The Feeling Skills Gap: The Role of Empathy in Voice-Driven AI for Service Recovery

ABSTRACT

As companies increasingly deploy voice-driven artificial intelligence (AI) to enhance efficiency in service interactions, its potential negative impact on perceived customer orientation remains underexplored. Drawing on the AI constraints framework and the feeling economy perspective, this research examines how and when voice-driven AI affects perceptions of customer orientation in service recovery. Three experimental studies demonstrate that customers perceive service providers as less customer-oriented when recovery is handled by a voice-driven AI agent rather than a human, because of AI's inherent tendency to simplify complex emotional experiences into quantifiable parameters (i.e., parametric reductionism), failing to deliver authentic empathy during service recovery. Mediation analyses confirm that perceived empathy explains the link between agent type and downstream outcomes, including satisfaction, repurchase intentions, and behavioral loyalty. We also identify a boundary condition: the negative effect of voice-driven AI on perceived customer orientation emerges only when there is a task–ability mismatch, where the recovery task requires emotional intelligence (i.e., feeling-oriented skills) that the voice-driven AI agent cannot convincingly deliver, but not when the task calls for analytical problem-solving (i.e., thinking-oriented skills), where AI's capabilities are well aligned with task demands. We further discuss theoretical and practical implications for the use of voice-driven AI in service recovery.

Keywords: Artificial intelligence; Customer orientation; Empathy; Feeling Economy; Service recovery; Parametric Reductionism; Voice-driven AI.

1. Introduction

Companies are increasingly turning to voice-driven artificial intelligence (AI) agents to handle customer service interactions, substantially changing how consumers engage with service providers (Han et al., 2023). From handling requests to complaints (Zaki & Al-Romeedy, 2024), voice AI agents, powered by natural language processing and speech synthesis technologies, can emulate human-like interactions with increasing sophistication (Li et al., 2023). While early versions of these agents were relatively basic and limited to scripted-based interactions, recent advancements have made them highly capable, enabling more natural and engaging conversations (Roy & Naidoo, 2021; Sheehan et al., 2020). In customer service, voice AI can surpass human employees in speed, consistency, and scalability, adapting to new tasks at a pace far exceeding that of human employees (Kecht et al., 2023). These capabilities have driven widespread adoption of conversational AI tools, with projections indicating a substantial growth in the market size of AI agents reaching USD 36 billion by 2030 (Valuates Reports, 2021).

In service recovery, one concern in these interactions is whether AI agents can convey customer orientation — that is, the perception that a service provider genuinely understands and prioritizes customer needs (Blocker et al., 2011; Brady & Cronin, 2001; Kennedy et al., 2003). In human-led service recovery, customer orientation affects customer satisfaction and trust, largely due to the ability of human agents to demonstrate empathy. Although voice-driven AI agents can now integrate emotional cues through human-like tones and inflections, it remains unclear whether these agents can effectively generate perceptions of empathy to maintain customer orientation or whether their limitations in emotional expression will instead reinforce the perceived divide between AI and human agents, particularly in service recovery.

To address this gap, this research investigates whether, how, and when voice-driven AI agents influence customer perceptions and behaviors in service recovery. Across three experimental studies, we demonstrate that AI-driven service recovery leads to lower customer orientation and satisfaction perceptions, which in turn affect behavioral loyalty and repurchase intentions. Drawing on the feeling economy (Huang et al., 2019) and AI constraints (Valenzuela et al., 2024) frameworks, argue that the negative effects of AI-driven service recovery happen due to a perceived lack of empathy and AI's failure to account for the unique characteristics of human interactions, a phenomenon known as parametric reductionism (Valenzuela et al., 2024). Thus, we examine whether the type of ability required for the recovery task (feeling *vs.* thinking) moderates these effects, serving as a boundary condition for voice-driven AI's effectiveness in service recovery. Our findings show that when the task requires feeling-based (*i.e.*, empathetic) skills, customers perceive human agents as more customer-oriented, whereas for thinking-based (*i.e*., technical) tasks, AI may be perceived as performing on par with human agents due to its competence in analytical problem-solving.

