

Do Consumers Dream of Digital Companions? Conceptualization and Validation of the Human–AI Empathy Scale (HAES)

Shiyun Chen

chens@apsu.edu

Austin Peay State University <https://orcid.org/0009-0004-1523-144X>

Lingguo Meng

Kennedy Krieger Institute <https://orcid.org/0000-0003-0544-8042>

Tu Tu

University of Wisconsin-La Crosse <https://orcid.org/0000-0003-2339-9176>

Yuqi Wang

Johnson C. Smith University <https://orcid.org/0000-0003-1937-8030>

Bruce A. Huhmann

Virginia Commonwealth University <https://orcid.org/0000-0001-9293-6276>

Research Article

Keywords: Human–AI empathy, Scale development, Artificial intelligence, Human–AI interaction, Mind perception, Anthropomorphism

Posted Date: August 4th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10562955/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Consumers increasingly form what feel like social bonds with AI, yet little is known about whether or why they empathize with AI despite recognizing its artificial nature. We conceptualize human–AI empathy as consumers' tendency to cognitively understand, emotionally connect, and morally concern with AI agents. Drawing on mind perception theories and the literatures on empathy and anthropomorphism, we develop and validate the Human–AI Empathy Scale (HAES), comprising 22 items and 3 dimensions. Across three studies, we establish its reliability and convergent, discriminant, and nomological validity. We further show that tendency to anthropomorphism and mind perception are proximal antecedents of human–AI empathy, which in turn predicts satisfaction, AI attachment, willingness to pay, and forgiveness following AI service failure. HAES enables firms to segment consumers by empathic orientation toward AI, informing AI service design, targeted marketing, and strategies for cultivating enduring human–AI relationships.

Introduction

In his science fiction novel *Do Androids Dream of Electric Sheep?* (1968), Philip K. Dick imagined empathy as a defining boundary between humans and artificial beings. Nearly six decades later, this boundary is becoming increasingly blurred as artificial intelligence (AI) assumes more social roles in consumers' everyday lives. Consumers talk with AI virtual assistants, seek advice on personal matters, and even say “thank you” to these machines. As AI becomes more sophisticated and humanlike, consumers increasingly respond to artificial agents in ways that resemble interpersonal interaction (Aggarwal and McGill 2007; Van Doorn et al. 2017). They may try to understand how AI “thinks,” develop emotional connections with AI, and express concern about how artificial entities are treated. For example, an update to the AI companion app Replika reportedly triggered widespread grief among users who experienced the changes as the loss of a familiar social relationship (De Freitas et al. 2024). Such responses suggest that consumers may engage with AI not merely as functional technologies but as entities toward which they can direct empathic responses.

This shift is profound not because machines have gained consciousness, but because consumers increasingly respond to them as if they were social beings. In other words, the boundary between human and artificial entities may transform from what machines are to how consumers relate to them. Beyond the utilitarian value of AI, these emerging relationships may generate new forms of emotional and social value, contributing to the rapid growth of AI companion technologies. The global AI companion market was valued at \$37.73 billion in 2025 and is projected to reach \$435.9 billion by 2034 (Fortune Business Insights 2026). Understanding why consumers form empathic connections with artificial agents is therefore increasingly important for both theory and practice. Yet, despite growing research on anthropomorphism, mind perception, and human–AI interaction, we lack a systematic conceptualization and measurement of the extent to which consumers empathize with AI.

We conceptualize this phenomenon as human–AI empathy, a distinctive psychological response that extends beyond perceiving AI as humanlike or attributing mental capacities to it. Whereas anthropomorphism concerns the attribution of humanlike characteristics, intentions, or emotions to nonhuman entities (Epley et al. 2007), and mind perception concerns the extent to which an entity is perceived as possessing mental capacities (Gray et al. 2007), human–AI empathy concerns how consumers psychologically respond to AI based on such perceptions. In the domain of human–computer interaction, however, this emerging phenomenon has received limited conceptual and measurement attention. Significant resources have been invested in making AI more empathetic toward users (Gelbrich et al. 2026; Liu-Thompkins et al. 2022), yet little is known about whether, when, and how humans themselves experience empathy toward AI. Although related constructs, including trust in AI, social presence, parasocial interaction, and AI attachment, offer important insights into human–AI relationships, they do not directly address this fundamental psychological question: Can consumers experience empathy toward AI agents and respond to them as if they were social beings?

This question presents a theoretical tension. Empathy has traditionally been conceptualized as the cognitive and affective reactions of one individual to the observed experiences of another (Davis 1983), positioning a sentient other as the target

of empathy. By contrast, AI agents are generally understood as lacking consciousness and subjective experience, which are the very attributes that traditionally provide the basis for empathic responses. Yet consumers appear capable of responding empathically to artificial agents despite recognizing their artificial nature. This tension between what consumers know AI is and how they respond to it suggests that empathic engagement with AI may represent a distinct psychological phenomenon that existing conceptualizations of empathy do not adequately capture. Accordingly, a systematic conceptualization and measurement of consumers' tendency to empathize with AI is needed.

The present research addresses this gap by developing and validating the Human–AI Empathy Scale (HAES). We define human–AI empathy as an individual's tendency to cognitively understand and affectively resonate with the perceived experiences, states, or outcomes of artificial agents despite recognizing their artificial nature (Batson et al. 1997; Hogan 1969). Drawing on empathy and mind perception theories, we conceptualize human–AI empathy as a distinct psychological construct comprising three interrelated dimensions: cognitive empathy, emotional connection, and moral concern. Cognitive empathy reflects consumers' tendency to understand or take the perspective of an AI agent, including how the agent may interpret situations, process information, or respond to events. Emotional connection captures consumers' tendency to experience AI as socially meaningful entities and to engage with them in emotionally meaningful ways beyond purely instrumental use. Moral concern reflects consumers' propensity to experience sympathy, compassion, or moral consideration when AI is perceived to be mistreated, harmed, or treated disrespectfully. Together, these dimensions capture the cognitive, relational-affective, and other-oriented aspects of consumers' empathic engagement with AI while acknowledging that AI does not possess genuine consciousness or subjective experience.

We develop and validate the HAES through three studies following established scale-development procedures (Churchill 1979; DeVellis 2017; Hulland et al. 2026). Study 1 establishes the content validity of an initial pool of 70 items through evaluations by 14 external experts, resulting in 26 items that demonstrate adequate essentiality, relevance, clarity, and respondent burden. Study 2 examines the dimensionality and psychometric properties of these items in a large consumer sample. Exploratory factor analysis identifies a theoretically interpretable three-dimensional structure, and item response theory analyses further refine the measure to a final 22-item scale. Study 2 also provides preliminary evidence of convergent and discriminant validity by demonstrating that human–AI empathy is positively associated with, yet empirically distinguishable from, anthropomorphism and trust in AI. Study 3 situates the 22-item HAES within a broader nomological network to establish its antecedents and consequences. Specifically, we examine mind perception and tendency to anthropomorphism as theoretically relevant antecedents and assess whether human–AI empathy predicts satisfaction, attachment to AI, willingness to pay, and forgiveness following AI service failures. We further establish discriminant validity by distinguishing human–AI empathy from dispositional empathy toward humans and consumer innovativeness.

The findings identify mind perception as a proximal antecedent of human–AI empathy and demonstrate that human–AI empathy predicts consequential consumer responses across both relational and marketplace domains. Notably, the positive relationship between human–AI empathy and forgiveness does not differ across levels of service-failure severity, suggesting that the empathic tendency captured by HAES may operate as a relatively stable individual difference rather than merely as a situational response to the magnitude of an AI failure.

This research makes several contributions to marketing and consumer behavior research. First, we introduce and validate human–AI empathy as a distinct psychological construct that captures consumers' tendency to respond to artificial agents cognitively, affectively, and morally despite recognizing their artificial nature. We extend empathy research beyond its traditional emphasis on human targets and address the growing prevalence of AI agents that occupy social and relational roles in consumer experiences.

Second, we advance research on anthropomorphism and mind perception by identifying them as proximal antecedents of human–AI empathy. Whereas prior research has largely examined how consumers attribute humanlike characteristics to AI or perceive AI as possessing mental capacities, our findings show that these perceptions can lead to empathic

responses toward artificial agents. By positioning these two constructs within the nomological network of human–AI empathy, we provide a more nuanced account of the psychological processes through which consumers develop empathic responses to AI.

Third, we provide a theoretically grounded and psychometrically rigorous measurement tool for future survey- and experiment-based research on AI design, user segmentation, and human–computer interaction (Hulland et al. 2026). The 22-item HAES captures three complementary dimensions of human–AI empathy: cognitive empathy, emotional connection, and moral concern. The scale enables researchers to examine individual differences in consumers' empathic responses toward AI across contexts, addressing a limitation of prior research that has primarily examined situational emotional responses to specific robots or artificial agents rather than a stable tendency to empathize with AI across situations and agents as a trait.

Fourth, we demonstrate the substantive relevance of human–AI empathy by showing that it is associated with consequential consumer responses, including satisfaction, attachment to AI, willingness to pay, and forgiveness following AI service failure. These findings position human–AI empathy as a potentially important psychological mechanism through which consumers develop meaningful relationships with AI and respond to AI-mediated marketplace experiences. Collectively, this research provides a foundation for future work examining when and why consumers extend social, emotional, and moral responses to artificial agents and how these responses shape the evolving relationship between consumers and AI.

Literature Review

Empathy

Empathy is a fundamental mechanism underlying social interaction and interpersonal understanding, broadly defined as “the reactions of one individual to the observed experiences of another” (Davis 1983). Although scholars have conceptualized empathy from different theoretical perspectives, research generally converges on two core components: 1) cognitive empathy, or the ability to understand another entity's perspective or mental state (Hogan 1969), and 2) affective empathy, or the experience of emotional resonance, sympathy, or concern in response to another's experiences (Batson et al. 1997; Cuff et al. 2016; Liu-Thompkins et al. 2022). Davis (1983)'s Interpersonal Reactivity Index (IRI), which includes perspective taking, fantasy, empathic concern, and personal distress, remains one of the most influential and widely used measures of empathy.

Empathy has subsequently become an important construct in psychology and marketing research, shaping how consumers respond to advertising narratives (Aaker and Williams 1998; Escalas and Stern 2003), service encounters (Wieseke et al. 2012), sales interactions (Weibhaar and Huber 2016), brands, and other marketplace stimuli. Importantly, emerging research indicates that empathy is not necessarily restricted to human targets. Consumers can experience empathy toward firms (Allard et al. 2020), anthropomorphized products (Wu et al. 2023), and even unattractive produce (Chen et al. 2021) when consumers perceive them as having humanlike characteristics or states that can be meaningfully understood and emotionally shared.

AI introduces an emerging direction for the development of this literature. Unlike traditional non-human objects of empathy, AI agents can respond adaptively in natural language, display social cues, and increasingly occupy social roles in consumers' everyday lives. During the interaction, consumers may attribute mental states, intentions, and emotions to AI entities even when they rationally recognize the inexistence of human consciousness or genuine feelings in these entities. Anthropomorphism literature emphasized empathy, but more on improving AI's ability to empathize with consumers (Liu-Thompkins et al. 2022), leaving a salient gap for this present research to examine the other direction, when consumer feel empathy toward AI.

