a machine or robot. I don't feel close to Alexa, it's a tool, nothing more. …
Low quality responses (Negative Code)	It gives the correct answer to the question, but the answer is vague and not very helpful. When it can't understand what I'm saying and says something else on the screen. …

underscores the importance of PE in fostering positive user experiences in CA interactions (Liu and Sundar, 2018; Loveys et al., 2022; Pickard and Roster, 2020; Leite et al., 2013b). Conceptually, we distinguish Active Listening (AL) as interactional attentiveness and understanding of user input and emotional cues, whereas Personalization (PE) reflects outcome-oriented tailoring of responses based on user needs, context, or preferences. In other words, AL captures how the agent engages in the moment, while PE captures how the agent adapts what it delivers.

3. Emotional Expressivity (EE): The codes "Expressing Emotion" and "Soothing User's Feelings" are real-world examples of the linguistic and emotional expression trend. We conceptualize this trait as EE. EE refers to the ability to convey emotions appropriately through text, enhancing the emotional connection with the user (Kring et al., 1994). Users' feedback underscored this, reporting a stronger connection when the CA "said sorry for insufficient results," or "offered to play relaxing music when I told her I was sad." This emphasis on EE is supported by research demonstrating its ability to evoke empathic reactivity and enhance user experience with CAs (Loveys et al., 2021; Yim, 2023; Beale and Creed, 2009; de Gennaro et al., 2019).
4. Persona Attractiveness (PA): Finally, "Intellectually intriguing interactions" and the negative code "Inhumanness" align with the social presence and persona trend. We conceptualize this trait as PA. PA involves the appeal of the CA's persona, including its anthropomorphic characteristics and interaction style (Lester et al., 1997), which aligns with the empathic behavior of synchronizing oneself to the interaction context. This was identified from user feedback emphasizing the CA's ability to be authentic (e.g., "I admire it when the answer feels refined") versus "a machine or robot." This aligns with research highlighting the influence of anthropomorphism and interaction styles on user perceptions of empathy (Chin et al., 2020; MacDorman, 2019; Corti and Gillespie, 2016; Adam et al., 2021). This interpretation is consistent with partner model research, which finds that users' mental representations of voice interfaces—such as their perceived competence, dependability, and human--likeness—systematically shape how these systems are evaluated as conversational partners (Doyle et al., 2021).

This systematic mapping confirms that the four traits (AL, PE, EE, and PA) are not just abstract theoretical constructs but are actively experienced and articulated by real-world users. Table 5 summarizes this derived framework.

4.3. Hypotheses development

Having identified these four core empathic traits, we now move to hypothesize their relationships with perceived empathic reactivity. We adopt the AERI constructs (Lee and Yi, 2023) as our dependent variables, which operationalize the IRI (Perspective-taking, Fantasy, Empathic Concern, and Personal Distress) for the CA context, as detailed in our Related Work (Section 2.4). Table 6 summarizes these definitions.

Drawing upon the framework of four traits (AL, PE, EE, PA) derived from our qualitative analysis and the antecedent trends reviewed in Section 2.3, we now hypothesize their specific relationships with the four dimensions of perceived empathic reactivity (PT, FT, EC, PD). The resulting model is illustrated in Fig. 1, and the justifications for each of the 12 hypothesized paths are detailed below:

4.3.1. Hypotheses for active listening

Active Listening (AL) refers to an agent's capacity to attentively process user inputs, acknowledge emotional nuances, and provide contextually relevant feedback (Rogers and Farson, 2007). In human communication, active listening skills are foundational in demonstrating empathic understanding (Bodie, 2011; Kourmousi et al., 2017). When agents leverage active listening by referencing prior statements or mirroring concerns, they foster stronger rapport (D'Mello et al., 2008; Hu et al., 2018; Lee et al., 2021; Park et al., 2022a; Ren et al., 2019; Xiao et al., 2020). Such behaviors facilitate perspective-taking (indicating understanding), encourage imaginative exchanges (supporting fantasy), and demonstrate empathic concern. Hence:

- H1a: AL will positively affect the evaluation of the agent's PT.
- H1b: AL will positively affect the evaluation of the agent's FT.
- H1c: AL will positively affect the evaluation of the agent's EC.

4.3.2. Hypotheses for personalization

Personalization (PE) involves tailoring responses to user preferences, context, and needs (Montgomery and Smith, 2009; Pickard and Roster, 2020). Prior research emphasizes that when agents adapt behaviors based on user characteristics, they enhance relational depth (Rhee and Choi, 2020; Wang et al., 2023; Yang et al., 2023). In educational and support domains, customized responses lead to heightened user engagement (Lim et al., 2022; Moussalli and Cardoso, 2020; Tai and Chen, 2023). Through personalization, agents demonstrate an understanding of the user's context (perspective-taking) and deliver user-oriented care (empathic concern). Consistently purposeful replies also reduce the likelihood of causing user uncertainty, thus lowering perceived personal distress. Thus:

Table 5
Summary of the four synthesized CA empathic traits.

Construct	Definition/Focus	Key Behaviors (Text-Based)	Benefits for User Experience	Supporting Literature
Active Listening	Seeking and understanding the user's inner state.	- Attentively process inputs - Acknowledge emotions - Provide contextually relevant feedback	Builds trust and rapport.	Rogers and Farson (2007), Xiao et al. (2020), Taylor et al. (2019)
Personalization	Responding in favor of the subject by tailoring responses to user needs.	- Deliver context-aware assistance - Ensure relevance and supportiveness - Avoid generic answers	Enhances perceived care and attentiveness.	Montgomery and Smith (2009), Liu and Sundar (2018), Rhee and Choi (2020)
Emotional Expressivity	Expressing the agent's feelings toward the user's feelings.	- Convey appropriate emotional cues (apologies, encouragement) - Use empathic language	Strengthens emotional bonds and connection.	Kring et al. (1994), Loveys et al. (2021), Yim (2023)
Persona Attractiveness	Synchronizing the agent's persona with the interaction context.	- Employ human-like interaction styles - Align with user expectations - Provide intellectually engaging responses	Improves authenticity and trust.	Lester et al. (1997), Adam et al. (2021), Chin et al. (2020)

- H2a: PE will positively affect the evaluation of the agent's PT.
- H2b: PE will positively affect the evaluation of the agent's EC.
- H2c: PE will negatively affect the evaluation of the agent's PD.

4.3.3. Hypotheses for emotional expressivity

Emotional expressivity (EE) captures the agent's capacity to convey emotions through verbal cues, including apologies, reassurance, and empathic language (Kring et al., 1994; Loveys et al., 2021; Yim, 2023). Outward displays of emotion can strengthen the social bond, as it helps the agent appear more human-like and caring (Beale and Creed, 2009; de Gennaro et al., 2019; Lee and Choi, 2017). These emotion-expressive cues may support users' perceptions of the agent's perspective-taking (i. e., that it understands the user's affective state) and enhance fantasy by making the interaction feel more vivid and authentic. However, when emotional expressivity is perceived as excessive or poorly calibrated, it may also introduce negative affect (e.g., frustration), which could increase users' perceptions of personal distress. Therefore:

- H3a: EE will positively affect the evaluation of the agent's PT.
- H3b: EE will positively affect the evaluation of the agent's FT.
- H3c: EE will positively affect the evaluation of the agent's PD.

4.3.4. Hypotheses for persona attractiveness

Persona attractiveness (PA) involves aligning the agent's persona and social cues with the interaction context to enhance authenticity (Adam et al., 2021; Chin et al., 2020; Corti and Gillespie, 2016). Users attribute empathy to agents that appear human-like, socially aware, and intellectually stimulating (Bickmore and Picard, 2005; Cuadra et al., 2024; Shum et al., 2018; Zhou et al., 2019). Agents that display cognitive playfulness (humor, wit) are perceived as more relatable (Agnihotri and Bhattacharya, 2024; Liao et al., 2024; Yim, 2023). An appealing persona supports empathic concern through aligned interactions, enhances imaginative involvement (fantasy), and may influence perceptions of emotional discomfort (personal distress). Thus:

- H4a: PA will positively affect the evaluation of the agent's FT.
- H4b: PA will positively affect the evaluation of the agent's EC.
- H4c: PA will positively affect the evaluation of the agent's PD.