In doing so, this study makes three main contributions to the literature. First, we offer empirical validation for the parametric reductionism bias introduced by Valenzuela et al. (2024) in their framework about how AI constraints human experience. Valenzuela et al. (2024) introduced parametric reductionism as a theoretical lens, arguing that AI, by design, simplifies complex and relational aspects of human experience into discrete, quantifiable parameters. To the best of our knowledge, no empirical study has yet demonstrated how this bias manifests in consumer-facing interactions. We demonstrate that customers perceive service recovery attempts by voice-driven AI agents as emotionally hollow (e.g., less empathetic), even when such efforts include symbolic gestures such as apologies or refunds.

Second, we contribute to the literature by bridging the AI Constraints framework (e.g., Valenzuela et al., 2024) and discussions around AI-based service recovery, by offering parametric reductionism as an explanation on why service recovery efforts from AI agents often backfire. While prior research (e.g., Zhu et al., 2023; Liang et al., 2024) has noted that AI struggles to elicit favorable customer reactions during service recovery, the reason why has remained underdeveloped. We argue that this failure stems from customers' recognition of the inauthentic, formulaic, or scripted nature of AI's emotional displays. This bias prevents AI from forming genuine emotional connections with customers (e.g., perceived empathy), thereby limiting the quality of the human experience in service interactions, translating into a lack of perceived empathy. We show that perceived empathy, in turn, mediates the effects of agent type on service recovery outcomes (i.e., perceptions of a company being customer-oriented, satisfaction, and behavioral loyalty).

Third, drawing on the feeling economy framework (e.g., Huang et al., 2019; Vorobeva et al., 2022), we identify a boundary condition for the negative effects of voice-driven AI in service recovery. Specifically, we show that these effects emerge when there is a mismatch between the emotional demands of the recovery task and the capabilities demonstrated by the AI agent. When the task requires feeling-oriented skills and the AI agent lacks the capacity to convincingly deliver them, customer evaluations and downstream outcomes are negatively affected. However, when the task instead draws on thinking-oriented skills, such as analytical or procedural recovery, this mismatch is less pronounced, and customer responses remain comparable to human-led service recovery. That is, we show that service recovery led by voice-driven AI agents only harms customer evaluations and service outcomes when the task performed requires feeling skills. Hence, when voice-driven AI agents handle tasks that align with their thinking-oriented strengths, such as analytical or procedural recovery actions, this perceptual mismatch is absent, and customer evaluations remain unaffected. Our work also offers insights for practitioners in helping to determine when AI-driven agents can be integrated without diminishing the emotional quality of customer interactions.

In what follows, we discuss how the use of AI agents in service recovery constraints interactions between customers and companies, thereby harming perceptions of companies' customer orientation and service outcomes. We then discuss how the lack of empathy often perceived in voice-driven AI agents accounts for these effects, and we explain that the negative outcomes of the use of voice-driven AI agents in service recovery only happen when there is a mismatch between the task performed and the type of ability required to perform it.

2. Theoretical Background

2.1. AI Agents Curb Service Recovery: Impacts for Perceived Customer Orientation and Downstream Outcomes

Customer orientation refers to a business approach that prioritizes meeting the needs and expectations of customers (Blocker et al., 2011), aligning all aspects of a company's operations, products, and services with the preferences and requirements of its target audience. This philosophy has often been associated with positive organizational performance (Kennedy et al., 2003), generating significant value for companies (Blocker et al., 2011). It is strategically relevant for companies as it fosters a deeper understanding of customer needs, enabling businesses to tailor their offerings accordingly. By being customer-oriented, companies can build stronger relationships with their customers, ultimately leading to increased customer satisfaction (Brady & Cronin, 2001), loyalty (Blocker et al., 2011), and even heightened well-being (Alabed et al., 2023).