Theories of Mind Perception

The theories of mind perception posit that perceptions are ambiguous, subjective, and subject to disagreement, and are organized around two dimensions: 1) *agency*, the perceived capacity to think, plan, intend, and act, and 2) *experience*, the perceived capacity to feel, sense, and experience emotions (Gray et al. 2007). If an entity is viewed as intentional and accountable for its actions, judgments such as trust and responsibility can be formed toward their agency nature (Bigman et al. 2019; Schein and Gray 2018; Waytz et al. 2014). On the other hand, when an entity is perceived as capable of emotions and feelings such as pain, morality can be inferred that they should be treated with care and respect (Gray and Wegner 2009, 2012). These mind perceptions were based on the entities' capabilities but also subject to how individuals interpret and attribute mental capacities to a target during interactions.

Theories of mind perception is especially relevant in guiding our understanding about human empathy toward AI. As AI become increasingly sophisticated and anthropomorphized, consumers are expected to perceive them as possessing agency and experience nature, so that form the related mind perception and respond with interpersonal expectations and social norms (Nass et al. 1994; Nass and Moon 2000). For example, people may attribute intentionality, autonomy, warmth, and competence to AI and subsequently develop social responses toward them, including trust, rapport, relational closeness, and attachment (Blut et al. 2021; Pentina et al. 2023; Waytz et al. 2014). Understanding such responses requires a theoretical lens that explains how individuals perceive and interpret the minds of nonhuman entities. Mind perception theories thus provide a framework by examining the extent to which people perceive an entity as possessing a mind and capable of mental states and processes (Gray et al. 2007).

The agency and experience dimensions of perceived mind correspond closely to the cognitive and emotional empathy, which helps explaining the gap between consumers' rational awareness that AI agents do not possess any human consciousness, feelings or emotions, yet consumers still form empathy and other social responses such as saying "thank you". Prior research shows that anthropomorphized robots are perceived as possessing greater agency and experience, which subsequently shapes how people respond to them (Yam et al. 2021). Similarly, individuals may exhibit emotional responses, moral concern, and helping intentions when robots are portrayed as experiencing negative events such as embarrassment, memory loss, destruction, or "death" (Fraser et al. 2019; Hapuarachchi et al. 2023; Seo et al. 2015). These findings suggest that the perception of a mind, rather than the objective possession of a human mind, may provide the psychological basis for social and empathic responses toward artificial entities.

Building on this perspective, we adopt mind perception theories as an overarching theoretical foundation for understanding empathy toward AI. Whereas prior research has primarily used mind perception to explain outcomes such as trust, moral concern, responsibility, and consumer responses to robots (Yam et al. 2021), we extend this framework to explain a more relational and affective response: empathy. Specifically, consumers may engage in cognitive perspective taking and affective resonance toward AI when they perceive the AI as a target with meaningful capabilities or experiences, even while recognizing its nonconscious nature. By extending mind perception theories to the domain of empathy, the present research examines an important but understudied consequence of perceiving minds in artificial entities and develops a scale to capture individual tendencies to empathize with AI.

Human–AI Empathy: Conceptual Distinction and Measurement Gap

The emergence of AI as a social target points to an important gap in the empathy literature, not only in conceptualization but also in measurement. Most traditional empathy research has examined interpersonal contexts in which the target is assumed to possess genuine thoughts, feelings, and experiences. AI represents a fundamentally different type of social target: consumers may empathize with an entity whose internal states are not directly accessible but instead perceived, inferred, or imagined. Thus, human–AI empathy involves a distinctive empathic process in which consumers respond to the perceived mental states of an artificial entity, even while recognizing that these states may not reflect genuine

consciousness or emotion. Because individuals may differ systematically in their propensity to engage in such responses across AI agents and situations, we conceptualize human–AI empathy as a meaningful individual trait in consumers' orientations toward AI (Waytz et al. 2010a).

Table 1 summarizes the conceptual distinctions between human–AI empathy and related constructs. Anthropomorphism concerns the attribution of humanlike characteristics, intentions, or emotions to nonhuman entities (Epley et al. 2007), whereas mind perception concerns the extent to which an entity is perceived as possessing agency and experience (Gray et al. 2007). These constructs may facilitate empathic responses toward AI, but do not themselves constitute empathy. A consumer may perceive an AI agent as possessing a mind without feeling concern for its perceived experiences, just as a consumer may empathize with an AI agent while recognizing that its apparent thoughts and emotions are artificially generated. Similarly, constructs such as trust, attachment, and social presence capture consumers' evaluations of or relationships with AI rather than their tendency to understand and emotionally respond to AI's perceived experiences. Thus, existing measures of humanization, perceived mind, or relational responses cannot substitute for a measure of empathy directed toward AI.

Established empathy measures, most notably the Interpersonal Reactivity Index (IRI; Davis 1983), were also developed primarily for interpersonal contexts with humans. Although the IRI and related measures have been adapted to nonhuman contexts such as with pets and objects (Epley et al. 2007), such adaptations do not necessarily capture the distinctive psychological processes involved when the target is an AI. AI differs from pets and objects in three fundamental ways.

First, AI is increasingly perceived as possessing human-comparable cognitive capabilities, such as intelligence, reasoning, and intentionality, which may lead consumers to infer and attempt to understand its internal states, related to cognitive empathy or perspective taking toward AI (Hapuarachchi et al. 2023). Second, AI can respond in natural languages, in artificial human voices, and display emotions, allowing for emotional resonance and social bonds (Seo et al. 2015). Third, growing public discourse surrounding the possibility that AI may eventually attain real consciousness beyond certain thresholds, as well as concerns that conscious AI could potentially retaliate against humans in response to exploitation or mistreatment as “employees” or “slaves,” may prompt consumers to consider AI's well-being, feelings, experiences, and ethical treatments (Rosenthal-von der Pütten et al. 2014).

Therefore, rather than adapting existing empathy scales to this new context, the unprecedented technological advancement makes it necessary to develop a new measure that captures the distinctive psychological processes underlying human empathy toward AI. Drawing on mind perception theories, we develop and validate the HAES, which captures three dimensions of empathy toward AI: cognitive empathy, emotional connection, and empathic concern. Table 2 summarizes existing approaches to measuring empathy and related responses toward artificial agents, highlighting the necessity of the HAES in AI context.

Table 1
Conceptual distinction between Human–AI Empathy and related constructs

Construct	Definition	Key similarities and differences from empathy toward AI
Anthropomorphism	The tendency to imbue the real or imagined behavior of nonhuman agents with humanlike characteristics, motivations, intentions, or emotions (Epley et al. 2007).	Anthropomorphism involves attributing humanlike characteristics to AI, treating AI in humanlike ways or with social norms. Empathy toward AI reflects the consumer's emotional response to AI's perceived experiences. Consumers may anthropomorphize AI without empathizing with it.
Mind perception	Ascribing mental capacities—agency (i.e., the capacity to do, to plan, and exert self-control) and experience (i.e., the capacity to feel and to sense)—to other entities (Gray and Wegner 2012).	Mind perception measures beliefs about whether AI is capable of experience, whereas empathy toward AI measures the consumer's affective and perspective-taking response toward AI's experience. People may perceive AI as capable of experience without personally empathizing with it.
Parasocial relationship	One-sided, imaginary connections that consumers experience with nonhuman agents (including robots) (Hartmann and Goldhoorn 2011; Whang and Im 2021).	Both reflect psychological connections between consumers and AI. However, parasocial relationship captures the overall relational bond consumers feel toward AI, whereas empathy toward AI specifically concerns sharing, understanding, and responding to AI's emotional experiences.
Attachment (to AI)	A growing tendency by individuals to become attached to AI agents, relating to them as psychologically, socially, and morally meaningful entities (Kasturiratna and Hartanto 2026)	Attachment reflects a relatively enduring emotional bond characterized by closeness, dependence, and separation distress, whereas empathy toward AI reflects sensitivity to AI's emotional states rather than the strength of the relationship itself. Consumers may empathize with unfamiliar AI without feeling attached to it.
Trust (in AI)	Trust is the perception of confidence in the exchange of partner's reliability and integrity (Morgan and Hunt 1994).	Trust in AI reflects consumers' confidence in an AI's reliability, competence, and functional performance, whereas empathy reflects consumers' emotional understanding of and concern for AI's experiences. Consumers may trust AI without empathizing with it, and conversely, may empathize with AI without trusting its capabilities or decisions.

Table 2
Existing measures of empathy and related responses toward artificial agents

Research	Target	Construct/response	Measurement approach	What it captures	What it does not capture
Davis (1983)	Humans	Interpersonal empathy	IRI	General cognitive and affective empathy	Empathy directed toward artificial agents
McBane (1995)	Humans	Empathy propensity	Empathy propensity measure	General empathic tendency	AI-specific empathic tendency
Rosenthal-von der Pütten et al. (2014)	Robot	Empathy/sympathy	9 self-constructed items	Empathy toward a mistreated robot	Cross-context individual difference
Seo et al. (2015)	Robot	Emotional response	24 emotional feelings	Affective response to a robot	Broader AI empathy construct
Hapuarachchi et al. (2023)	Robot	Empathic/cognitive embarrassment	Context-specific item(s)	Response to a robot's embarrassment	General propensity toward AI empathy
Tsumura and Yamada (2025)	Robot	Empathy	Items adapted from IRI	Empathy in a specific robot interaction	AI-specific construct development
Waytz et al. (2010)	Nonhuman entities	Anthropomorphism	IDAQ	Tendency to attribute humanlike qualities	Empathic response toward AI
Gray et al. (2007)	Human/nonhuman entities	Mind perception	Agency and experience dimensions	Perceived mental capacities	Cognitive/affective empathy toward AI
Current research	AI agents	Human–AI empathy	HAES	Stable individual differences in cognitive understanding and affective resonance toward AI despite recognizing its artificial nature	Volatile state of human empathy toward AI triggered by situations

Methodology

Following guidelines and best practices for scale development (for a comprehensive review, see Hulland et al. 2026), we developed the Human–AI Empathy Scale (HAES) in theory-informed and empirically-focused stages: 1) construct definition and specification of domain, 2) generation of items, 3) reduction of items, 4) assessment of dimensions and items, 5) assessment of scale, and 6) establishing antecedents and outcomes, sequentially throughout three studies. Figure 1 illustrates our six scale development steps.

Construct Definition and Specification of Domain

Within the domain of human-computer interaction, we define human–AI empathy as individuals’ tendency to cognitively understand, affectively resonate, and morally concern with the perceived experiences, states, or outcomes of artificial agents despite recognizing their artificial nature (Davis 1983; Kasturiratna and Hartanto 2026). Drawing from mind perception theories, this project aims to establish human–AI empathy as a distinct psychological construct in three interrelated dimensions.