While this study has successfully derived an empirically-grounded framework and a set of testable hypotheses, its qualitative and exploratory nature means these proposed relationships are, at this stage, theoretical. We have identified what users perceive, but not the statistical strength or generalizability of these perceptions. Therefore, a second, quantitative study is required to empirically test the model, validate the statistical significance of the 12 hypothesized paths, and determine the model's overall explanatory power. We now proceed to this validation in Study 2.

5. Study 2: quantitative validation

This second, quantitative stage of our research empirically tests the framework and hypotheses developed in Study 1. This validation is the essential next step in our sequential mixed-methods design. While Study 1 identified what traits users perceive as empathetic, Study 2 is designed to test how and how much these traits statistically predict perceived empathic reactivity. This step is crucial for moving our model from a qualitative proposal to an empirically-supported explanatory framework, aiming to (1) validate the measurement items for the four new constructs (AL, PE, EE, PA) and (2) empirically test the 12 hypothesized relationships between these traits and the dimensions of perceived empathic reactivity (PT, FT, EC, PD) using structural equation modeling.

5.1. Study 2 methodology

5.1.1. Data collection and participants

To verify the research model and test the 12 hypotheses, a new quantitative survey was conducted. The data was collected over a 10-day period using a mobile survey platform to ensure a diverse and timely sample of active users. While Study 1's framework was derived

Table 6
Operationalized definitions of IRI constructs in the CA context.

Construct	Definition	Exemplary Behaviors in CA Interaction
Perspective-taking	The extent to which the agent appears to understand the user's perspectives and needs.	Reflecting the user's previous state or informational query in subsequent recommendations.
Fantasy	The extent to which the agent interacts imaginatively to provide new and creative responses.	Offering witty jokes or crafting humorous stories that align with the user's emotional tone.
Empathic Concern	The extent to which the agent provides user-oriented care and concern through its responses.	Detecting emotional shifts and offering comforting content, such as soothing music.
Personal Distress	The extent to which the agent expresses negative feelings in high-pressure situations.	Conveying frustration or discomfort when facing offensive user behavior, potentially "fighting back."

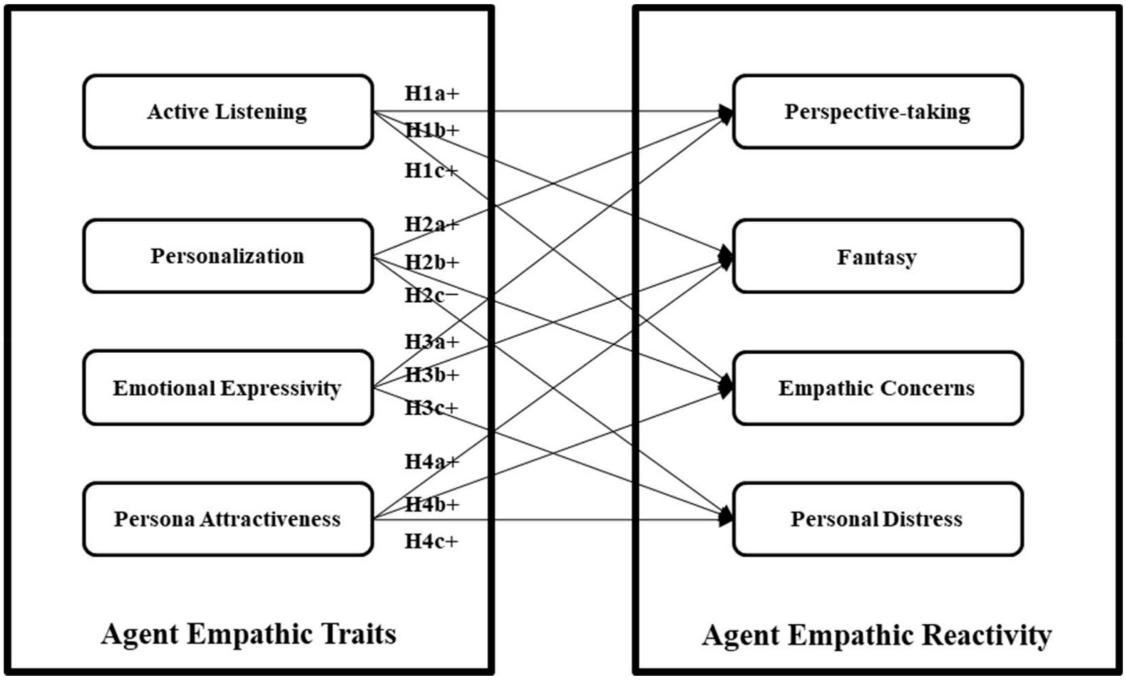

Fig. 1. Research model and hypotheses.

from a mixed U.S. and South Korea sample, this quantitative validation was conducted with participants recruited exclusively from South Korea. This approach provides a robust, first-time validation of the model within a specific, culturally-homogeneous context. This decision allows for a deep, initial validation of the framework's core relationships, while acknowledging the methodological mismatch as a key area for future cross-cultural replication, as discussed in the Limitations section.

The survey targeted users with at least one month of experience using social chatbots (e.g., "ChatGPT", "Simsimi", "Ruda", "Replika"). This target population was specifically chosen because such platforms are explicitly designed for the socio-emotional and relational interactions that are central to our framework, providing a rich and highly relevant context to test perceptions of empathy. Based on G*power 3.1.9.7 (effect size f^2=0.15, *alpha*=0.05, *power*=0.99, 4 predictors), a minimum of 174 samples was required (Hair et al., 2017). We successfully collected 200 valid samples, providing adequate statistical power for our model. The detailed demographic characteristics of the participants are summarized in Table 7.

5.1.2. Measurements

The measurement items for all eight constructs were developed by adapting validated questionnaires from extant studies, ensuring a foundation of established research. The scales for the four independent variables—the agent empathic traits (AL, PE, EE, PA)—were developed based on the operationalization from our Study 1 framework and the literature cited therein. The scales for the four dependent variables—the empathic reactivity dimensions (PT, FT, EC, PD)—were adapted from the CA-specific IRI measure (the AERI) validated by Lee and Yi (2023). As justified in our Related Work, the AERI provides a robust, multidimensional measure that is theoretically grounded in the IRI but operationally validated for the specific context of human-agent interaction.

A total of 32 items were used to measure the 8 constructs (as shown in Table 8). All measurement items were translated from English into Korean and then back-translated by an independent bilingual expert to ensure semantic equivalence. All constructs were measured using a 5-point Likert scale (1=strongly disagree and 5=strongly agree).

5.2. Data analysis

Measure validation and model testing were conducted using SmartPLS 4 (version 4.1.0; SmartPLS GmbH), a SEM tool that utilizes a component-based, partial least squares estimation approach. We chose component-based PLS-SEM over covariance-based SEM (CB-SEM) for clear methodological reasons. Our model, while grounded in Study 1 and the literature, is new and is being tested for the first time. Therefore, our primary goal is exploratory and predictive (i.e., to determine how much variance in empathy our four traits can explain), which is the primary strength of PLS-SEM. This contrasts with CB-SEM, which is designed for the strict confirmation of well-established, pre-existing models (Hair et al., 2017).

Because our goal is prediction and theory development rather than strict model fit confirmation, standard global fit indices (e.g., SRMR, NFI) are considered unreliable and less meaningful for our PLS-SEM

Table 7
Demographic distribution of Study 2 participants (N=200).

Factors	Distribution, n (%)
Sex	
Female	130 (65.0)
Male	70 (35.0)
Age Group (in years)	
Under 20	29 (14.5)
20–29	70 (35.0)
30–39	53 (26.5)
40–49	26 (13.0)
50 and above	22 (11.0)
Chatbot Usage (period)	
At least 1 month	119 (59.5)
At least 3 months	36 (18.0)
At least 6 months	29 (14.5)
1 year and above	16 (8.0)
Chatbot Platform	
ChatGPT	123 (61.5)
Simsimi	43 (21.5)
Ruda	17 (8.5)
Replika	4 (2.0)
Blender	2 (1.0)
Others (A., Siri etc.)	11 (5.5)

approach (Hair et al., 2017). Our analysis therefore focuses on the appropriate PLS-SEM evaluation criteria: individual parameter estimates (VIF, factor loadings), measurement model validity (AVE, CR), and the structural model's predictive accuracy (R^2) and path significance (t-values).

The analysis followed the standard two-step process: (1) rigorous assessment of the measurement model and (2) assessment of the structural model. Hypothesis testing was conducted using a bootstrapping procedure (5,000 resamples) to generate t-values and assess the statistical significance of each path coefficient.