Despite the strategic relevance and positive outcomes associated with customer orientation, it is important to acknowledge that service failures can still occur. Service failures, defined as instances where service delivery falls short of the desired outcome, remain a challenge for businesses despite their customer-oriented approach (Smith et al., 1999). These failures can manifest in various ways, such as incorrect orders or inefficient service. Service recovery refers to the actions taken by a company to address service failures and prevent negative customer responses (Andreassen, 2000; Chih et al., 2025), such as dissatisfaction. Service recovery strategies can be categorized into symbolic recovery (*e.g.*, non-monetary actions such as apologies and explanations) and utilitarian recovery (*e.g.*, tangible compensation like refunds or discounts) (Zaki & Al-Romeedy, 2024; Zhu et al., 2023). The ultimate goal is to transform negative customer emotions (*e.g.*, dissatisfaction) into positive ones (*e.g.*, satisfaction). Effective service recovery is crucial as it strengthens customer relationships, fosters trust and commitment, encourages repeat patronage, and generates positive word-of-mouth (DeWitt et al., 2008). Conversely, poor recovery attempts can intensify the negative effects of failures and lead customers to abandon the company for competitors.

While traditional service recovery has relied on human interventions, the emergence of AI in service settings introduces a new layer of complexity to these challenges.. One of the most notable advancements is the emergence of AI-based agents, such as chatbots or voicebots, that engage with customers through text or voice (Crolic et al., 2022; Heo & Lee, 2018; Hsu & Lin, 2023) in increasingly human-like interactions (Chen et al., 2024; Schuetz & Venkatesh, 2020). These agents are progressively replacing human professionals in frontline service roles (Crolic et al., 2022; Moorman et al., 2024; Sands et al., 2021), offering companies significant operational advantages. They deliver consistent performance by avoiding human limitations such as fatigue, frustration, or sickness, while outpacing the learning curve of human agents (Kecht et al., 2023). Beyond efficiency, AI agents provide 24/7 availability, handle multiple customer interactions simultaneously, maintain perfect memory of customer histories, and reduce operational costs through scalable deployment.

Despite AI-based agents' numerous advantages to companies, they still face significant limitations in service interactions: they excel at tasks that require data processing and logical reasoning, but still struggle to match humans in contexts that require nuanced emotional understanding (Huang & Rust, 2024; Zhang et al., 2023), such as service recovery efforts.

One attempt to address this limitation and bring AI closer to human-like emotional understanding has been the development of voice-driven AI agents, which aim to simulate human interactions more closely through speech, tone, and prosody. Unlike traditional text-based chatbots, voice-driven agents can convey emotional nuances using vocal tone (Van Der Goot et al., 2024) and inflection (Chih et al., 2025), potentially enabling a more natural and expressive interaction. However, emerging research and empirical evidence suggest that voice-driven AI continues to face limitations in emotionally charged contexts. While voice cues such as tone and inflection enhance the perception of empathy, they do not fully compensate for the absence of genuine emotional understanding. For instance, Zhu et al. (2023) found that, during service recovery, voice-based AI agents produced lower satisfaction and revisit intentions compared to human agents, highlighting that the medium alone does not close the empathy gap. Similarly, Zaki and Al-Romeedy (2024) observed that although service recovery efforts by voice AI agents can prompt customer forgiveness, they do not always enhance perceived empathy from the agent itself. These findings indicate that while voice can mimic empathy, it does not ensure that customers perceive the AI agent as genuinely empathetic.

A main reason voice-driven AI agents struggle in emotionally sensitive service recovery scenarios is parametric reductionism, a process where AI simplifies complex, dynamic aspects of human experience into fixed, computational, and scripted parameters (Valenzuela et al., 2024). In doing so, AI agents reduce relational depth into programmed routines, translating nuanced customer preferences, emotions, and behaviors into standardized, quantifiable variables. This simplification may help AI scale operations efficiently, but it also creates a perceptual disconnect: customers often recognize the formulaic nature of these interactions, leading to impressions of insincerity or emotional hollowness (e.g., Grewal et al., 2017). As a result, parametric reductionism undermines the emotional resolution customers seek during service recovery and constrains the authenticity of the human experience in interactions more broadly.