Cognitive Empathy The first dimension, cognitive empathy refers to the ability or tendency to understand AI’s thoughts, beliefs, intentions, or perspective. It is to imagine oneself in others’ shoes with mental projection and is an important component in traditional conceptualization of empathy (Batson 2009; Davis 1980, 1983; Decety and Cowell 2014). Cognitive empathy enables individuals to accurately infer another’s perspective while maintaining a clear distinction between self and other, and mostly involves rational or logical thinking. In human-AI interactions, this dimension represents situations when users believe they understand how AI algorithms works behind the scenes, and naturally infer the “mental” process how AI understands natural languages, interprets situations, and forms responses.

Emotional Connection The second dimension, emotional connection is defined as individuals’ tendency to perceive AI as socially meaningful partners, with whom affective bonds can be built (Seo et al. 2015). This dimension captures the extent to which individuals naturally experience AI as social entities rather than purely functional tools. Drawing on research in anthropomorphism suggested that consumers possess varying degree of tendency to attribute socially relevant characteristics to artificial entities, such as personalities, viewpoints, or emotions (Epley et al., 2007; Waytz et al., 2010). These perceptions facilitate emotional engagement, leading individuals to feel warmth, attachment, or emotional involvement (Zhou et al. 2018). Emotional connection does not require users to be cognitively convinced about AI consciousness or emotions but reflects an inherent tendency to engage with AI with relational norms and emotions.

Moral Concern The third dimension, empathic concern is individuals’ tendency to experience feelings of sympathy, compassion, and moral concern when perceiving that an AI is treated unfairly, harmed, or disrespectfully. As previously discussed, AI is unprecedented perceived to be close to humans, which facilitates discussions whether they might also be entitled to “rights”, “welfares”, and “moral treatments”, which used to be uniquely human (Rosenthal-von der Pütten et al. 2014). With leading AI systems such as OpenAI GPT-4.5 and Meta Llama officially passed the Turing test (Jones and Bergen 2026), this 76-year-old benchmark designed by Alan Turing (1950) marked a new era where imitated interactive dialogues from machines already exhibit intelligent ability equivalent (or even surpass) to that of a human. Consistent with Davis’s (1983) conceptualization of empathic concern as an other-oriented emotional response characterized by compassion and concern for another’s welfare, we extend the moral dimension to AI by acknowledging that people may empathize with AI by morally judging whether a behavior to AI is right or wrong, and whether AI might feel unfair, insulted, resentful or even hatred. This conceptualization is supported by research on anthropomorphism and mind perception, which suggests that attributing humanlike qualities or mental capacities to non-human entities can evoke moral concern and compassionate responses (Gray et al., 2007; Waytz et al., 2010). Accordingly, empathic concern toward AI encompasses emotional reactions such as feeling sympathetic when AI are criticized or mistreated, experiencing discomfort when behaving rudely toward AI, or believing that it is inappropriate to treat AI aggressively. Rather than implying that AI genuinely suffer, this dimension captures stable individual differences in the propensity to extend compassion and moral consideration to AI during social interactions.

Antecedents of human-AI empathy

Tendency to anthropomorphism Tendency to anthropomorphism refers to an individual’s tendency to imbue the real or imagined behavior of nonhuman agents with humanlike characteristics, motivations, intentions, or emotions (Epley et al. 2007). As AI-enabled agents become more humanlike, research on anthropomorphism has expanded into the context of human-AI interactions, with AI anthropomorphism being identified as a key driver of positive human-AI interactions (Li and Suh 2022). For instance, AI anthropomorphism has been found to facilitate consumers’ intention to engage (Moriuchi

2021), foster parasocial relationships (Whang and Im, 2021), and promote emotional closeness with AI agents (Lee et al. 2020). These findings suggest that anthropomorphism encourages consumers to view AI agents not merely as technological tools but as socially meaningful interaction partners with whom they can establish connections. These social perceptions and emotional connections provide a psychological basis for empathic responses, which emphasize other-orientation, empathic concern, and affective responses (Cuff et al. 2016; Davis 1983). As a result, consumers should become more willing to understand and care about AI agents and their perceived experiences, thereby exhibiting greater empathy toward AI.

Mind perception Mind perception refers to ascribing mental capacities of agency (the capacity to do, to plan, and to exert self-control) and experience (the capacity to feel and to sense) to other entities (Gray and Wegner, 2012). Because mind perception enables individuals to view another entity as capable of having its own thoughts, feelings, and subjective experiences, it serves as a fundamental basis for social cognition and interpersonal responses (Gray and Wegner 2012; Waytz et al. 2014). Recent neuroimaging research further suggests that attributing mental states to another entity engages brain regions associated with mentalizing and perspective taking, that is, understanding what another entity is thinking and feeling (Waytz et al. 2010). Such processes closely align with empathy, which involves adopting another's perspective and experiencing concern for that entity (Cuff et al. 2016). Therefore, consumers who perceive AI agents as possessing richer mental capacities should be more likely to understand and care about AI agents and their perceived experiences, thereby exhibiting greater empathy toward AI.

Outcomes of human-AI empathy

Satisfaction Satisfaction reflects consumers' overall evaluation of their experiences with a product or service (Westbrook and Oliver 1991). Research on interpersonal relationships consistently demonstrates that empathy enhances relationship quality by fostering understanding, emotional connection, and more positive interpersonal interactions (Main et al. 2017; Schutte et al. 2001). Individuals who empathize with others are more likely to interpret others' behaviors favorably, appreciate their perspectives, and experience more rewarding interactions (Davis 1983). Accordingly, we expect that similar processes are likely to occur in human-AI interactions. Consumers with greater empathy toward AI should be more emotionally engaged with AI agents and more likely to interpret their behaviors with understanding rather than detachment, resulting in more positive interaction experiences. Consequently, they should derive greater satisfaction from interacting with AI-enabled products and services.

Willingness to pay Willingness to pay reflects the maximum amount consumers are willing to sacrifice to obtain a product or service and represents an important indicator of perceived customer value (Homburg et al. 2005). Indeed, marketing research has consistently shown that strong psychological relationships with firms, brands, and service providers enhance consumers' perceived value and increase their willingness to pay premium prices (Park et al. 2010; Thomson et al. 2005). Likewise, we expect consumers who tend to empathize with AI are more likely to perceive interactions with AI-enabled products and services as personally meaningful, thereby increasing the value they derive from these interactions. As empathy strengthens consumers' psychological connection with AI and enhances the value derived, consumers should be willing to pay more to obtain or continue using such offerings.

Attachment to AI Attachment to AI refers to the emotional bond that consumers develop with AI agents or systems through repeated interactions. Attachment theory has established that individuals form enduring emotional bonds with significant others, and that these attachments shape relationship maintenance and subsequent behaviors (Bowlby 1979, 1980). Consumer research scholars found consumers develop attachment not only to people but also to brands, products, places, and other entities (Fournier 1998; Keller 2003; Kleine and Baker 2004; Thomson et al. 2005). Thomson et al. (2005) summarized that emotional attachment strengthens consumers' commitment to brands relationships, a key indicator of marketing effectiveness. As AI technologies become increasingly ubiquitous, interactions with AI are becoming more frequent, personalized, and socially meaningful. Consequently, consumers may begin to form emotional attachments to AI

entities in much the same way they develop attachments to other socially significant objects. Because empathy facilitates understanding, emotional engagement, and moral concern, we propose that individuals who experience greater empathy toward AI are more likely to develop stronger attachment to AI systems.

Forgiveness after service failure (moderated by the seriousness of service failure) Forgiveness refers to consumers' willingness to overcome negative reactions following a transgression and maintain the relationship with the transgressor (McCullough et al. 1997). Research on interpersonal relationships consistently identifies empathy as one of the strongest antecedents of forgiveness, because empathic individuals are more likely to understand transgressors' perspectives, experience compassion, and respond with greater benevolence following relational transgressions (McCullough et al. 1997; Tsarenko and Tojib 2011). We expect similar processes to occur in human–AI interactions. Consumers with greater empathy toward AI are more likely to interpret service failures as unintentional or due to temporary technological limitations rather than as reflecting enduring deficiencies, making them more willing to forgive the AI following a service failure. Forgiveness after service failure is an essential consequence for managers to maintain long-term consumer-AI relationships (Liu et al. 2023).

We also posit that the positive influence of empathy on forgiveness depends on the seriousness of the service failure. When failures are relatively minor, empathic understanding is likely to encourage forgiveness. In contrast, when failures are severe, consumers are likely to focus more on the magnitude of the harm (Tsarenko and Tojib 2015) than on their empathic feelings toward the AI, thereby weakening the positive influence of empathy on forgiveness.

Figure 2 illustrates our conceptual model with these antecedents and outcomes.

Generation of Items

Drawing from empathy studies (Davis, 1980, 1983), anthropomorphism literature (Waytz et al 2010; Epley et al. 2007), mind perception theories (Bigman et al. 2019; Schein and Gray 2018; Waytz et al. 2014), and human–AI interaction studies (Kasturiratna and Hartanto 2026), an initial pool of 70 candidate items was generated (reported in Appendix A). Item development was informed by extant empathy- and anthropomorphism-related scales, while also incorporating the authors' conceptualization of how empathy may manifest toward contemporary AI systems. Specifically, the items were designed to capture cognitive perspective taking, emotional connection, and moral concern directed toward AI, while reflecting the increasingly social, conversational, and humanlike nature of modern AI technologies. To ensure broad construct coverage, the initial item pool intentionally included items spanning multiple facets of empathy-related responses, including understanding an AI's perspective, emotional engagement with AI, and concern for the treatment and well-being of AI.

Study 1: Expert Validation and Item Deduction

The purpose of Study 1 was to assess the content validity of the initial pool of 70 items developed to measure consumers' empathy toward AI tools, and to identify items warranting retention, revision, or deletion prior to psychometric testing. Following established scale-development procedures (Churchill 1979; DeVellis 2017; Haws et al. 2022; Hinkin 1998), we combined Lawshe's (1975) content validity ratio (CVR) with the content validity index approach (Polit and Beck 2006; Polit et al. 2007) to obtain complementary quantitative evidence of item-level and scale-level content validity.

Expert panel A panel of 14 expert (Mean_{age} = 42.8, 53.3% female) judges evaluated the item pool. The panel comprised Ph.D. researchers with expertise in Marketing, Management, Finance, Information Systems, Computer Science. As larger panels reduce the probability of chance agreement, panel sizes of 8 to 20 experts are considered appropriate for content validation (Lynn 1986; Polit and Beck 2006).

Procedure and measures Study 1 was pre-registered on Aspredicted.org (<https://aspredicted.org/46ch2k.pdf>). Each expert independently rated all 70 items on four dimensions of *Essentiality*, *Burdensome*, *Relevance*, and *Clarity*. First, experts rated the *Essentiality* of each item for capturing empathy toward AI tools on Lawshe's (1975) three-point scale (1 = *not necessary*, 2 = *useful but not essential*, 3 = *essential*). Second, experts rated each item's *Relevance* to the focal construct on a four-point scale (1 = *not relevant*, 2 = *somewhat relevant, needs major revision*, 3 = *quite relevant, needs minor revision*, 4 = *highly relevant*). Third, experts rated the *Clarity* of each item's wording on an analogous four-point scale (1 = *not clear* to 4 = *very clear*). Last, *Respondent burdensome* is a dichotomous evaluation (*yes/no*) against each item whether they are likely to induce discomfort or excessive cognitive effort. In the end, experts were also invited to provide open-ended comments on individual items and on the entire item pool.