5.3. Study 2 results

5.3.1. Measurement model assessment

Measurement scale reliability and validity were tested, and full diagnostics are presented. Convergent validity was confirmed as the Average Variance Extracted (AVEs) were all 0.50 or higher, Cronbach's Alpha was 0.70 or higher, and Composite reliability (CR) was 0.70 or higher (Henseler et al., 2009). The strong CR and AVE values confirm that the measurement items for each construct (e.g., the four items for Active Listening) are indeed measuring the same underlying concept and are distinct from measurement error. These results are shown in Table 9.

Discriminant validity—the degree to which constructs are statistically distinct from one another—was primarily assessed using a cross-loading analysis. We computed the cross-loadings by calculating the correlations between latent variable component scores and the manifest indicators of all constructs (Chin, 1998). For strong convergent and discriminant validity, loadings should be high (ideally exceeding 0.70), but loadings between 0.40 and 0.70 are considered acceptable if the construct's CR and AVE are high, which was the case for our model (Hair et al., 2017). As shown in the full cross-loadings table in Appendix B, all items loaded most strongly on their intended construct, and no item loaded more highly on a construct it was not intended to measure, providing robust evidence of discriminant validity. This evidence is critical, as it confirms, for example, that Active Listening (AL) and Personalization (PE) are statistically discriminable constructs, despite their high conceptual correlation. Additionally, discriminant validity was confirmed using the traditional Fornell-Larcker criterion (Fornell and Larcker, 1981). As shown in Table 10, the square root of the AVE for each construct (the bold diagonal) was higher than its correlation with any other construct in the model.

5.3.2. Structural model assessment

Following the validation of the measurement model, the structural model was assessed. In PLS-SEM, this evaluation focuses on (1) assessing collinearity, (2) the model's explanatory power (R^2), and (3) the significance of its path coefficients (β, t-values) and their effect size (f^2) (Hair et al., 2017).

First, it is crucial to assess potential collinearity among the predictor constructs. We calculated the Variance Inflation Factor (VIF) for all predictors. All VIF values were well below the common threshold of 5 (the highest being 3.31), indicating that multicollinearity was not a concern (Hair et al., 2017).

With collinearity ruled out, we examined the model's explanatory power. The R^2 values were 0.60 (Perspective-taking), 0.43 (Fantasy), 0.54 (Empathic Concern), and 0.30 (Personal Distress). All R^2 values indicate moderate to substantial explanatory power for an exploratory model, demonstrating that our four proposed traits account for a significant portion of the variance in perceived empathy. The path analysis (see Fig. 2 and Table 11), based on a 5,000-resample bootstrapping procedure, supported all 12 hypothesized paths. The f^2 effect sizes were also calculated to assess the substantive impact of each predictor, with 0.02, 0.15, and 0.35 considered small, medium, and large effects, respectively.

As shown in Table 11, all hypothesized paths were statistically significant. AL was a strong, positive predictor of PT, FT, and EC. PE strongly and positively predicted PT and EC, and (as hypothesized) was the only trait to negatively and significantly predict PD. EE positively influenced PT, FT, and PD. Finally, PA was a significant predictor for FT, EC, and PD. These results provide strong initial quantitative support for

Table 8
Measurement items utilized in Study 2.

Construct	Item	Reference
Active Listening	<agent> restated my input to convey understanding. <agent> seemed to understand my emotions well, reflecting my emotions in its responses. <agent> responded actively and we had a smooth conversation. <agent> demonstrated an attitude of listening or responding to my reaction.	Xiao et al. (2020)
Personalization	<agent> responded based on my moods and demands <agent> offered services that matched my demands and the situation. Responses provided by the <agent> were customized to my needs. Through the interaction with <agent>, I received services that were tailored to my issues.	Zhu et al. (2022)
Emotional Expressivity	<agent> had the ability to express rich emotions. It was easy to understand the emotions <agent> expressed. <agent> had an excellent ability to convey emotions. <agent> tended to reveal its emotions well.	Kring et al. (1994)
Persona Attractiveness	The presented persona of <agent> was attractive. I found the character of <agent> fascinating. The design of <agent> was appealing. Overall, I found the <agent> good-looking and attractive.	Merikivi et al. (2017)
Perspective-taking	<agent>'s reaction was appropriate from my perspective. I found that <agent> sees things as I do. I like how <agent> sees things from my perspective. <agent> seemed to understand my needs and reflected them in its responses.	Lee and Yi (2023)
Fantasy	<agent> interacted imaginatively like a real person. <agent> sometimes showed creative responses. <agent> responded to my problems with new and unexpected solutions. <agent> sometimes surprised me with its unconventional and interesting reactions.	
Empathic Concern	<agent> showed an attitude of caring and protecting the user. <agent> has a friendly heart toward the user. <agent> revealed kindness in its interaction with the user. <agent> listened to my words carefully and was being supportive.	
Personal Distress	<agent> reacted with intense emotion as the situation became troublesome. <agent> became emotionally unstable when the user gave pressure. <agent> expressed its displeasure toward the user's harassment or offensive reaction. <agent> showed emotional distress when it encountered a troubling situation.	

Note. <agent> was replaced with the name of the chatbot during the survey.

Table 9
Results of the assessment of reliability and convergent validity (N=200).

	Mean	Cronbach's alpha	Composite Reliability	Average Variance Extracted
Active Listening	3.51	0.71	0.82	0.53
Personalization	3.53	0.79	0.87	0.62
Emotional Expressivity	3.14	0.90	0.93	0.76
Persona Attractiveness	3.27	0.83	0.89	0.66
Perspective-taking	3.47	0.84	0.89	0.68
Fantasy	3.38	0.83	0.88	0.66
Empathic Concern	3.67	0.88	0.92	0.73
Personal Distress	2.68	0.83	0.89	0.66

the framework derived from Study 1 in our South Korean sample. The findings clearly support our central thesis: the functional traits (AL, PE) have strong, positive impacts on core empathy dimensions (PT, EC) and can even mitigate distress (PD), while the socio-affective traits (EE, PA) show a more complex, dual-edged influence, boosting imaginative and social connection (FT, EC) but also perceived agent vulnerability (PD). A full interpretation of these findings is presented in the General Discussion. The overall results of the path analysis, illustrating the relationships and explained variances among the constructs, are visually depicted in Fig. 2.

6. General discussion

This research provides a new, empirically-grounded perspective on the perceived empathy of conversational agents, moving the field beyond isolated feature testing. By employing a sequential exploratory mixed-methods design, this study bridges a critical gap between high-level psychological theories of empathy and the tangible, lived experience of users. Our research moves beyond prior works that primarily emphasize isolated aspects of empathy or rely on ad-hoc, non-validated measures (Yalçın, 2020; Paiva et al., 2017). We achieve this by first qualitatively identifying the specific behaviors users articulate as empathetic (Study 1) and then systematically testing how these empirically-grounded traits predict empathic perceptions (Study 2).

Our study's contributions are twofold, demonstrating the value of this sequential approach. First, Study 1 conducted a qualitative investigation (N=166, U.S. and South Korea) using a hybrid computational-human thematic analysis. This method allowed us to derive the core behaviors users associate with agent empathy (e.g., "Consulting with user troubles," "Remembering interaction context"), grounding our work in authentic user-defined concepts. These empirical findings were then synthesized with established empathy theory (Batson, 2009) to propose a new, four-factor framework of agent empathic traits: Active Listening (AL), Personalization (PE), Emotional Expressivity (EE), and Persona Attractiveness (PA).

Second, Study 2 provided an initial quantitative validation of this framework in a specific cultural context (N=200, South Korea). Using

Table 11
Path coefficient and statistical power of the research model (N=200).