Therefore, while voice-enabled AI agents are often designed to humanize interactions through vocal features such as tone and inflection, these surface-level enhancements often fail to deliver effective perceptions of customer orientation. Customers may appreciate the conversational fluidity but still perceive these responses as algorithmic rather than genuinely felt, leading to what Mende et al. (2019) describe as a “deception cost”—the more an AI strives to be human-like, the more its emotional shortcomings stand out. For instance, in contexts requiring real emotional engagement—like service recovery—voice-driven AI that feigns empathy can highlight its own limitations, reducing perceived genuineness and trust.

Thus, simply equipping AI with a human-sounding voice does not necessarily enhance customer orientation and may even backfire. While AI agents have been designed to mimic human-like interactions (e.g., through voice and other anthropomorphic cues), recent findings suggest that such cues may actually exacerbate customer disappointment when the AI agent fails to meet heightened emotional expectations. For instance, Mende et al. (2019) show that consumers experience psychological discomfort and compensatory responses when humanoid agents trigger social expectations they cannot fulfill.

AI-based service recovery is particularly limited in emotionally charged situations, such as service recovery or contexts requiring improvisation (Bag et al., 2021), which leads customers to prefer human agents over AI ones, even when AI demonstrates superior task efficiency (Dietvorst et al., 2015). Although customers may show some tolerance for AI service failures when agents incorporate humor (Xu & Liu, 2022), and may attribute less responsibility to robots due to their perceived lack of task control (Leo & Huh, 2020), limitations persist in service recovery. Human agents excel in such situations due to their strong interpersonal communication skills (Luo et al., 2019) and their ability to adapt using emotional intelligence, thereby enhancing the customer experience (Kim et al., 2021; Puntoni et al., 2021). They can provide high-quality and personalized interactions through genuine emotional expressions, like eye contact and sincere smiles (Ariffin, 2013). Their friendliness and sincerity foster positive emotions in customers, facilitating forgiveness (Kemper & Lazarus, 1992). In addition, humans' ability to handle emotional tasks, establish relationships, and provide authentic interactions that feel familiar to customers (*i.e.,* in-group effect) (Park & Rothbart, 1982) creates an emotional authenticity that strongly influences how customers evaluate a service provider's commitment to understanding and serving their needs – in other words, their perceived customer orientation. When service recovery efforts are handled by AI agents, who are seen as lacking emotional intelligence and connection, customers may perceive the company as prioritizing efficiency over genuine customer care.

Equally, this effect can manifest in downstream consequences related to the service recovery experience. Beyond immediate resolution, service recovery also affects broader customer evaluations and future behavior towards companies. When recovery efforts are perceived as emotionally inauthentic, as often occurs with AI agents, customers may view the company as less trustworthy, which can undermine satisfaction, loyalty, and repurchase intentions (de Matos et al., 2007; Maxham & Netemeyer, 2002). This has already been consistently demonstrated also in non-AI service recovery: satisfaction reflects the customer's affective response to how well the service provider addressed their complaint (Gustafsson, 2009); loyalty and repurchase intentions, in turn, are influenced by whether customers feel that their needs were genuinely understood and prioritized (DeWitt, Nguyen &

Marshall, 2008, Grewal, Roggeveen & Tsiros, 2008). When voice-driven AI agents fall short in delivering affective reassurance, customers may attribute the failure to the company as a whole, leading to weaker relational bonds and increased likelihood of churn (Mattila, 2001; Tax et al., 1998). Thus, the inability of AI agents to handle service recovery effectively may also affect downstream consequences for service providers, namely a) satisfaction (Gustafsson, 2009), b) behavioral loyalty (DeWitt et al., 2008), and c) repurchase intentions (Grewal et al., 2008). Hence, we formally propose:

> *H1*. Service recovery efforts by voice-driven AI (*vs.* human) agents lead to lower perceived customer orientation and less positive service outcomes.