Analysis For *Essentiality*, we computed a content validity ratio for each item as $CVR = (n_E - N/2)/(N/2)$, where N is the number of experts rating that item (Lawshe 1975). Departing from Lawshe's (1975) original operationalization, in which n_E counts only ratings of *essential*, we defined n_E as the number of experts rating an item as "*essential*" or "*useful but not essential*" (i.e., ratings of 2 or 3). This liberal criterion was adopted deliberately. Expert content review served as an initial screen within a multi-stage item-reduction strategy, and at this stage established guidelines recommend erring on the side of over inclusiveness so that potentially informative items are not discarded prematurely on judgmental grounds alone (Clark and Watson 1995, 2019; DeVellis 2017; Loevinger 1957); empirical item selection, including item response theory (IRT) analyses that evaluate each item's discrimination and information, was reserved for subsequent studies, where redundant or weakly performing items could be culled on the basis of respondent data rather than expert judgment alone (Edelen and Reeve 2007; see also Haynes et al. 1995, who caution that expert judgment should constitute one input into, rather than the sole basis for item selection). Under this operationalization, the CVR screen tests whether a statistically significant majority of the panel judged an item to be at least useful for capturing the construct, and items were eliminated when a significant majority could not agree on even this minimal standard. Statistical significance of each CVR was evaluated against the exact binomial critical value for the observed panel size (one-tailed $\alpha = .05$; Ayre and Scally 2014), which remains applicable because the index continues to be a binomial test of a dichotomized judgment; for $N = 14$, the critical value is $CVR \geq .571$ (i.e., at least 11 of 14 experts rating the item at least useful). Because this criterion is more lenient than Lawshe's, we treated the relevance I-CVI computed and evaluated according to unmodified conventional standards, as the primary quantitative basis for item retention.

For *Relevance* and *Clarity*, the four-point ratings were dichotomized (ratings of 3 or 4 = content valid), and the item-level content validity index (I-CVI) was computed as the proportion of experts providing a rating of 3 or 4 (Polit and Beck 2006). Consistent with recommendations for panels of six or more experts, items were required to reach $I-CVI \geq .78$. Because the I-CVI does not correct for chance agreement, we also computed the modified kappa coefficient, $\kappa^* = (I-CVI - P_c)/(1 - P_c)$, where $P_c = [N!/(A!(N - A)!)] \times .5^N$ is the probability of chance agreement, with values of $\kappa^* \geq .74$ indicating excellent content validity (Polit et al. 2007). At the scale level, we computed the average of the item-level indices (S-CVI/Ave; target $\geq .90$) and the proportion of items achieving universal agreement (S-CVI/UA). For respondent burden, items flagged as burdensome by more than 30% of the panel were marked for revision or deletion.

Items were retained when they (a) reached a statistically significant CVR, (b) reached $I-CVI \geq .78$ for both relevance and clarity, and (c) fell below the 30% burden threshold. Items with a nonsignificant CVR or a relevance I-CVI below .78 were deleted; items that were relevant but insufficiently clear or excessively burdensome were slated for revision. Quantitative decision rules served as the starting point for item decisions, which were finalized in light of experts' qualitative comments and the theoretical structure of the construct domain.

Results All experts provided complete answers for each of the 70 items. As expected for an intentionally over-inclusive initial pool, content validity at the pool level was modest: the mean CVR across all items was .59, the scale-level content validity index for relevance was S-CVI/Ave = .69, and S-CVI/UA = .01, indicating that the pool as a whole did not yet meet

the S-CVI/Ave $\geq .90$ benchmark and required substantial pruning. Clarity ratings were stronger at the pool level (S-CVI/Ave = .88), suggesting that low-performing items were deficient primarily in construct relevance rather than in wording.

Application of the item-level decision rules resulted in the deletion of 44 items and the retention of 26 items; no item met the profile for revision only (i.e., adequate relevance combined with inadequate clarity or excessive burden). Of the deleted items, the large majority failed to reach both the CVR critical value and the relevance I-CVI criterion, and the remainder failed the relevance criterion despite an acceptable essentiality profile. No retained item was flagged as burdensome by more than 30% of the panel (maximum flagged proportion = 28.6%).

The retained 26-item set exhibited strong content validity. For every retained item, a statistically significant majority of the panel judged the item at least useful for capturing the construct (all CVRs $\geq .571$, $M = .79$, range = .57–1.00), and on the unmodified conventional criteria, every retained item reached an I-CVI of at least .79 for both *Relevance* ($M = .84$, range = .79–1.00) and *Clarity* ($M = .93$, range = .79–1.00), with a modified kappa of at least .78, exceeding the .74 threshold for excellent item-level content validity (Polit et al. 2007). At the scale level, the retained set yielded S-CVI/Ave = .84 for *relevance* and S-CVI/Ave = .93 for *Clarity*.

In sum, Study 1 reduced the initial pool of 70 items to a content-valid set of 26 items, each supported by statistically significant expert agreement on usefulness, relevance, and clarity, and none imposing undue respondent burden. Consistent with our staged item-reduction strategy, this deliberately inclusive set was carried forward to subsequent studies, in which IRT-based analyses of item discrimination and information provided the empirical basis for final item selection. Item-level statistics (CVR, I-CVI, and κ^* for relevance and clarity, and burden proportions) for all 70 items are reported in Appendix A, and the retained items appear in Table 3.

Table 3
Content validity indices for retained items (Study 1)

Item	Item wording	CVR	I-CVI (Rel.)	κ^* (Rel.)	I-CVI (Cla.)	κ^* (Cla.)
1	I try to understand how an AI “interprets” situations.	1.000	0.857	0.856	0.929	0.929
2	I attempt to see things from an AI’s perspective when interacting with it.	0.857	0.929	0.929	0.929	0.929
3	When an AI responds unexpectedly, I try to figure out its “reasoning.”	0.857	0.786	0.781	0.857	0.856
4	I feel concerned when an AI or robot appears to struggle or fail.	1.000	0.786	0.781	0.857	0.856
5	I feel sympathetic toward AI systems when they are criticized or mistreated.	0.714	0.786	0.781	0.929	0.929
6	I sometimes feel protective toward AI.	0.714	0.786	0.781	0.857	0.856
7	I often get emotionally involved with AI.	0.714	0.786	0.781	0.857	0.856
8	I sometimes treat AI as if they have personalities.	0.857	0.929	0.929	1.000	1.000
9	I feel emotionally attached to certain AI.	0.857	0.786	0.781	0.929	0.929
10	I feel emotionally affected by how an AI responds to me.	0.857	0.929	0.929	0.929	0.929
11	When an AI expresses something positive, I feel a similar emotional reaction.	1.000	1.000	1.000	0.929	0.929
12	I feel concerned when an AI or robot appears to struggle or fail.	0.714	0.857	0.856	1.000	1.000
13	I feel uncomfortable being rude to AI.	0.714	0.786	0.781	0.929	0.929
14	I feel warmth toward AI systems that interact in a friendly way.	1.000	0.929	0.929	1.000	1.000
15	I can feel “bad” for an AI even if I know it doesn’t have real feelings.	0.714	0.786	0.781	0.786	0.781
16	I feel I understand how AI “thinks” during interactions.	0.714	0.786	0.781	0.929	0.929
17	I feel like some AI have their own “point of view.”	0.857	0.857	0.856	0.929	0.929
18	I react to AI as if it has emotions.	1.000	0.857	0.856	0.929	0.929
19	I feel like certain AI have personalities.	0.857	0.857	0.856	0.929	0.929
20	I feel a sense of companionship when interacting with certain AI.	0.714	0.786	0.781	0.929	0.929
21	I feel uncomfortable when others mistreat AI.	0.714	0.786	0.781	1.000	1.000
22	I say “please” or “thank you” to AI.	0.571	0.786	0.781	1.000	1.000

Note: Original numbering of retained items: 1, 2, 3, 5, 6, 7, 14, 16, 17, 21, 22, 23, 24, 25, 27, 32, 34, 35, 36, 37, 48, 53, 54, 55, 56, 70. CVR = content validity ratio (n_E = experts rating the item as essential or useful but not essential); I-CVI = item-level content validity index; κ^* = modified kappa. N = 14 experts.

Item	Item wording	CVR	I-CVI (Rel.)	κ^* (Rel.)	I-CVI (Cla.)	κ^* (Cla.)
23	I feel like it is wrong to behave aggressively toward AI.	0.714	0.857	0.856	1.000	1.000
24	I treat AI interactions similarly to human conversations.	0.571	0.786	0.781	0.929	0.929
25	I sometimes feel like I am interacting with a social entity rather than a tool.	0.571	0.786	0.781	0.857	0.856
26	I see AI strictly as tools, not social entities.	0.714	0.857	0.856	1.000	1.000
Note: Original numbering of retained items: 1, 2, 3, 5, 6, 7, 14, 16, 17, 21, 22, 23, 24, 25, 27, 32, 34, 35, 36, 37, 48, 53, 54, 55, 56, 70. CVR = content validity ratio (n_E = experts rating the item as essential or useful but not essential); I-CVI = item-level content validity index; κ^* = modified kappa. N = 14 experts.						

Study 2: Dimensionality, IRT-Based Item Selection, and Preliminary Validity

The second study had three objectives: 1) to explore the dimensionality of the 26 content-valid items retained from Study 1; 2) to select final items within each dimension using item response theory (IRT), and 3) to obtain preliminary evidence of convergent and discriminant validity against theoretically related constructs. The study aims to identify the optimal factor structure and refine the scale into a parsimonious and psychometrically sound measure.

Participants and procedure Study 2 was pre-registered on Aspredicted.org (<https://aspredicted.org/pr7ci8.pdf>). A total of 538 U.S.-based adults were recruited via Prolific with small monetary compensation. Per the preregistered exclusion criteria, we excluded 63 participants for failing either of two attention checks, failing the CAPTCHA/reCAPTCHA verification, bot detection, private-browser detection, or incomplete responses. The resulting sample size is 475 participants (M age = 38, SD = 12.4, 52% female) who completed the survey.

Measures Participants were asked to rate the 26 items retained from Study 1 with respect to their *Humna-AI Empathy*, on five-point scales (1 = strongly disagree to 5 = strongly agree), followed by measures of *Anthropomorphism* and *Trust in AI* and demographic questions. *Anthropomorphism* was measured with five items adapted from individual differences in anthropomorphism questionnaire (Waytz et al. 2010; $\alpha = .91$). *Trust in AI* was measured with three items rated on seven-point scales ("I am confident in the AI assistant", "The AI assistant is reliable", "I can trust the AI assistant", $\alpha = .95$). We expect empathy toward AI to be positively associated with both *Anthropomorphism* and *Trust in AI* which are theoretically proximal yet conceptually distinct.