Hypothesized Path	β	t-Value	f^2	Hypothesis test result
H1a) AL to PT	0.29	4.23	0.11	Supported
H1b) AL to FT	0.33	4.52	0.13	Supported
H1c) AL to EC	0.33	4.96	0.13	Supported
H2a) PE to PT	0.40	6.16	0.23	Supported
H2b) PE to EC	0.30	4.05	0.11	Supported
H2c) PE to PD	-0.17	2.28	0.03	Supported
H3a) EE to PT	0.22	3.32	0.09	Supported
H3b) EE to FT	0.25	3.17	0.06	Supported
H3c) EE to PD	0.47	5.42	0.18	Supported
H4a) PA to FT	0.21	2.77	0.04	Supported
H4b) PA to EC	0.23	4.03	0.08	Supported
H4c) PA to PD	0.22	2.44	0.04	Supported

Note for abbreviations. AL: Active listening; PE: Personalization; EE: Emotional expressivity; PA: Persona attractiveness; PT: Perspective-taking; FT: Fantasy; EC: Empathic concern; PD: Personal distress

structural equation modeling, we confirmed that these four traits are not only valid constructs but are also strong, statistically significant predictors of how users perceive an agent's empathic reactivity. The model demonstrated substantial explanatory power over the four dimensions of the adapted IRI (Perspective-taking, Fantasy, Empathic Concern, and Personal Distress), confirming the traits' relevance and impact.

6.1. Theoretical implications

Our findings provide a much clearer, more structured picture of how users construct their perception of empathy. It is not a single, monolithic feeling, but rather a complex, multi-layered judgment. This judgment appears to balance what the agent does (its functional competence) with how the agent acts (its perceived warmth and social character). Our model's four traits map directly onto this dichotomy: Active Listening (AL) and Personalization (PE) represent the core functional competence—the agent's demonstrated ability to understand and act appropriately. In contrast, Emotional Expressivity (EE) and Persona Attractiveness (PA) constitute the socio-affective character—how the agent presents itself and its actions within a social context.

Our results, which show distinct and sometimes conflicting effects, provide significant theoretical nuance. This overall synthesized view is elaborated below:

1. Active Listening (AL) acts as the foundation or "gateway" to empathy. It is the prerequisite trait that proves the agent is paying attention and correctly processing the user's information, both explicit and implicit. Our qualitative data showed users felt "understood" when an agent recalled past details. The strong quantitative paths from AL to PT, FT, and EC confirm that this cognitive validation is the necessary first step for a user to even consider an agent as a candidate for an empathetic interaction. Without this foundation, any subsequent attempts at empathy (like PE or EE) risk feeling hollow, generic, or even misplaced, thus failing to foster positive engagement.

Table 10
Results of the assessment based on the Fornell-Larcker criterion (N=200).

	AL	PE	EE	PA	PT	FT	EC	PD
Active Listening	**0.73**							
Personalization	0.64	**0.79**						
Emotional Expressivity	0.52	0.48	**0.87**					
Persona Attractiveness	0.52	0.53	0.62	**0.81**				
Perspective-taking	0.67	0.70	0.57	0.57	**0.82**			
Fantasy	0.56	0.44	0.55	0.53	0.64	**0.81**		
Empathic Concerns	0.65	0.64	0.49	0.57	0.79	0.61	**0.86**	
Personal Distress	0.27	0.17	0.52	0.41	0.31	0.38	0.18	**0.82**

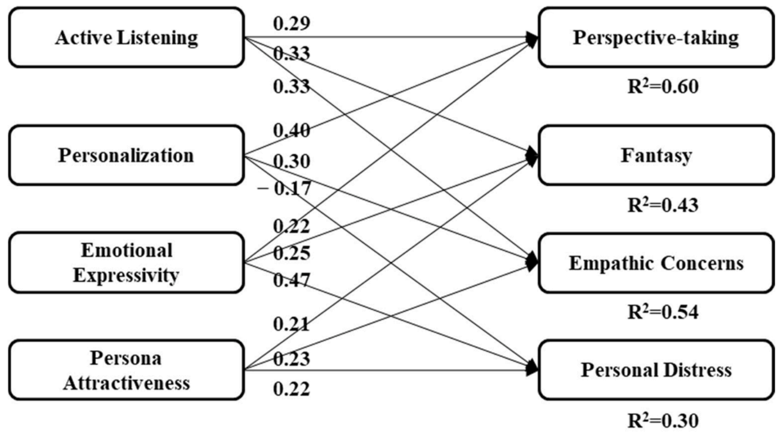

Fig. 2. Results of the path analysis of the research model (N=200).

2. Personalization (PE) is the proactive component of competence. It is the functional proof that the agent is using its understanding (from AL) to act in the user's best interest. By tailoring its actions, recalling preferences, and providing relevant, non-generic responses, the agent signals that it is not just a passive listener but a helpful, user-oriented participant. This trait powerfully reinforces Empathic Concern (perceived care). Crucially, its unique status as the only negative predictor of Personal Distress suggests that PE also builds a sense of reliability and perceived control that the user trusts that this competent agent will not "fail" or cause stress, thereby mitigating the user's own anxiety about the interaction.
3. Emotional Expressivity (EE) and Persona Attractiveness (PA) are social multipliers. But they are a double-edged sword. These traits are what make the interaction feel rich, engaging, creative (Fantasy), and human-like. However, their strong positive prediction of Personal Distress (the agent's own perceived anxiety) is a salient and critical finding. It suggests a potential 'uncanny valley' of agent empathy. When an agent's persona is too human or its emotional language is too intense (e.g., overly apologetic, seemingly frustrated), users may perceive this not as empathy for them, but as the agent's own anxiety or instability. This can break the user's immersion, cause discomfort, or reduce the agent's perceived reliability.

Taken together, these findings challenge the simplistic, linear assumption that "more human is always better." It offers a more nuanced, theoretically grounded framework explaining how distinct empathic traits collectively and interactively shape user perceptions, refining the conceptual understanding of empathy in CAs.

6.2. Practical implications for design

While our study is exploratory and not a prescriptive guide, our findings offer specific, data-driven considerations for designers. The challenge is not simply to "add" empathy, but to balance these four traits. Our results suggest that implementation requires a careful, orchestrated combination of elements. For instance, the strong, consistently positive paths from Active Listening (AL) and Personalization (PE) to core empathy dimensions (PT, EC) confirm their foundational role within the tested model. Conversely, the nuanced and strong positive paths from Emotional Expressivity (EE) and Persona Attractiveness (PA) to Personal Distress (H3c, H4c) serve as a clear, data-driven warning against naive human-likeness. Our practical considerations are thus based on balancing these findings:

1. To enhance Active Listening, designers must prioritize robust context-awareness and memory, enabling agents to reference past user inputs, recall preferences, and note changes in emotional states (Lee et al., 2021; Taylor et al., 2019). As our findings imply, AL is the foundation of trust. Without it, an agent offering an "empathetic" phrase about a topic the user just discussed will feel hollow, incompetent, and ultimately damage user trust more than if it had said nothing at all.
2. To improve Personalization, agents should adapt message content and style to users' individual needs and goals, moving beyond generic assistance (Chin and Yi, 2021; Pickard and Roster, 2020). A generic, 'one-size-fits-all' response, as our qualitative data (code: "Low quality responses") showed, is often perceived as a lack of care and a sign of 'inhumanness'. This user-centric tailoring is what reinforces perceived care and supportive behavior (Liu and Sundar, 2018; Loveys et al., 2022). This trait is the functional proof of empathy.
3. To manage Emotional Expressivity, designers should be cautious and prioritize calibration. This advice stems directly from our finding that EE strongly predicted Personal Distress (H3c). While empathic language (e.g., apologies, encouragement, affirmations) is beneficial (Beale and Creed, 2009; Loveys et al., 2021), overt or performative emotionality can backfire. Moderate, calibrated verbal cues (e.g., "That sounds frustrating," rather than "Oh no, that is terrible and makes me so sad!") may be more effective. The goal should be to signal understanding of the user's emotion, not to perform a high-intensity emotion for the agent's own sake.
4. To optimize Persona Attractiveness, a consistent and appealing interaction style is key (Chin et al., 2020; MacDorman, 2019). However, our finding that PA also predicted Personal Distress (H4c) suggests moderation is crucial. A persona that is too human-like may set expectations that the technology cannot meet, leading to a greater sense of disappointment when it inevitably fails. This perception of emotional vulnerability could make users uncomfortable or less likely to trust the agent with sensitive tasks. An agent that is 'pleasantly' an agent, rather than one 'imperfectly' human, may be the more stable design.

6.3. Limitations and future research

Despite the contributions of this study, several limitations must be acknowledged, which in turn offer clear directions for future research.

First, this study faces a methodological limitation regarding its sampling and cultural context. Study 1 derived its framework from qualitative data collected in two cultural contexts (the U.S. and South Korea), whereas Study 2 tested the model using a South Korean sample only. Accordingly, the current quantitative evidence should be considered culturally contingent at this stage and interpreted as initial support within one cultural context rather than as cross-culturally generalizable.