2.2. The Mediating Role of Perceived Empathy

Empathy has become increasingly relevant in understanding how customers evaluate service recovery attempts. Research shows that when AI agents use emotional language, adopt a warm tone, or simulate concern, customers often report higher satisfaction and show greater willingness to forgive service failures (Yun & Park, 2022; Fan et al., 2024; Zhang et al., 2024). These effects occur even in the absence of tangible compensation, suggesting that symbolic gestures rooted in empathy can be influential (Song et al., 2023; Zhou & Chang, 2024). The feeling economy perspective helps explain this shift. As companies automate routine and analytical tasks, social and emotional skills remain more difficult to replicate through machines (Huang & Rust, 2024; Mari et al., 2024). Within this context, the human capacity to recognize and respond to emotions gains renewed importance for service delivery (Bagozzi et al., 2022; Huang et al., 2019), and such interpersonal demands also influence hiring and compensation decisions, especially in roles that involve direct customer interaction (Gabbott et al., 2011; Bagozzi et al., 2022).

Despite this growing recognition of empathy's role in customer evaluations, companies continue to utilize voice-driven AI agents in service recovery, including emotionally sensitive interactions. This choice is often driven by efficiency and scalability, rather than alignment with the emotional demands of the task. However, we argue, service recovery represents a particularly delicate moment in the customer journey, one in which expectations around emotional support are heightened (DeWitt & Brady, 2003). When companies fail to meet these expectations, they risk triggering customer dissatisfaction, eroding trust, and undermining long-term engagement (Hess Jr. et al., 2003). While AI agents may be programmed to simulate concern, their ability to deliver emotionally resonant interactions remains limited, especially when compared to human agents. As a result, the use of voice-driven AI in such contexts may come into tension with what customers expect from a service provider in moments of failure.

In addition, the effectiveness of AI-delivered empathy is not universal. While AI agents can replicate certain empathy cues, these signals do not always translate into customer perceptions of genuine concern. A discrepancy then arises, because customers evaluate service agents not just by their words or gestures, but by how authentic and capable of understanding human emotions they appear. Prior research supports this distinction: Han et al. (2023) show that although positive emotion expressed by chatbots can trigger emotional contagion, it simultaneously leads to expectation-disconfirmation, especially among customers with transactional mindsets. Haupt et al. (2023) similarly find that empathy-seeking messages from chatbots enhance perceived warmth but have limited impact on perceptions of competence, suggesting that customers perceive AI empathy as surface-level rather than fully relational. Together, these findings highlight that simply signaling empathy does not ensure customers will perceive it as authentic or meaningful—particularly when the agent lacks the human-like qualities that traditionally convey such emotional understanding.

In addition, following a service failure, customers anticipate resolution and, importantly, expect companies to show genuine understanding through reassurance, apologies, and personalized engagement (Andreassen, 2000; Basso & Pizzutti, 2016). While human agents can adapt their communication style to match the customer's emotional state, providing a more personalized and empathetic experience, AI agents struggle to comprehend and respond appropriately to the emotional aspects of a customer's experience (Bergner et al., 2023). Moreover, while voice-driven AI tools offer efficiency and speed in addressing customer concerns, human agents' innate empathy and emotional intelligence (Huang & Rust, 2022) may make them more effective in service recovery. Human agents possess the ability to empathize with customers, actively listen to their concerns, and tailor their responses based on the emotional nuances of each situation. This level of emotional intelligence is essential for building trust and fairness (Andreassen, 2000; Smith et al., 1999), elements that remain difficult for AI systems to convincingly replicate (Huang et al., 2019).

------------- Insert Table 1 Around Here-----------

These observations point to an explanatory mechanism: beyond the objective presence of empathy cues, customers’ subjective perception of empathy shapes how they evaluate the service recovery agent. Across service recovery contexts, empathy functions as a relational signal that communicates the agent’s customer orientation—that is, their willingness and ability to prioritize customer needs and concerns. Consistent with this perspective, empathy has been shown to increase forgiveness, reduce dissatisfaction, and foster customer loyalty in post-failure interactions (Wieseke et al., 2012; Cartabuke et al., 2019). This applies to human agents and extends to AI-driven interactions: when customers perceive AI agents as empathetic, they are more likely to interpret the service provider as committed to delivering value and maintaining the relationship. Building on this logic, we posit that perceived empathy serves as the underlying mechanism linking agent type—whether human or voice-driven AI—to perceived customer orientation and, ultimately, to positive service outcomes. Formally:

> *H2*. Perceived empathy mediates the relationship between agent type (human *vs.* voice-driven AI), perceived customer orientation, and positive service outcomes.