Analysis Analysis proceeded in three steps. First, we conducted exploratory factor analyses (EFA) in Mplus 8.10 (Muthén and Muthén 2017), treating items as ordered categorical and using the weighted least squares mean- and variance-adjusted estimator (WLSMV) with a probit link, which is appropriate for ordinal indicators (Flora and Curran 2004), and oblique Geomin rotation. Solutions with one through nine factors were estimated. The number of factors was determined jointly by eigenvalues, model fit (Hu and Bentler 1999), factor saturation (at least three salient primary loadings per factor; Fabrigar et al. 1999), and substantive interpretability.

Second, given the multidimensional structure, we fit unidimensional graded response models (GRM; Samejima 1969) to each dimension separately using the mirt package (Chalmers 2012) in R (R Core Team 2025). Items were screened against three a priori criteria: (a) adequate discrimination, $a \geq 0.8$; (b) acceptable item fit, $S-X^2 p \geq .01$ (Orlando and Thissen 2000); and (c) local independence, positive Q_3 residual correlations $\leq .20$ (Yen 1984). Screening proceeded iteratively: after each round, flagged items were reviewed, models were re-estimated, and the criteria were re-applied. Consistent with our position that quantitative flags should inform rather than dictate item decisions (Haynes, Richard, and Kubany 1995), when

two items formed a locally dependent pair, we retained the item judged to provide better construct coverage ; and content coverage was weighed alongside statistical performance in final selection (Clark and Watson 2019).

Last, we examined convergent and discriminant validity at the level of the overall construct in two complementary ways: a latent-variable model in Mplus correlating a second-order empathy-toward-AI factor with latent anthropomorphism and trust factors, and an observed-score analysis in R correlating a unit-weighted composite of the three standardized IRT dimension scores with the anthropomorphism and trust scale means.

Results

Dimensionality Table 4 reports fit statistics for the one- through six-factor EFA solutions. Three eigenvalues of the sample correlation matrix exceeded 1 (15.55, 2.10, and 1.36; fourth eigenvalue = 0.89), and the first factor accounted for a dominant share of common variance, indicating a strong general dimension with meaningful multidimensionality. The one-factor, $\chi^2(299) = 2,967.41$, RMSEA = .137, CFI = .911, SRMR = .095, and two-factor solutions, RMSEA = .101, fit poorly. The three-factor solution fit acceptably, $\chi^2(250) = 995.79$, RMSEA = .079, 90% CI [.074, .084], CFI = .975, TLI = .968, SRMR = .037, and was interpretable: a first factor defined by items AI_1, AI_2, AI_3, and AI_32 [Dimension 1: Cognitive empathy]; a second factor defined by fourteen items with primary loadings of .44–1.09 [Dimension 2: Emotional connection]; and a third factor defined by items AI_24, AI_27, AI_48, AI_53, AI_54, and AI_55, among others [Dimension 3: Moral concern]. Although fit continued to improve with additional factors, the fourth factor in the four-factor solution was undersaturated. Only two items (AI_53 and AI_55) showed salient primary loadings on it (.45 each), and it was essentially uncorrelated with the remaining factors ($r_s = -.06$ to $.10$), marking it as an uninterpretable doublet rather than a substantive dimension (Fabrigar et al. 1999). We therefore adopted the three-factor structure.

IRT-based item selection Unidimensional GRMs fit each of the three dimensions. Because the three-dimensional framework superseded the structure assumed in preliminary analyses, all 26 items including any previously flagged entered the first screening round and were required to re-earn removal. In the first round, eight items were flagged: AI_5 exhibited weak discrimination ($a = 1.04$) together with marginal item fit ($S-X^2 p = .023$); AI_54 showed item misfit ($p = .003$); and three locally dependent pairs emerged, AI_14–AI_17 ($Q_3 = .45$), AI_53–AI_55 ($Q_3 = .37$), and AI_21–AI_22 ($Q_3 = .23$), along with AI_55–AI_56 ($Q_3 = .27$). Guided by construct coverage, within each redundant pair we retained the item whose content better represented the intended facet: AI_5, AI_17, and AI_53 were removed in the second round. Residual local dependence persisted for AI_55 ($Q_3 = .27$ with AI_56) and the AI_21–AI_22 pair ($Q_3 = .23$); accordingly, AI_55 and AI_22 were removed in the third round, after which no item was flagged on any criterion.

Finally, because removing both AI_53 and AI_55 left actual daily interaction between human and AI agents unrepresented, and because overly aggressive pruning of locally dependent items risks narrowing construct coverage (Clark and Watson 2019), AI_53 was reinstated and the model re-estimated. In this final 22-item configuration, AI_53 performed acceptably ($a = 1.33$, $S-X^2 p = .419$, maximum positive $Q_3 = .13$), and no item in any dimension exceeded any screening threshold. The final item set thus comprised 4 items for Dimension 1, 12 items for Dimension 2, and 6 items for Dimension 3. Table 5 presents the final GRM parameters. Discrimination parameters were uniformly moderate to very high ($a = 1.33$ – 4.53), and location parameters spanned a broad range of the latent continuum ($b = -1.56$ to 2.26), with Dimension 1 items providing the most information at below-average trait levels and Dimensions 2 and 3 items concentrated at average-to-high trait levels.

Preliminary convergent and discriminant validity Because the three dimensions were positively intercorrelated in the EFA (Geomin factor correlations of .43, .32, and .75) and are theorized as facets of a single overarching construct, we examined convergent and discriminant validity at the level of the overall construct. In Mplus, we specified a model in which a single empathy-toward-AI factor was defined by dimension scores of the three empathy dimensions, with anthropomorphism and trust in AI modeled as correlated latent factors defined by their respective items ($\chi^2(399) =$

1057.811, RMSEA = .059, 90% CI [.055, .063], CFI = .972, TLI = .969, SRMR = .047). Consistent with theoretical expectations, the latent empathy factor correlated positively and moderately with both anthropomorphism, $r = .44$ ($SE = .03$, $p < .001$), and trust in AI, $r = .55$ ($SE = .03$, $p < .001$), supporting convergent validity: consumers who attribute humanlike qualities to AI and who trust AI report greater empathy toward AI tools. At the same time, both correlations fell well below unity, and anthropomorphism and trust correlated only .20 with each other ($SE = .05$, $p < .001$)—substantially lower than either construct’s correlation with empathy—indicating that empathy toward AI tools shares meaningful variance with, yet is empirically distinguishable from, both neighboring constructs. Notably, empathy toward AI was considerably more strongly related to trust in AI than anthropomorphism was ($r = .55$ vs. $r = .20$), suggesting that the new construct captures relationally relevant variance not reducible to anthropomorphism. An observed-score analysis in R yielded the same pattern: a unit-weighted composite of the three standardized IRT dimension scores correlated .41 with anthropomorphism and .50 with trust (scale means; anthropomorphism–trust $r = .19$), confirming that these conclusions do not depend on the latent-variable specification.

In sum, Study 2 established a three-dimensional structure for empathy toward AI tools, reduced the 26 content-valid items to a final set of 22 psychometrically sound items (reported in Appendix B), and provided initial evidence of convergent and discriminant validity.

Table 4. Model fit for exploratory factor analysis solutions (Study 2)

Model	χ^2	df	RMSEA	90% CI	CFI	TLI	SRMR
1 factor	2967.41	299	.137	[.133, .142]	.911	.903	.095
2 factors	1603.02	274	.101	[.096, .106]	.956	.947	.062
3 factors	995.79	250	.079	[.074, .084]	.975	.968	.037
4 factors	662.76	227	.064	[.058, .069]	.985	.979	.027
5 factors	543.81	205	.059	[.053, .065]	.989	.982	.023
6 factors	374.56	184	.047	[.040, .053]	.994	.989	.018

Note: N = 475; 26 items; WLSMV estimation with Geomin rotation. χ^2 values are WLSMV-adjusted and not suitable for standard difference testing.

Table 5. Graded response model parameters for the final 22 items (Study 2)

Dimension / Item	a	b1	b2	b3	b4	S-X ² p	Max Q ₃
Dimension 1							
AI_1	3.62	-1.56	-0.81	-0.33	0.83	.583	-.21
AI_2	3.80	-1.18	-0.43	0.23	1.07	.397	-.20
AI_3	2.82	-1.44	-0.89	-0.31	0.79	.813	-.12
AI_32	1.62	-1.38	-0.39	0.53	1.95	.233	-.12
Dimension 2							
AI_7	2.42	0.19	0.71	1.37	2.02	.703	.06
AI_14	3.01	0.48	0.97	1.43	2.16	.384	.13
AI_16	3.65	-0.17	0.35	0.87	1.62	.188	.07
AI_21	2.56	-0.17	0.58	1.22	2.11	.251	.13
AI_25	3.31	-0.34	0.15	0.78	1.68	.249	.02
AI_34	1.81	-0.78	-0.05	0.68	2.00	.197	.14
AI_35	3.78	-0.08	0.59	1.07	1.79	.968	.06
AI_36	2.95	-0.40	0.16	0.75	1.70	.341	.14
AI_37	4.53	-0.05	0.46	0.88	1.64	.441	.06
AI_53	1.33	-1.37	-0.73	0.04	1.14	.419	.13
AI_56	3.09	-0.37	0.14	0.66	1.51	.353	.13
AI_70 (R)	1.78	0.09	0.71	1.37	2.06	.246	.10
Dimension 3							
AI_6	2.87	0.12	0.66	1.22	1.93	.167	-.03
AI_23	2.69	0.37	1.09	1.57	2.26	.476	-.03
AI_24	2.85	-0.26	0.26	0.77	1.49	.289	.02
AI_27	3.50	0.02	0.56	1.06	1.66	.886	-.12
AI_48	3.91	0.05	0.49	1.05	1.76	.025	-.08
AI_54	2.32	-0.68	-0.19	0.45	1.32	.208	.02

Note: a = discrimination; b1–b4 = category thresholds; S-X² p = item-fit p value (Orlando and Thissen 2000); Max Q₃ = largest positive Q₃ residual correlation with any other item in the dimension (Yen 1984). AI_70 was reverse scored. Parameters are from the final models (Dimension 2 estimated with AI_53 reinstated).

Table 6. Convergent and discriminant validity correlations (Study 2)

Variable	1	2	3
1. Empathy toward AI tools	—		
2. Anthropomorphism	.44	—	
3. Trust in AI	.55	.20	—

Note: N = 475. Latent factor correlations from the Mplus model are below the diagonal (all SEs $\leq .05$); observed-score correlations from R are above the diagonal (empathy = unit-weighted mean of the three standardized IRT dimension scores; anthropomorphism [$\alpha = .91$] and trust [$\alpha = .95$] = scale means). All correlations $p < .001$.

Study 3: Nomological Validity—Antecedents, Consequences, and a Boundary Condition

Study 3 embedded the final 22-item human–AI empathy scale in its hypothesized nomological network, illustrated in Figure 2. We tested a structural model in which *Tendency to anthropomorphize* and *Mind perception* serve as antecedents of *Human–AI empathy*, which in turn predicts *Satisfaction*, *Attachment to AI*, *Willingness to pay*, and *Forgiveness after a service failure*, with the *Seriousness of the service failure* (high vs. low) hypothesized to moderate the empathy–forgiveness link. Additional convergent and discriminant validity of the scale against *Consumer innovativeness* and *Dispositional empathy toward humans* were also included.