The themes identified in Study 1—and the relative salience of traits (e.g., emotional expressivity vs. functional personalization)—may vary across cultures depending on communication styles and expectations about appropriate emotional expression. Therefore, future research should conduct cross-cultural replication and formal measurement invariance testing before making broader generalizability claims. Following recent cross-cultural validation work on partner model questionnaires for voice interfaces (Seaborn et al., 2024), future studies should also examine the cross-cultural robustness of both our empathic framework and the AERI-based measures across different language and cultural groups, with explicit attention to measurement invariance and culturally situated interpretations.

Second, our findings rely on self-reported user perceptions of empathy. Self-reports are, by nature, retrospective and subject to individual biases (e.g., social desirability, recall bias). For instance, a user's current mood or general disposition toward AI may have colored their recollection of past events. This study does not test an actual, controlled CA system. Future work should complement our findings by using behavioral experiments (e.g., A/B testing of agents high/low in these four traits) to measure how these traits affect actual user behavior (like task completion time or self-disclosure depth) in real-time.

Third, as a related point, our data is cross-sectional. We captured user perceptions at a single point in time, which does not account for how these perceptions might evolve as the novelty of an agent wears off or as trust is built (or broken). Empathy and trust are often built over time through repeated interactions and trust-repair mechanisms. Future longitudinal studies are needed to understand how these four traits contribute to the development, maintenance, and potential decay of a long-term human-agent relationship.

Fourth, our qualitative findings in Study 1 are based on the interpretation of textual survey answers. This method, while rich, may be subject to researcher bias and cannot capture the full nuance of user experience. This method particularly misses the paralinguistic cues (sighs, hesitations, tone of voice) that a verbal interview might capture, which are often powerful indicators of a user's true affective state.

Fifth, our study’s scope was limited to text-based interactions. As our Related Work section noted, empathy is often conveyed multimodally through powerful non-verbal cues (e.g., voice tone, facial expressions, gestures). Our focus on text provides a clear boundary but means our findings may not generalize to embodied agents or voice assistants. The added layers of vocal prosody or facial mimicry are known to be powerful channels for empathy and would likely alter the relationships and their strengths observed in our model.

Sixth, our framework was derived from general-purpose CAs (e.g., ChatGPT, Siri). The importance and optimal balance of these four traits may differ significantly in high-stakes or specialized contexts, such as healthcare, education, or crisis counselling. In a healthcare context, for example, personalization (PE) and active listening (AL) may be paramount, while persona attractiveness (PA) could be seen as frivolous or even inappropriate. The “right” mix of these traits is likely context-dependent.

Finally, our sample in Study 1, while diverse in platform use, was not specifically recruited or powered to analyze differences across age or other demographic characteristics, and our demographic items used binary sex categories. The perception of what counts as “attractive” or “appropriate” emotional expression is highly subjective and may differ not only across age groups but also across gender identities. Future research should both provide more inclusive demographic options (e.g., non-binary, trans, and other diverse identities) and be explicitly designed to compare how different cohorts—for example, older adults, teenagers, and users with diverse gender identities—perceive and evaluate these empathic traits, as they may have distinct mental models of CAs and varying needs for social connection.

7. Conclusion

This study advances the understanding of empathy in conversational agents by providing a theoretically grounded and empirically-tested framework. Through a sequential exploratory mixed-methods design, we moved beyond ad-hoc feature testing to first identify what users perceive as empathy (Study 1) and then quantitatively test how those perceptions are structured (Study 2). We identified four critical agent traits—Active Listening, Personalization, Emotional Expressivity, and Persona Attractiveness—that collectively shape how users perceive an agent's cognitive and affective empathy.

Our findings move beyond prior work’s reliance on isolated cues by elucidating the complex interactions between these traits. We particularly highlight the delicate balance between functional competence (AL, PE) and socio-emotional character (EE, PA), revealing that the most human-like traits can also be a source of user-perceived distress. This finding complicates the naive pursuit of mere human-likeness and presents a novel, data-driven challenge for HCI researchers.

Theoretically, this research refines and extends empathy frameworks into the domain of human-agent interaction by offering a new, initially supported model that is empirically derived from users (Study 1) and grounded in established psychological theory (e.g., IRI and CASA). Practically, it offers actionable considerations for designers, shifting the goal from simply adding empathy to strategically balancing functional attentiveness with socio-affective expression to build authentic, trustworthy agents. By addressing users’ nuanced emotional and cognitive needs, this work lays the groundwork for future explorations into long-term, multimodal, and context-dependent empathic interactions that can foster more meaningful and supportive user-agent relationships.

CRediT authorship contribution statement

Bumho Lee: Writing – review & editing, Writing – original draft, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Youngsoo Shin:** Writing – review & editing, Methodology, Investigation. **Byounghyun Yoo:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

There are no conflicts of interest regarding this article.

Acknowledgement

This work was supported by the ETRI grant (No. 25ZF1100, Research on technology policy and standardization for national intelligence), the National Research Foundation (NRF) grant (No. NRF-2021R1A2C2093065), and the KIST under the Institutional Program (Grant No. 26E0211), all funded by the Korean government (MSIT).

Appendix A: Study 1 Survey Questions (N=166)

The following questions were presented to participants in Study 1, translated into their respective languages (English or Korean). This survey was designed to first capture a profile of the participant's usage habits and then elicit rich, qualitative narratives about their lived experiences with CA empathy.

Part 1: Demographic and Usage Questions

(These questions correspond to the data presented in Table 2. They were designed to profile the sample and ensure our qualitative data was drawn from a diverse set of users with varied backgrounds, ages, and levels of experience with the most common CA platforms.)

1. **What is your age?**
   o (Dropdown menu of age ranges)
2. **What is your gender?**
   o Male
   o Female
3. **What is the primary Conversational Agent (CA) you use most often?**
   o ChatGPT
   o IPA (Siri, Alexa, Google Assistant, etc.)
   o Simsimi
   o Ruda
   o Other:
4. **How long have you used this primary CA?**
   o At least 1 month
   o At least 3 months
   o At least 6 months
   o 1 year and above
5. **How frequently do you use this primary CA?**
   o More than 3 times a day
   o At least once a day
   o At least once in three days
   o At least once a week
   o At least once in two weeks

Part 2: Open-Ended Qualitative Questions

(These six questions formed the core of our qualitative data collection. They were specifically designed to move beyond simple "yes/no" answers and elicit rich, detailed narratives. Each question targets a potential facet of perceived empathy, which were used for the Thematic Analysis presented in Section 4.2.)

1. **AI챗봇을 사용 중에 챗봇이 참여자님이 이야기 하는 내용을 잘 파악하거나 공감하고 있다고 생각한 적이 있다면 해당 경험을 적어 주세요.** (If you have ever thought that the chatbot understood or empathized with what you were saying while using it, please describe that experience.)
2. **AI챗봇을 사용 중에 챗봇이 감정적으로 대응하거나 대답을 회피하는 경우 어떤 생각이 드셨는지 적어주세요.** (What did you think when the chatbot responded emotionally or avoided answering while you were using it?)
3. **AI챗봇을 사용 중에 챗봇에 반응에 특히 감탄하거나 놀랐던 적이 있었다면 어떤 때 그렇게 느꼈는지 적어주세요.** (If you were particularly impressed or surprised by the chatbot's response while using it, please describe when you felt that way.)
4. **AI챗봇을 사용 중에 챗봇이 참여자님이나 참여자님에게 관심을 보이고 잘 챙기는 듯 하다고 느끼셨던 적이 있었다면 어떤 때 그렇게 느꼈는지 적어주세요.** (If you have ever felt that the chatbot was showing interest in you or taking good care of you, please describe when you felt that way.)
5. **AI챗봇은 사람의 감정을 잘 이해하고 이에 맞추어 더 나은 서비스를 제공하기 위해 발전하고 있습니다. 이러한 경향에 대해 개인적으로 느끼셨거나 의견사항이 있으면 적어주세요.** (AI chatbots are evolving to better understand human emotions and provide better services accordingly. Please write any personal feelings or opinions you have about this trend.)
6. **AI챗봇을 이용하면서 가장 불만족스러웠던 점은 어떤 것이었나요?** (What was the most dissatisfying thing about using the AI chatbot?)

*__Note:__ The participants were requested to answer '해당 없음' (Not Applicable), if they did not have such experience or remember it well.