2.3 The Task–Ability Mismatch as a Boundary Condition for Parametric Reductionism

As our earlier observations show, while AI agents have become increasingly sophisticated, important performance gaps persist between humans and AI in service recovery. Building on the Feeling Economy framework (Huang et al., 2019; Vorobeva et al., 2022), we distinguish between two broad categories of service recovery tasks: those that rely on analytical reasoning (thinking tasks) and those that involve emotional understanding (feeling tasks). While AI agents are well-suited for thinking tasks, such as providing technical explanations or executing procedural solutions, they remain less effective in contexts that call for spontaneous empathy, nuanced interpersonal cues, and relational sensitivity (Vorobeva et al., 2022). In these feeling-oriented interactions, human agents tend to be perceived as more attuned to customers' emotional needs and convey higher perceptions of empathy, which is especially important in service recovery.

The issue is not that AI is inherently unfit for all service recovery roles, but that mismatches occur when the agent's capabilities do not align with the emotional demands of the task. When voice-driven AI agents attempt to perform feeling tasks, customers may detect a lack of emotional authenticity, particularly when expressions of empathy feel scripted or standardized. This reflects what Valenzuela et al. (2024) describe as parametric reductionism, the process through which AI systems translate subjective, dynamic human experiences into simplified, predefined parameters. As a result, even if voice-driven AI agents simulate empathy through tone or language, customers often perceive a gap between what is expressed and what is genuinely felt. This misalignment, we propose, may not arise in thinking tasks, where standardized responses are more acceptable and emotional depth is not expected.

Empirical evidence from several service contexts supports this distinction. In domains such as hospitality (Khoa et al., 2023), healthcare (Bagozzi et al., 2022) and education (Belpaeme et al., 2018), studies show that human agents' proficiency in feeling tasks enhances trust, satisfaction, and loyalty. Customers prefer human agents when empathy and emotional connection are required, reinforcing the unique role of human emotional intelligence in service interactions (Bagozzi et al., 2022). On the other hand, AI agents have demonstrated advantages in thinking-oriented tasks involving data analysis, pattern recognition, and repetitive processing (Davenport et al., 2020; Kozinets & Gretzel, 2021), where emotional connection is secondary to efficiency and accuracy. Attempting to apply feeling skills through voice-driven AI agents in these contexts fails to meet customer expectations and, as we propose, may backfire: customers detect the mismatch between task demands and agent capabilities, perceiving AI-delivered empathy as inauthentic or hollow.

Therefore, we argue that the ability required to perform the task acts as a boundary condition in agent evaluations during service recovery. Specifically, while the relationship between agent type and perceived customer orientation is mediated by perceived empathy, this path is contingent on the emotional demands of the task. When the task is feeling-oriented, requiring emotional sensitivity and relational nuance, the limitations of AI in expressing authentic empathy become salient, diminishing perceived customer orientation and related downstream outcomes. In contrast, when the task is thinking-oriented, demanding analytical precision rather than emotional resonance, customers are less likely to devalue AI agents for their affective shortcomings. In such contexts, the mismatch between task demands and agent capabilities is less evident, and we expect no significant differences in customer evaluations across agent types. Formally, we propose:

H3. The mismatch (vs. match) between the ability required to perform the task (feeling vs. thinking) and the agent's demonstrated capabilities moderates the relationship between agent type and perceived empathy, such that the indirect effect of agent type on customer orientation via perceived empathy is stronger when there is a mismatch (i.e., feeling tasks handled by AI agents) than when there is a match (i.e., thinking tasks handled by AI agents or feeling tasks handled by humans).

-------------------------Please Add Figure 1 about here-------------------------

3. Overview of studies

Our theoretical framework makes several predictions that we empirically test. First, we propose that service recovery strategies – whether delivered by human or voice-driven AI agents – significantly impact perceived customer orientation and positive service outcomes (*i.e.*, satisfaction, behavioral loyalty, and repurchase intentions). Second, we posit that perceived empathy mediates this effect. Third, we propose that these effects are contingent on a boundary condition: the type of ability required to perform the service recovery task(feeling *vs.* thinking).