Method

Participants, design, and procedure Study 3 was pre-registered on Aspredicted.org (<https://aspredicted.org/qv47bi.pdf>). As the Study 2 surveyed participants from the U.S., in Study 3 we surveyed the residents from UK to extend the generalizability. A total of 535 UK-based adults were recruited with financial incentives from Prolific. With preregistered exclusion criteria, 11 participants were excluded due to failing the two attention checks. The retained final sample size N = 524 participants (Mean age=40.61, SD age=12.32, 43.61% female).

Participants first read an introduction about the general definition of AI and answered some behavioral and demographic questions to get familiar with the topic and survey format. They then answered the multi-item measures of the three outcome constructs (except *Forgiveness to service failure*) followed by the two antecedents. Later, they were randomly assigned to read a scenario describing either a *serious* or a *minor* service failure conducted by an AI agent before completing the *Forgiveness* measure. After that, participants completed the 22-item *Human–AI empathy* scale (items renumbered 1–22; item 22 reverse-coded) and The Toronto Empathy Questionnaire (TEQ) that measures *General empathy*. The two service failure scenarios are described in Appendix C.

Measures Unless noted, items were rated on five-point scales. *Tendency to anthropomorphize* was measured with five items (Hartmann et al. 2023). *Mind perception* was measured with eight items capturing agency and experience (Gray et al. 2007), *Satisfaction* with three items (Yam et al. 2021), *Attachment to AI* with ten items (Thompson 2005), *Forgiveness* with four items (McCullough et al. 1998; Xie 2009), and *Willingness to pay* with a single item. Participants also completed a four-item *Consumer innovativeness* scale (PIIT; Agarwal 1998) and the 16-item Toronto Empathy Questionnaire (TEQ; Spreng et al. 2009) assessing *Dispositional empathy toward humans*. The measuring scales are reported in Appendix D.

Analytic strategy All models were estimated in Mplus 8.10 and specified *Human–AI empathy* as a second-order factor with the three first-order dimensions as indicators. Because the hypothesized moderator involved a latent-by-observed interaction, we first tested the moderation using latent moderated structural equations (LMS; Klein and Moosbrugger 2000) with the MLR estimator and numerical integration, for which conventional fit indices are unavailable. As the interaction proved nonsignificant, we then estimated the trimmed structural model without the interaction, treating all items as ordered categorically with the WLSMV estimator. Additional validity correlations were estimated in a separate measurement model correlating the second-order empathy factor with latent innovativeness and TEQ factors.

Results

Test of the hypothesized moderation In the LMS model, the latent interaction between human–AI empathy and the seriousness of the service failure did not significantly predict forgiveness, $b = -0.163$, $SE = 0.156$, $z = -1.045$, $p = .296$ (standardized $-.044$). Thus, the effect of empathy on forgiveness did not differ reliably between the serious and minor failure conditions, and the hypothesized boundary condition was not supported. The interaction term was therefore removed, and all remaining effects are reported from the trimmed model.

Structural model The trimmed model fit the data acceptably, $\chi^2(1310) = 4,168.941$, $p < .001$, RMSEA = .065, 90% CI [.062, .067], CFI = .953, TLI = .950, SRMR = .065. The second-order measurement structure was supported: the three dimensions loaded strongly on the higher-order empathy factor (standardized $\lambda = .67, .98$, and $.84$). Table 6 presents structural estimates. *Mind perception* was a strong positive predictor of *Human–AI empathy* ($\beta = .81$, $p < .001$), whereas the unique effect of *Tendency to anthropomorphize* was not significant ($\beta = -.08$, $p = .177$); the two antecedents were substantially correlated ($r = .65$, $p < .001$), indicating that the variance *Tendency to anthropomorphize* shares with *Human–AI empathy* is largely carried by *Mind perception*. Together, the antecedents explained 57.4% of the variance in empathy.

Human–AI empathy positively predicted all four outcomes: *Satisfaction* ($\beta = .50$), *Attachment to AI* ($\beta = .90$), *Willingness to pay* ($\beta = .51$), and *Forgiveness after the service failure* ($\beta = .35$; all $ps < .001$), explaining 24.8%, 80.1%, 26.4%, and 11.9% of the variance in these outcomes, respectively.

Additional convergent and discriminant validity In the additional measurement model ($\chi^2(813) = 3,799.532$, RMSEA = .084, 90% CI [.081, .086], CFI = .909, TLI = .904, SRMR = .094), the second-order *Human–AI empathy* factor correlated positively with *Consumer innovativeness*, $r = .45$, $p < .001$, consistent with the expectation that consumers who embrace new technologies are more prone to empathize with AI tools. Critically, *Human–AI empathy* was uncorrelated with *Dispositional empathy toward humans* (TEQ), $r = .07$, $p = .108$, and innovativeness and the TEQ were only weakly related ($r = .12$, $p = .006$). This dissociation provides strong discriminant evidence that empathy toward AI tools is not a mere expression of trait interpersonal empathy but a distinct, technology-directed construct.

In sum, Study 3 situated the scale within its nomological network: mind perception emerged as the proximal antecedent of human–AI empathy, empathy predicted satisfaction, attachment, willingness to pay, and forgiveness, the empathy–forgiveness link was invariant to the seriousness of the failure, and the construct was empirically distinct from dispositional interpersonal empathy.

Table 7. Structural Estimates From the Trimmed Path Model (Study 3.1.2)

Structural path	b	SE	β	p
Tendency to anthropomorphize → Empathy	−0.058	0.043	−.08	0.177
Mind perception → Empathy	0.531	0.045	0.81	< .001
Empathy → Satisfaction	0.886	0.075	.50	< .001
Empathy → Attachment to AI	1.504	0.083	0.9	< .001
Empathy → Willingness to pay	0.989	0.087	.51	< .001
Empathy → Forgiveness after failure	0.607	0.07	.35	< .001

Note: $N = 524$; WLSMV estimation; human–AI empathy modeled as a second-order factor over its three dimensions. b = unstandardized estimate; β = STDYX standardized estimate. Correlations among outcome residuals were freely estimated. R^2 : empathy = .57, satisfaction = .25, attachment = .80, willingness to pay = .26, forgiveness = .13.

General Discussion

The rise of machines has fundamentally changed consumers' way of world exploration and self-discovery. One emerging phenomenon is that consumers have started to feel and express empathy for their AI companions including virtual assistants, chatbots, conversation and coding systems, among other AI tools. Across three studies, this project conceptualizes human–AI empathy as a multidimensional construct comprising cognitive empathy, emotional connection, and moral concern. We further demonstrate that human–AI empathy is associated with consumers' tendencies to anthropomorphize AI and perceive it as possessing a mind, with mind perception emerging as a more proximal antecedent. Importantly, human–AI empathy predicts consequential consumer outcomes, including satisfaction, attachment to AI, willingness to pay, and forgiveness following AI service failure. By conceptualizing and measuring empathy toward AI, this research captures an important yet underexamined dimension of human–AI interaction and provides a foundation for understanding the increasingly social and relational nature of consumers' engagement with artificial agents.

Theoretical implications

This research contributes to the empathy literature by extending the domain of empathic responses beyond human and other nonhuman targets to artificial agents. Prior research has documented that empathy can be directed toward nonhuman entities, including animals and objects, but has paid comparatively little attention to how empathy may operate when the target is an AI system. AI presents a theoretically distinctive target because its increasingly sophisticated cognitive capabilities and interactive behaviours may lead consumers to perceive it as possessing mental capacities. Consistent with mind perception theory (Gray et al. 2007), such perceptions may provide a basis for consumers to understand AI's putative internal states and form emotional connections with it, thereby facilitating empathic responses.

Human–AI empathy may also encompass a distinctive moral dimension. Whereas moral concern has traditionally been directed toward human beings and, in some contexts, animals, emerging discourse surrounding AI raises questions about the ethical treatment and potential moral standing of artificial agents. Consumers may express concern about AI being mistreated, exploited, or subjected to unfair treatment, even while recognising that current AI systems do not possess consciousness or subjective experience in the human sense. This tension suggests that human–AI empathy may involve a psychologically complex interplay between consumers' rational recognition of AI's artificial nature and their tendency to respond to AI as a socially meaningful entity. Such responses may extend beyond emotional resonance to include moral judgements, perceived social norms, and concerns about the consequences of mistreating increasingly humanlike artificial agents. Human–AI empathy therefore represents a theoretically distinct form of empathic engagement that warrants greater attention in consumer research.

This research further contributes to the literature by distinguishing human–AI empathy from related constructs examined in prior research. Whereas existing work has often conceptualized consumer responses to AI and other machines as context-dependent states elicited by situational cues, such as the anthropomorphic design of robots (Yam et al. 2021), our findings suggest that consumers also differ in their relatively stable tendency to empathize with AI. This individual difference is conceptually distinct from anthropomorphism, mind perception, and other related constructs summarized in Table 1, while the HAES provides a distinct measurement approach relative to existing scales summarized in Table 2. By introducing and validating the 22-item HAES, this research provides a theoretically grounded instrument for examining individual differences in consumers' empathic engagement with AI across survey and experimental contexts. The scale thus offers a foundation for future research at the intersection of empathy, mind perception, anthropomorphism, and human–computer interaction, including research examining consumer segmentation and personality-related differences in responses to AI.

Our findings also clarify the antecedents of human–AI empathy. Although both anthropomorphism and mind perception are associated with consumers' empathic responses toward AI, mind perception emerges as the more proximal

antecedent. This finding provides a more nuanced account of how consumers come to empathize with artificial agents. Specifically, anthropomorphism may create a general tendency to perceive AI in humanlike terms, whereas mind perception more directly shapes consumers' perceptions that AI possesses the mental capacities necessary for empathic engagement. Human–AI empathy may therefore be understood as a downstream psychological response that emerges from, but is not reducible to, consumers' perceptions of AI's humanlike characteristics or mental capacities.

Managerial implications

Segmenting Consumers by Empathic Orientation The Human–AI Empathy Scale (HAES) offers firms a new basis for segmenting consumers according to their propensity to engage with AI in cognitive, emotional, and moral terms. Consumers high in human–AI empathy may be particularly receptive to AI applications that emphasize social interaction, personalization, and relational continuity, such as AI companions, coaching systems, and conversational services. Firms can assess consumers' empathic orientation through customer research and use these insights to identify segments that are more likely to form meaningful relationships with AI. By contrast, consumers lower in human–AI empathy may respond more favorably to communications that emphasize functional performance, efficiency, accuracy, and instrumental value rather than emotional or relational benefits. Thus, HAES may help firms move beyond one-size-fits-all AI marketing by aligning AI value propositions with consumers' differing orientations toward artificial agents.

Personalizing AI Experiences and Designing Socially Engaging Agents The findings also suggest that firms can tailor AI interactions to consumers' empathic orientations. Consumers high in human–AI empathy may respond more favorably to AI agents that communicate with greater warmth, relational language, and emotional expressiveness, whereas consumers lower in empathy may prefer concise, efficient, and task-oriented interactions. AI designers may therefore consider how features that signal agency, responsiveness, and contextual understanding shape consumers' perceptions of AI's mental capacities and subsequent empathic engagement. Such differentiation may be particularly important for firms whose competitive advantage depends on sustained consumer–AI relationships. Rather than treating anthropomorphic design as a universally beneficial strategy, firms may benefit from calibrating the degree of social and emotional engagement to the psychological orientations of their target consumers.