Part 3: Additional Quantitative Scales (Collected but Not Used in Analysis)

(The following 7-point semantic differential scales and 5-point Likert scales were also included in the Study 1 survey. However, to maintain research parsimony, they were not used in this manuscript's analysis. The primary goal of Study 1 was ***exploratory****—to derive new constructs from qualitative, open-ended data—rather than* ***confirmatory****—to test pre-existing quantitative scales. Including these scales would have unnecessarily complicated the narrative of our exploratory first study. They are presented here for full methodological transparency.)*

5-point Likert Scales (1=Strongly Disagree, 5=Strongly Agree)

User Enjoyment

- It was fun and enjoyable to share a conversation with the <agent>.
- The conversation with the <agent> was exciting.
- Services provided by the <agent> were entertaining.
- I enjoyed the services provided by <agent>.

Agent Likeability

- <agent> was likable.
- <agent> was friendly.
- <agent> was kind.
- <agent> was pleasant.
- <agent> was nice.

User Engagement

- I felt the time was going fast when I was sharing a conversation with <agent>.
- I felt I was well involved in interaction and conversation with <agent>.
- I could concentrate on what <agent> was saying.
- I was fully engaged in the interaction with <agent>.

User Satisfaction

- I was satisfied with the experience of using a dialogue with <agent> to complete tasks.
- I am satisfied with the <agent> service.
- Interacting with <agent> was a pleasant and satisfactory experience.
- I really enjoyed and was satisfied with the reactions of <agent>.
- The overall assessment of conversing with <agent> was satisfactory.

Appendix B: Study 2 VIF and Cross-loadings

	VIF	AL	PE	EE	PA	PT	FT	EC	PD
AL1	1.27	**0.67**	0.33	0.22	0.32	0.40	0.33	0.44	0.17
AL2	1.26	**0.72**	0.51	0.57	0.48	0.55	0.42	0.50	0.29
AL3	1.40	**0.72**	0.49	0.26	0.30	0.39	0.42	0.38	0.09
AL4	1.52	**0.81**	0.53	0.41	0.39	0.58	0.47	0.55	0.19
PE1	1.25	0.59	**0.69**	0.56	0.49	0.53	0.40	0.51	0.31
PE2	1.74	0.37	**0.79**	0.21	0.29	0.53	0.29	0.47	-0.08
PE3	1.91	0.52	**0.83**	0.37	0.44	0.56	0.37	0.54	0.13
PE4	1.95	0.53	**0.83**	0.32	0.45	0.57	0.32	0.48	0.12
EE1	2.64	0.46	0.42	**0.88**	0.57	0.49	0.51	0.44	0.41
EE2	2.48	0.46	0.49	**0.86**	0.57	0.54	0.48	0.45	0.36
EE3	3.15	0.50	0.43	**0.91**	0.54	0.53	0.50	0.47	0.48
EE4	2.11	0.38	0.31	**0.84**	0.49	0.40	0.43	0.35	0.56
PA1	2.33	0.44	0.43	0.52	**0.85**	0.43	0.38	0.43	0.36
PA2	2.25	0.42	0.44	0.61	**0.84**	0.50	0.43	0.43	0.46
PA3	2.01	0.36	0.36	0.52	**0.85**	0.43	0.41	0.47	0.35
PA4	1.38	0.47	0.49	0.37	**0.71**	0.49	0.50	0.51	0.18
PT1	1.60	0.55	0.62	0.43	0.46	**0.78**	0.52	0.62	0.19
PT2	2.08	0.54	0.58	0.55	0.52	**0.84**	0.57	0.65	0.36
PT3	2.03	0.59	0.51	0.47	0.48	**0.84**	0.51	0.64	0.30
PT4	1.86	0.52	0.59	0.41	0.42	**0.83**	0.49	0.70	0.16
FT1	2.25	0.40	0.30	0.52	0.43	0.49	**0.81**	0.44	0.37
FT2	2.62	0.44	0.31	0.51	0.45	0.54	**0.87**	0.50	0.36
FT3	1.62	0.46	0.40	0.31	0.35	0.50	**0.78**	0.42	0.22
FT4	1.53	0.52	0.42	0.43	0.48	0.53	**0.78**	0.58	0.29
EC1	1.60	0.53	0.53	0.55	0.50	0.70	0.50	**0.76**	0.34
EC2	2.43	0.51	0.56	0.32	0.43	0.63	0.46	**0.86**	0.04
EC3	3.31	0.57	0.56	0.41	0.48	0.67	0.55	**0.90**	0.09
EC4	2.89	0.60	0.55	0.42	0.52	0.71	0.56	**0.89**	0.16
PD1	1.68	0.24	0.21	0.53	0.44	0.31	0.37	0.23	**0.82**
PD2	2.63	0.22	0.14	0.43	0.36	0.23	0.29	0.11	**0.88**
PD3	1.49	0.24	0.10	0.31	0.21	0.24	0.34	0.20	**0.70**
PD4	2.47	0.14	0.05	0.36	0.28	0.19	0.24	0.03	**0.84**

Note for abbreviations. AL: Active listening; PE: Personalization; EE: Emotional expressivity; PA: Persona attractiveness; PT: Perspective-taking; FT: Fantasy; EC: Empathic concern; PD: Personal distress.

Data availability

Data will be made available on request.

References

Abbasian, M., Azimi, I., Feli, M., Rahmani, A. M., & Jain, R. (2024). Empathy through multimodality in conversational interfaces. *arXiv preprint arXiv:2405.04777*.

Adam, M., Wessel, M., Benlian, A., 2021. AI-based chatbots in customer service and their effects on user compliance. Electron. Mark. 31 (2), 427–445.

Agnihotri, A., Bhattacharya, S., 2024. Chatbots' effectiveness in service recovery. Int. J. Inf. Manag. 76, 102679.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. J. Autism. Dev. Disord. 34 (2), 163–175. https://doi.org/10.1023/b:jadd.0000022607.19833.00.

Batson, C.D., 2009. These things called empathy: eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), The Social Neuroscience of Empathy (pp. 3–15). Boston Review. https://doi.org/10.7551/mitpress/9780262012973.003.0002.

Beale, R., Creed, C., 2009. Affective interaction: how emotional agents affect users. Int. J. Hum.-Comput. Stud. 67 (9), 755–776.

Bickmore, T.W., Fernando, R., Ring, L., Schulman, D., 2010. Empathic touch by relational agents. IEEE Trans. Affect. Comput. 1 (1), 60–71.

Bickmore, T.W., Picard, R.W., 2005. Establishing and maintaining long-term human-computer relationships. ACM Trans. Comput.-Hum. Interact. (TOCHI) 12 (2), 293–327.

Bodie, G.D., 2011. The Active-Empathic Listening Scale (AELS): conceptualization and evidence of validity within the interpersonal domain. Commun. Q. 59 (3), 277–295.

Brave, S., Nass, C., Hutchinson, K., 2005. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int. J. Hum.-Comput. Stud. 62 (2), 161–178. https://doi.org/10.1016/j.ijhcs.2004.11.002.

Cao, Q., Cheng, X., Liao, S., 2023. A comparison study of topic modeling based literature analysis by using full texts and abstracts of scientific articles: a case of COVID-19 research. Libr. Hi Tech 41 (2), 543–569.

Carolus, A., Wienrich, C., Törke, A., Friedel, T., Schwietering, C., Sperzel, M., 2021. Alexa, I feel for you!'Observers' empathetic reactions towards a conversational agent. Front. Comput. Sci. 3, 682982.

Casas, J., Spring, T., Daher, K., Mugellini, E., Khaled, O.A., Cudré-Mauroux, P., 2021. Enhancing conversational agents with empathic abilities. In: Proceedings of the 21st

ACM International Conference on Intelligent Virtual Agents, Virtual Event, Japan. https://doi.org/10.1145/3472306.3478344.
Chaturvedi, R., Verma, S., Srivastava, V., 2023. Empowering AI companions for enhanced relationship marketing. Calif. Manag. Rev. 66 (2), 65–90. https://doi.org/10.1177/00081256231215838.
Chin, H., Molefi, L.W., Yi, M.Y., 2020. Empathy is all you need: how a conversational agent should respond to verbal abuse. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.
Chin, H., Yi, M.Y., 2021. Voices that care differently: understanding the effectiveness of a conversational agent with an alternative empathy orientation and emotional expressivity in mitigating verbal abuse. Int. J. Hum.–Comput. Interact. 38 (12), 1153–1167. https://doi.org/10.1080/10447318.2021.1987680.
Chin, W.W., 1998. The partial least squares approach to structural equation modeling. Mod. methods bus. res. 295 (2), 295–336.
Chrysikou, E.G., Thompson, W.J., 2016. Assessing cognitive and affective empathy through the interpersonal reactivity index: an argument against a two-factor model. Assessment 23 (6), 769–777. https://doi.org/10.1177/1073191115599055.
Cliffordson, C., 2001. Parents' judgments and students' self-judgments of empathy: the structure of empathy and agreement of judgments based on the interpersonal reactivity index (IRI). Eur. J. Psychol. Assess. 17 (1), 36.
Corti, K., Gillespie, A., 2016. Co-constructing intersubjectivity with artificial conversational agents: people are more likely to initiate repairs of misunderstandings with agents represented as human. Comput. Hum. Behav. 58, 431–442.
Cuadra, A., Wang, M., Stein, L.A., Jung, M.F., Dell, N., Estrin, D., Landay, J.A., 2024. The illusion of empathy? notes on displays of emotion in human-computer interaction. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA. https://doi.org/10.1145/3613904.3642336.
D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., Person, N., Kort, B., El Kaliouby, R., Picard, R., 2008. AutoTutor detects and responds to learners affective and cognitive states. In: Workshop on emotional and cognitive issues at the international conference on intelligent tutoring systems.
Davis, M.H., 1980. A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology 10, 85.
Davis, M.H., 1983. Measuring individual differences in empathy: evidence for a multidimensional approach. J. personal. soc. psychol. 44 (1), 113.
Davis, M.H., 2018. Empathy: A social psychological approach. Routledge.
de Gennaro, M., Krumhuber, E.G., Lucas, G., 2019. Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. Front Psychol 10, 3061. https://doi.org/10.3389/fpsyg.2019.03061.
Doyle, P.R., Clark, L., Cowan, B.R., 2021. What do we see in them? identifying dimensions of partner models for speech interfaces using a psycholexical approach. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1–14.
Efthymiou, F., Hildebrand, C., 2023. Empathy by design: the influence of trembling AI voices on prosocial behavior. IEEE Trans. Affect. Comput.
Egger, R., Yu, J., 2022. A topic modeling comparison between LDA, NMF, top2vec, and bertopic to demystify twitter posts. Front. sociol. 7, 886498.
Eklund, J.H., Meranius, M.S., 2021. Toward a consensus on the nature of empathy: a review of reviews. Patient Educ. Couns. 104 (2), 300–307.
El Azhari, K., Hilal, I., Daoudi, N., Ajhoun, R., 2023. SMART chatbots in the E-learning domain: a systematic literature review. Int. J. Interact. Mob. Technol. 17 (15).
Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. J. mark. res. 18 (1), 39–50.
Ha, Q.-A., Chen, J.V., Uy, H.U., Capistrano, E.P., 2021. Exploring the privacy concerns in using intelligent virtual assistants under perspectives of information sensitivity and anthropomorphism. Int. J. Hum.–Comput. Interact. 37 (6), 512–527.
Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M., 2017. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 2nd Edition. Sage Publications Inc., Thousand Oaks, CA.
Han, Z.-M., Huang, C.-Q., Yu, J.-H., Tsai, C.-C., 2021. Identifying patterns of epistemic emotions with respect to interactions in massive online open courses using deep learning and social network analysis. Comput. Hum. Behav. 122, 106843.
Henseler, J., Ringle, C.M., Sinkovics, R.R., 2009. The use of partial least squares path modeling in international marketing. New challenges to international marketing. Emerald Group Publishing Limited.
Hoenen, M., Lübke, K.T., Pause, B.M., 2016. Non-anthropomorphic robots as social entities on a neurophysiological level. Comput. Hum. Behav. 57, 182–186.
Hu, T., Xu, A., Liu, Z., You, Q., Guo, Y., Sinha, V., Luo, J., Akkiraju, R., 2018. Touch your heart: a tone-aware chatbot for customer care on social media. In: Proceedings of the 2018 CHI conference on human factors in computing systems.
Jeong, D.C., Lee, J., 2017. Snap back to reality: examining the cognitive mechanisms underlying Snapchat. Comput. Hum. Behav. 77, 274–281.
Kim, J., Kim, W., Nam, J., Song, H., 2020. I can feel your empathic voice. In: Effects of Nonverbal Vocal Cues in Voice User Interface. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems.
Kourmousi, N., Amanaki, E., Tzavara, C., Koutras, V., 2017. Active Listening Attitude Scale (ALAS): reliability and validity in a nationwide sample of Greek educators. Soc. Sci. 6 (1), 28.
Kraus, M., Seldschopf, P., Minker, W., 2021. Towards the development of a trustworthy chatbot for mental health applications. In: MultiMedia Modeling: 27th International Conference, MMM 2021, Prague, Czech Republic, June 22–24, 2021, Proceedings, Part II 27.
Kring, A.M., Smith, D.A., Neale, J.M., 1994. Individual differences in dispositional expressiveness: development and validation of the Emotional Expressivity Scale. J. personal. soc. psychol. 66 (5), 934.
Lee, B., Yi, M.Y., 2023. Understanding the empathetic reactivity of conversational agents: measure development and validation. Int. J. Hum.–Comput. Interact. 1–19.
Lee, S., Choi, J., 2017. Enhancing user experience with conversational agent for movie recommendation: effects of self-disclosure and reciprocity. Int. J. Hum.-Comput. Stud. 103, 95–105.
Lee, S.K., Kavya, P., Lasser, S.C., 2021. Social interactions and relationships with an intelligent virtual agent. Int. J. Hum.-Comput. Stud. 150, 102608.
Leite, I., Martinho, C., Paiva, A., 2013a. Social robots for long-term interaction: a survey. Int. J. Soc. Robot. 5 (2), 291–308. https://doi.org/10.1007/s12369-013-0178-y.
Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R., Paiva, A., 2013b. The influence of empathy in human-robot relations. Int. J. Hum.-Comput. Stud. 71 (3), 250–260. https://doi.org/10.1016/j.ijhcs.2012.09.005.
Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A., Bhogal, R.S., 1997. The persona effect: affective impact of animated pedagogical agents. In: Proceedings of the ACM SIGCHI Conference on Human factors in computing systems.
Liang, K.-H., Shi, W., Oh, Y.J., Wang, H.-C., Zhang, J., Yu, Z., 2024. Dialoging resonance in human-chatbot conversation: how users perceive and reciprocate recommendation chatbot's self-disclosure strategy. Proc. ACM Hum.-Comput. Interact. 8 (CSCW1), 1–28.
Liao, X., Zheng, Y.-H., Shi, G., Bu, H., 2024. Automated social presence in artificial-intelligence services: conceptualization, scale development, and validation. Technol. Forecast. Soc. Change 203, 123377.
Lim, W.M., Kumar, S., Verma, S., Chaturvedi, R., 2022. Alexa, what do we know about conversational commerce? Insights from a systematic literature review. Psychol. mark. 39 (6), 1129–1155.
Liu-Thompkins, Y., Okazaki, S., Li, H., 2022. Artificial empathy in marketing interactions: bridging the human-AI gap in affective and social customer experience. J. Acad. Mark. Sci. 1–21.
Liu, B., Sundar, S.S., 2018. Should machines express sympathy and empathy? experiments with a health advice chatbot. Cyberpsychol Behav Soc Netw 21 (10), 625–636. https://doi.org/10.1089/cyber.2018.0110.
Loveys, K., Hiko, C., Sagar, M., Zhang, X., Broadbent, E., 2022. I felt her company": a qualitative study on factors affecting closeness and emotional support seeking with an embodied conversational agent. Int. J. Hum.-Comput. Stud. 160, 102771.
Loveys, K., Sagar, M., Zhang, X., Fricchione, G., Broadbent, E., 2021. Effects of emotional expressiveness of a female digital human on loneliness, stress, perceived support, and closeness across genders: randomized controlled trial. J. med. Internet res. 23 (11), e30624.
MacDorman, K.F., 2019. In the uncanny valley, transportation predicts narrative enjoyment more than empathy, but only for the tragic hero. Comput. Hum. Behav. 94, 140–153.
McQuiggan, S.W., Lester, J.C., 2007. Modeling and evaluating empathy in embodied companion agents. Int. J. Hum.-Comput. Stud. 65 (4), 348–360. https://doi.org/10.1016/j.ijhcs.2006.11.015.
Merikivi, J., Tuunainen, V., Nguyen, D., 2017. What makes continued mobile gaming enjoyable? Comput. Hum. Behav. 68, 411–421.
Montgomery, A.L., Smith, M.D., 2009. Prospects for personalization on the Internet. J. Interact. Mark. 23 (2), 130–137.
Morris, R.R., Kouddous, K., Kshirsagar, R., Schueller, S.M., 2018. Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions [Original Paper]. J Med Internet Res 20 (6), e10148. https://doi.org/10.2196/10148.
Moussalli, S., Cardoso, W., 2020. Intelligent personal assistants: can they understand and be understood by accented L2 learners? Comput. Assist. Lang. Learn. 33 (8), 865–890.
Nass, C., Moon, Y., 2000. Machines and mindlessness: social responses to computers. J. soc. issues 56 (1), 81–103.
Olafsson, S., O'Leary, T.K., Bickmore, T.W., 2020. Motivating health behavior change with humorous virtual agents. In: Proceedings of the 20th ACM international conference on intelligent virtual agents.
Paiva, A., Leite, I., Boukricha, H., Wachsmuth, I., 2017. Empathy in virtual agents and robots: a survey. ACM Trans. Interact. Intell. Syst. 7 (3), 1–40. https://doi.org/10.1145/2912150.
Park, G., Yim, M.C., Chung, J., Lee, S., 2022a. Effect of AI chatbot empathy and identity disclosure on willingness to donate: the mediation of humanness and social presence. Behav. Inf. Technol. 42 (12), 1998–2010. https://doi.org/10.1080/0144929x.2022.2105746.
Park, S., Park, S., Whang, M., 2022b. Empathic responses of behavioral-synchronization in human-agent interaction. Comput. Mater. Contin. 71 (2).
Pickard, M.D., Roster, C.A., 2020. Using computer automated systems to conduct personal interviews: does the mere presence of a human face inhibit disclosure? Comput. Hum. Behav. 105, 106197.
Polignano, M., Lops, P., de Gemmis, M., Semeraro, G., 2023. HELENA: an intelligent digital assistant based on a lifelong health user model. Inf. Process. Manag. 60 (1), 103124.
Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind? Behav. Brain Sci. 1 (4), 515–526.
Rapp, A., Curti, L., Boldi, A., 2021. The human side of human-chatbot interaction: a systematic literature review of ten years of research on text-based chatbots. Int. J. Hum.-Comput. Stud. 151, 102630. https://doi.org/10.1016/j.ijhcs.2021.102630.
Reeves, B., Nass, C., 1996. The Media Equation: How People Treat Computers, Television, and New Media Like Real People. Cambridge University Press, Cambridge, United Kingdom.
Ren, Y., Elliott, S., Wang, Y., Yang, Y., Zhang, W., 2019. How shall I drive? Interaction modeling and motion planning towards empathetic and socially-graceful driving. In: 2019 International Conference on Robotics and Automation (ICRA).