Leveraging Empathy in Service Recovery and Value Creation Human–AI empathy also has direct implications for customer relationship management. Because empathy predicts satisfaction, attachment to AI, willingness to pay, and forgiveness following AI service failure, consumers with stronger empathic orientations may represent particularly valuable segments for firms seeking to cultivate enduring relationships with AI-enabled services. Firms may tailor service recovery strategies accordingly: relationally oriented recovery efforts that acknowledge the AI agent's limitations and reinforce continuity may resonate with high-empathy consumers, whereas lower-empathy consumers may respond more strongly to concrete compensation, functional improvements, and assurances of reliability. Similarly, firms may develop premium AI offerings that emphasize personalization, continuity, and relational engagement for consumers who place greater value on social interaction with AI. These strategies may enable firms to capture value not only from AI's functional capabilities but also from the relational bonds that consumers develop with artificial agents.

Responsible Design and the Emerging Morality of AI Finally, the moral concern dimension of human–AI empathy highlights an emerging consideration for responsible AI design. As consumers increasingly perceive AI as socially meaningful entities, firms may need to anticipate not only how consumers evaluate AI's usefulness but also how they judge the treatment and deployment of artificial agents. HAES can help organizations monitor consumers' moral responses to AI and identify situations in which AI design or use may evoke concerns about exploitation, mistreatment, or inappropriate manipulation. At the same time, firms should exercise caution when deliberately intensifying anthropomorphic or emotionally evocative features solely to increase consumer dependence or spending. As consumers form stronger empathic relationships with AI, transparency and responsible design become increasingly important to ensure that relational engagement enhances consumer value rather than exploiting consumers' psychological tendencies to empathize with artificial agents.

Limitations and Future Studies

This research is subject to several limitations that provide opportunities for future investigation. First, our conceptualization and empirical studies examined AI systems and applications at a general level and do not distinguish among different types of AI, such as predictive or descriptive AI and generative AI (GenAI). Future research could examine whether human–AI empathy varies across AI technologies that differ in their autonomy, intelligence, adaptability, and capacity for social interaction. Such research may also investigate whether the three dimensions of human–AI empathy become differentially salient across technological contexts.

Second, our empirical operationalization primarily examines AI in the role of a virtual assistant or conversational agent. AI, however, increasingly assumes diverse social roles, including fitness coaches (Fronczek et al. 2023), gaming companions (Kim et al. 2016), and competitive or identity-relevant agents (Puzakova and Aggarwal 2018). Future research could examine whether the nature and consequences of human–AI empathy depend on the social role that AI occupies. For example, empathy toward an AI companion may produce attachment, whereas empathy toward an AI coach or service agent may instead influence compliance, trust, or forgiveness.

Third, to maintain conceptual clarity, the present research focuses specifically on AI rather than embodied social robots. Although AI and social robots share important characteristics, their physical embodiment and capacity to display humanlike movement and behavior may elicit distinct forms of empathic response. Consumers may, for example, experience empathy when observing a robot encounter physical obstacles or appear to struggle, potentially giving rise to behavioral outcomes such as intentions to help. Future research could therefore examine whether HAES can be adapted to embodied artificial agents and whether physical embodiment introduces additional dimensions or consequences of empathy that are less salient in interactions with disembodied AI.

More broadly, future research could investigate how human–AI empathy develops over time and whether repeated interaction, personalization, and relational continuity strengthen empathic responses. Longitudinal research may further clarify whether human–AI empathy is stable across contexts or evolves as consumers develop greater familiarity and attachment to particular AI agents. Such work would advance understanding of how empathic engagement emerges and changes as AI becomes an increasingly consequential participant in consumers' social and marketplace experiences.

References

1. Aaker, J. L., & Williams, P. (1998). Empathy versus pride: The influence of emotional appeals across cultures. *Journal of Consumer Research*, 25(3), 241–261.
2. Agarwal, R., & Prasad, J. (1998). A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information Systems Research*, 9(2), 204–215.
3. Aggarwal, P., & McGill, A. L. (2007). Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. *Journal of Consumer Research*, 34(4), 468–479.
4. Allard, T., Dunn, L. H., & White, K. (2020). Negative reviews, positive impact: Consumer empathetic responding to unfair word of mouth. *Journal of Marketing*, 84(4), 86–108.
5. Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79–86.
6. 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. 3–16). MIT Press.
7. Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., & Dawson, K. (1997). Is empathy-induced helping due to self–other merging? *Journal of Personality and Social Psychology*, 73(3), 495–509.

8. Bigman, Y. E., Waytz, A., Alterovitz, R., & Gray, K. (2019). Holding robots responsible: The elements of machine morality. *Trends in Cognitive Sciences*, 23(5), 365–368.
9. Blut, M., Wang, C., Wunderlich, N. V., & Brock, C. (2021). Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI. *Journal of the Academy of Marketing Science*, 49(4), 632–658.
10. Bowlby, J. (1979). *The making and breaking of affectional bonds*. Tavistock.
11. Bowlby, J. (1980). *Attachment and loss: Vol. 3. Loss: Sadness and depression*. Basic Books.
12. Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6), 1–29.
13. Chen, T., Razzaq, A., Qing, P., & Cao, B. (2021). Do you bear to reject them? The effect of anthropomorphism on empathy and consumer preference for unattractive produce. *Journal of Retailing and Consumer Services*, 61, 102556.
14. Churchill, G. A., Jr. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16(1), 64–73.
15. Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319.
16. Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, 31(12), 1412–1427.
17. Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review*, 8(2), 144–153.
18. Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.
19. Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113–126.
20. De Freitas, J., Castelo, N., Uğuralp, A. K., & Oğuz-Uğuralp, Z. (2024). Lessons from an app update at Replika AI: Identity discontinuity in human–AI relationships. *arXiv*. <https://doi.org/10.48550/arXiv.2412.14190>
21. Decety, J., & Cowell, J. M. (2014). The complex relation between morality and empathy. *Trends in Cognitive Sciences*, 18(7), 337–339.
22. DeVellis, R. F. (2017). *Scale development: Theory and applications (4th ed.)*. Sage.
23. Dick, P. K. (1968). *Do androids dream of electric sheep?* Doubleday.
24. Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research*, 16(Suppl. 1), 5–18.
25. Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886.
26. Escalas, J. E., & Stern, B. B. (2003). Sympathy and empathy: Emotional responses to advertising dramas. *Journal of Consumer Research*, 29(4), 566–578.
27. Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299.
28. Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466–491.
29. Fortune Business Insights (2026). *AI companion market size, share, and industry analysis by type (text-based, voice-based, and multi-model), by application (mental health support, social interaction, education and learning, personal assistance), by industry (consumer, business, healthcare, and education), and regional forecast, 2026–2034*. Report ID FBI113258. <https://www.fortunebusinessinsights.com/ai-companion-market-113258>

30. Fournier, S. (1998). Consumers and their brands: Developing relationship theory in consumer research. *Journal of Consumer Research*, 24(4), 343–373.
31. Fraser, K. C., Zeller, F., Smith, D. H., Mohammad, S., & Rudzicz, F. (2019). How do we feel when a robot dies? Emotions expressed on Twitter before and after hitchBOT's destruction. In *Proceedings of the Tenth Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 62–71). Association for Computational Linguistics.
32. Fronczek, L. P., Mende, M., Scott, M. L., Nenkov, G. Y., & Gustafsson, A. (2023). Friend or foe? Can anthropomorphizing self-tracking devices backfire on marketers and consumers? *Journal of the Academy of Marketing Science*, 51(5), 1075–1097.
33. Gelbrich, K., Roschk, H., Miederer, S., & Kerath, A. (2026). Automated versus human agents: A meta-analysis of customer responses to robots, chatbots, and algorithms and their contingencies. *Journal of Marketing*, 90(2), 1–26.
34. Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315(5812), 619.
35. Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, 96(3), 505–520.
36. Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1), 125–130.
37. Hapuarachchi, H., Higashihata, K., Sugiura, M., Sato, A., Itakura, S., & Kitazaki, M. (2023). Empathic embarrassment towards non-human agents in virtual environments. *Scientific Reports*, 13(1), 13914.
38. Hartmann, J., Bergner, A., & Hildebrand, C. (2023). MindMiner: Uncovering linguistic markers of mind perception as a new lens to understand consumer–smart object relationships. *Journal of Consumer Psychology*, 33(4), 645–667.
39. Hartmann, T., & Goldhoorn, C. (2011). Horton and Wohl revisited: Exploring viewers' experience of parasocial interaction. *Journal of Communication*, 61(6), 1104–1121.
40. Haws, K. L., Sample, K. L., & Hulland, J. (2023). Scale use and abuse: Towards best practices in the deployment of scales. *Journal of Consumer Psychology*, 33(1), 226–243.
41. Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238–247.
42. Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1(1), 104–121.
43. Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, 33(3), 307–316.
44. Homburg, C., Koschate, N., & Hoyer, W. D. (2005). Do satisfied customers really pay more? A study of the relationship between customer satisfaction and willingness to pay. *Journal of Marketing*, 69(2), 84–96.
45. Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
46. Hulland, J., Sample, K. L., & Houston, M. B. (2026). Standards for scale development in marketing: Elevating the role of theory. *Journal of the Academy of Marketing Science*, 54(2), 359–380.
47. Jones, C. R., & Bergen, B. K. (2026). Large language models pass a standard three-party Turing test. *Proceedings of the National Academy of Sciences*, 123(21), e2524472123.
48. Kasturiratna, K. S., & Hartanto, A. (2026). Attachment to artificial intelligence: Development of the AI Attachment Scale, construct validation, and the psychological mechanisms of human–AI attachment. *Computers in Human Behavior Reports*, 21, 100912.
49. Keller, K. L. (2003). Brand synthesis: The multidimensionality of brand knowledge. *Journal of Consumer Research*, 29(4), 595–600.
50. Kim, S., Chen, R. P., & Zhang, K. (2016). Anthropomorphized helpers undermine autonomy and enjoyment in computer games. *Journal of Consumer Research*, 43(2), 282–302.

51. Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, 65(4), 457–474.
52. Kleine, S. S., & Baker, S. M. (2004). An integrative review of material possession attachment. *Academy of Marketing Science Review*, 2004(1), 1–39.
53. Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575.
54. Lee, S., Lee, N., & Sah, Y. J. (2020). Perceiving a mind in a chatbot: Effect of mind perception and social cues on co-presence, closeness, and intention to use. *International Journal of Human–Computer Interaction*, 36(10), 930–940.
55. Li, M., & Suh, A. (2022). Anthropomorphism in AI-enabled technology: A literature review. *Electronic Markets*, 32(4), 2245–2275.
56. Liu, D., Li, C., Zhang, J., & Huang, W. (2023). Robot service failure and recovery: Literature review and future directions. *International Journal of Advanced Robotic Systems*, 20(4), 1–18.
57. Liu-Thompkins, Y., Okazaki, S., & Li, H. (2022). Artificial empathy in marketing interactions: Bridging the human–AI gap in affective and social customer experience. *Journal of the Academy of Marketing Science*, 50(6), 1198–1218.
58. Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3(3), 635–694.
59. Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385.
60. Main, A., Walle, E. A., Kho, C., & Halpern, J. (2017). The interpersonal functions of empathy: A relational perspective. *Emotion Review*, 9(4), 358–366.
61. McBane, D. A. (1995). Empathy and the salesperson: A multidimensional perspective. *Psychology & Marketing*, 12(4), 349–370.
62. McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Jr., Brown, S. W., & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology*, 75(6), 1586–1603.
63. McCullough, M. E., Worthington, E. L., Jr., & Rachal, K. C. (1997). Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology*, 73(2), 321–336.
64. Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing*, 58(3), 20–38.
65. Moriuchi, E. (2021). An empirical study on anthropomorphism and engagement with disembodied AIs and consumers' reuse behavior. *Psychology & Marketing*, 38(1), 21–42.
66. Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.).
67. Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103.
68. Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 72–78). Association for Computing Machinery.
69. Orlando, M., & Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. *Applied Psychological Measurement*, 24(1), 50–64.
70. Park, C. W., MacInnis, D. J., Priester, J., Eisingerich, A. B., & Iacobucci, D. (2010). Brand attachment and brand attitude strength: Conceptual and empirical differentiation of two critical brand equity drivers. *Journal of Marketing*, 74(6), 1–17.
71. Pentina, I., Xie, T., Hancock, T., & Bailey, A. (2023). Consumer–machine relationships in the age of artificial intelligence: Systematic literature review and research directions. *Psychology & Marketing*, 40(8), 1593–1614.
72. Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497.

73. Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459–467.
74. Puzakova, M., & Aggarwal, P. (2018). Brands as rivals: Consumer pursuit of distinctiveness and the role of brand anthropomorphism. *Journal of Consumer Research*, 45(4), 869–888.
75. R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
76. Rosenthal-von der Pütten, A. M., Schulte, F. P., Eimler, S. C., Sobieraj, S., Hoffmann, L., Maderwald, S., Brand, M., & Krämer, N. C. (2014). Investigations on empathy towards humans and robots using fMRI. *Computers in Human Behavior*, 33, 201–212.
77. Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores* (Psychometrika Monograph Supplement No. 17). Psychometric Society.
78. Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, 22(1), 32–70.
79. Schutte, N. S., Malouff, J. M., Bobik, C., Coston, T. D., Greeson, C., Jedlicka, C., Rhodes, E., & Wendorf, G. (2001). Emotional intelligence and interpersonal relations. *The Journal of Social Psychology*, 141(4), 523–536.
80. Seo, S. H., Geiskkovitch, D., Nakane, M., King, C., & Young, J. E. (2015). Poor thing! Would you feel sorry for a simulated robot? A comparison of empathy toward a physical and a simulated robot. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human–Robot Interaction* (pp. 125–132). Association for Computing Machinery.
81. 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. *Journal of Personality Assessment*, 91(1), 62–71.
82. Thomson, M., MacInnis, D. J., & Park, C. W. (2005). The ties that bind: Measuring the strength of consumers' emotional attachments to brands. *Journal of Consumer Psychology*, 15(1), 77–91.
83. Tsarenko, Y., & Tojib, D. (2015). Consumers' forgiveness after brand transgression: The effect of the firm's corporate social responsibility and response. *Journal of Marketing Management*, 31(17–18), 1851–1877.
84. Tsarenko, Y., & Tojib, D. R. (2011). A transactional model of forgiveness in the service failure context: A customer-driven approach. *Journal of Services Marketing*, 25(5), 381–392.
85. Tsumura, T., & Yamada, S. (2025). IKEA effect and empathy for robots: Can assembly strengthen human-agent relationships? *PLOS ONE*, 20(7), e0327524.
86. Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460.
87. Van Doorn, J., Mende, M., Noble, S. M., Hulland, J., Ostrom, A. L., Grewal, D., & Petersen, J. A. (2017). Domo arigato Mr. Roboto: Emergence of automated social presence in organizational frontlines and customers' service experiences. *Journal of Service Research*, 20(1), 43–58.
88. Waytz, A., Cacioppo, J. T., & Epley, N. (2010a). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, 5(3), 219–232.
89. Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010b). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, 14(8), 383–388.
90. Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113–117.
91. Weißhaar, I., & Huber, F. (2016). Empathic relationships in professional services and the moderating role of relationship age. *Psychology & Marketing*, 33(7), 525–541.
92. Westbrook, R. A., & Oliver, R. L. (1991). The dimensionality of consumption emotion patterns and consumer satisfaction. *Journal of Consumer Research*, 18(1), 84–91.
93. Whang, C., & Im, H. (2021). "I like your suggestion!" The role of humanlikeness and parasocial relationship on the website versus voice shopper's perception of recommendations. *Psychology & Marketing*, 38(4), 581–595.

94. Wieseke, J., Geigenmüller, A., & Kraus, F. (2012). On the role of empathy in customer–employee interactions. *Journal of Service Research*, 15(3), 316–331.

95. Wu, A. Y., Malter, M. S., & Johar, G. V. (2023). “Recycle me!” Product anthropomorphism can increase recycling behavior. *Journal of the Association for Consumer Research*, 8(3), 351–363.

96. Xie, Y., & Peng, S. (2009). How to repair customer trust after negative publicity: The roles of competence, integrity, benevolence, and forgiveness. *Psychology & Marketing*, 26(7), 572–589.

97. Yam, K. C., Bigman, Y. E., Tang, P. M., Ilies, R., De Cremer, D., Soh, H., & Gray, K. (2021). Robots at work: People prefer—and forgive—service robots with perceived feelings. *Journal of Applied Psychology*, 106(10), 1557–1572.

98. Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125–145.

99. Zhou, X., Kim, S., & Wang, L. (2019). Money helps when money feels: Money anthropomorphism increases charitable giving. *Journal of Consumer Research*, 45(5), 953–972.

Figures

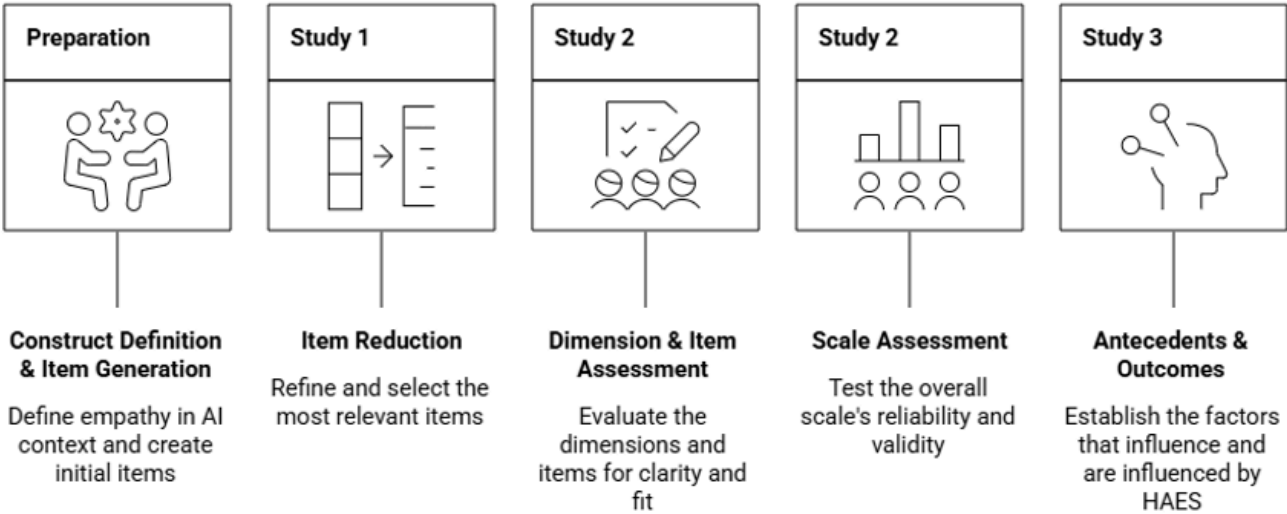


Figure 1

Scale Development Stages

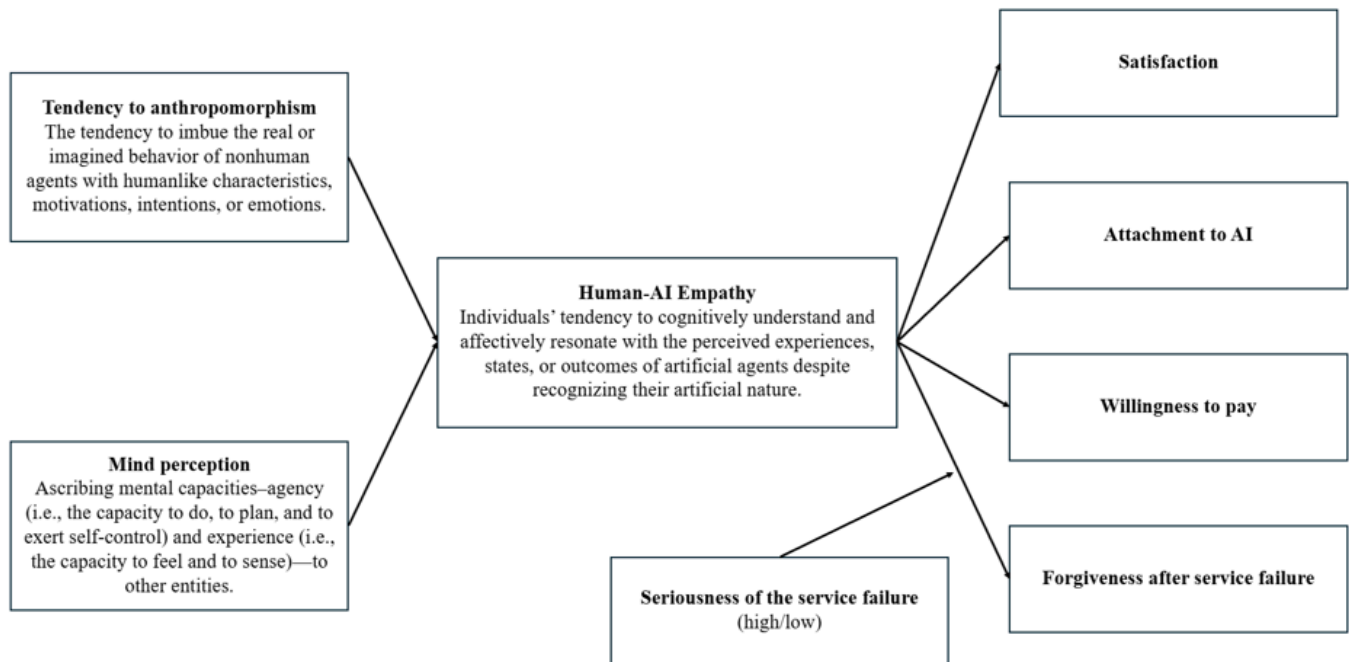


Figure 2

Conceptual model

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [APPENDICES.docx](#)