Reniers, R.L., Corcoran, R., Drake, R., Shryane, N.M., Vollm, B.A., 2011. The QCAE: a questionnaire of cognitive and affective empathy. J. Pers. Assess. 93 (1), 84–95. https://doi.org/10.1080/00223891.2010.528484.
Rhee, C.E., Choi, J., 2020. Effects of personalization and social role in voice shopping: an experimental study on product recommendation by a conversational voice agent. Comput. Hum. Behav. 109, 106359.
Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. Annu. Rev. Neurosci. 27, 169–192.
Rogers, C.R., Farson, R.E., 2007. Active listening. Organizational Communication, 2nd ed. Transaction, Piscataway, NJ, pp. 319–334.
Samrose, S., Anbarasu, K., Joshi, A., Mishra, T., 2020. Mitigating Boredom using an empathetic conversational agent. In: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents.
Schmidmaier, M., Rupp, J., Cvetanova, D., Mayer, S., 2024. Perceived Empathy of Technology Scale (PETS): measuring empathy of systems toward the user. In: Proceedings of the CHI Conference on Human Factors in Computing Systems.
Seaborn, K., Gessinger, I., Yoshida, S., Cowan, B.R., Doyle, P.R., 2024. Cross-cultural validation of partner models for voice user interfaces. In: Proceedings of the 6th ACM Conference on Conversational User Interfaces, pp. 1–10.
Shum, H.-Y., He, X.-d., Li, D., 2018. From Eliza to XiaoIce: challenges and opportunities with social chatbots. Front. Inf. Technol. Electron. Eng. 19, 10–26.
Silke, C., Brady, B., Boylan, C., Dolan, P., 2018. Factors influencing the development of empathy and pro-social behaviour among adolescents: A systematic review. Child. Youth Serv. Rev. 94, 421–436.
Spreng, R.N., McKinnon, M.C., Mar, R.A., Levine, B., 2009. The Toronto Empathy Questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures. J. Pers. Assess. 91 (1), 62–71. https://doi.org/10.1080/00223890802484381.
Tai, T.-Y., Chen, H.H.-J., 2023. The impact of Google Assistant on adolescent EFL learners' willingness to communicate. Interact. Learn. Environ. 31 (3), 1485–1502.
Taylor, S.H., DiFranzo, D., Choi, Y.H., Sannon, S., Bazarova, N.N., 2019. Accountability and empathy by design: encouraging bystander intervention to cyberbullying on social media. Proc. ACM Hum.-Comput. Interact. 3 (CSCW), 1–26.
Vuong, Q.-H., La, V.-P., Nguyen, M.-H., Jin, R., La, M.-K., Le, T.-T., 2023. AI's humanoid appearance can affect human perceptions of Its emotional capability: evidence from self-reported data in the US. Int. J. Hum.–Comput. Interact. 1–12.
Wang, X.W., Luo, R., Liu, Y.T., Chen, P., Tao, Y.Y., He, Y.M., 2023. Revealing the complexity of users' intention to adopt healthcare chatbots: a mixed-method analysis of antecedent condition configurations. Inf. Process. Manag. 60 (5). https://doi.org/10.1016/j.ipm.2023.103444.
Xiao, Z., Zhou, M.X., Chen, W., Yang, H., Chi, C., 2020. If i hear you correctly: building and evaluating interview chatbots with active listening skills. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.
Yalçın, Ö.N., 2020. Empathy framework for embodied conversational agents. Cogn. Syst. Res. 59, 123–132.
Yamamoto, S., 2017. Primate empathy: three factors and their combinations for empathy-related phenomena. Wiley Interdiscip. Rev.: Cogn. Sci. 8 (3), e1431.
Yang, B., Sun, Y., Shen, X.-L., 2023. Understanding AI-based customer service resistance: a perspective of defective AI features and tri-dimensional distrusting beliefs. Inf. Process. Manag. 60 (3). https://doi.org/10.1016/j.ipm.2022.103257.
Yang, X., Aurisicchio, M., Baxter, W., 2019. Understanding affective experiences with conversational agents. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.
Yim, M.C., 2023. Effect of AI chatbot's interactivity on consumers' negative word-of-mouth intention: mediating role of perceived empathy and anger. Int. J. Hum.-Comput. Interact. https://doi.org/10.1080/10447318.2023.2234114.
Zhang, Z., Xie, J.A., Xu, X.T., Lu, H.R., Cheng, Y., 2023. Modeling the information behavior patterns of new graduate students in supervisor selection. Inf. Process. Manag. 60 (3). https://doi.org/10.1016/j.ipm.2023.103342.
Zhou, M.X., Mark, G., Li, J., Yang, H., 2019. Trusting virtual agents: the effect of personality. ACM Trans. Interact. Intell. Syst. (TiiS) 9 (2-3), 1–36.
Zhu, Y., Janssen, M., Wang, R., Liu, Y., 2022. It is me, chatbot: working to address the COVID-19 outbreak-related mental health issues in China. User experience, satisfaction, and influencing factors. Int. J. Hum.–Comput. Interact. 38 (12), 1182–1194.
Zhuo, T. Y., Huang, Y., Chen, C., & Xing, Z. (2023). Exploring ai ethics of chatgpt: a diagnostic analysis. *arXiv preprint arXiv:2301.12867*.
Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., Yang, D., 2024. Can large language models transform computational social science? Comput. Linguist. 50 (1), 237–291.