We test these predictions through three experimental studies. Studies 1 and 2 support our predicted causal relationship between agent type, perceived customer orientation, and positive service outcomes (H1). These studies also demonstrate the mediating role of perceived empathy (H2). Study 3 explores how the type of ability required to perform the task (thinking *vs.* feeling) moderates our theorized effects (H3).

The experimental designs and procedures were preregistered (Study 2: AsPredicted.Org *#207809*; Study 3: AsPredicted.Org *#120155*). To enhance generalizability, we conducted our studies across multiple participant pools and geographical regions, including participants from both the United States and Europe.

4. Study 1

Study 1 investigates the causal effect of service recovery agent type (human *vs.* voice-driven AI) on perceived customer orientation, satisfaction, and behavioral loyalty, while examining perceived empathy as the underlying mechanism. This study provides initial empirical evidence as to whether consumers perceive a service provider as less customer-oriented when AI agents, rather than humans, handle service recovery.

4.1. Method

Design and Participants. Study 1 uses a single-factor (agent type: voice-driven AI *vs.* human), between-subjects experiment. A total of one hundred and thirty Portuguese participants were recruited via social media on a voluntary basis (94.6% female, 5.4% male; $M_{\text{age}} = 42.03$; $SD_{\text{age}} = 10.35$).

Procedures and Stimuli. The experiment used vignette scenarios as stimuli, a method commonly employed in research examining new technologies (Jörling et al., 2019). Participants were randomly assigned to read a scenario describing a phone call interaction with either a human or a voice-driven AI customer service agent named Mario. The interaction occurred in the immediate recovery phase, where the service provider became aware of the failure and provided fair restitution to the customer (Miller et al., 2000). The scenario asked participants to imagine they had received an online order of a household appliance via home delivery and discovered deep scratches upon opening the package. In both conditions (human and voice-driven AI), Mario followed the same sequence of steps for service recovery. To enhance the realism of our experiment, these steps were designed following two elements: first, we incorporated the fundamental service recovery strategies established in the academic literature, combining both emotional recovery (offering an apology and providing an explanation) and tangible recovery (resolving the failure and offering monetary compensation through a gift card) (Miller et al., 2000; Wei et al., 2020). Second, to improve the external validity of the scenario, we based our dialogue on a real customer service interaction script obtained from the largest e-commerce retailer in Portugal, which allowed us to capture authentic service recovery language and resolution patterns. The scenario concluded by specifying that the call lasted five minutes and confirming that the failure was successfully addressed with a replacement item delivered within two hours. The survey was translated from English to Portuguese using a double translation (Pavone et al., 2023). The complete stimulus of the experiment is presented in Appendix A.

Measures. We assessed all variables using a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree). Our primary dependent variable, perceived customer orientation, was measured using a five-item scale adapted from Homburg et al. (2009) (α = .89). To capture service outcomes, we measured customer satisfaction using three items adapted from Wieseke et al. (2012) (α = .81) and behavioral loyalty using three items from Shuqair et al. (2019) (α = .78). The hypothesized mediating variable, perceived empathy, was assessed using a four-item scale adapted from Wilder et al. (2014) (α = .92). To validate our experimental design, we included two additional measures. First, participants completed a manipulation check to verify their recognition of the agent type (voice-based AI versus human). Second, they evaluated the scenario's realism to ensure external validity. We concluded the survey by collecting demographic information. All measures have reliabilities greater than the recommended threshold of 0.7 (Juquelier et al., 2025) (see table 3).

4.2. Results

Manipulation checks. Results show that the manipulation worked as intended. Results from an independent samples t-test revealed that participants in the human condition showed higher agreement with the statement that the agent of the call was a human (*i.e.*, "a first-line assistant") ($M_{\text{human}} = 5.31$, $SD = 1.97$; $M_{\text{AI}} = 3.76$, $SD = 2.50$; $t(122.840) = 3.948$, $p < 0.001$). Similarly, participants in the voice-based AI condition indicated a higher level of agreement with the statement that the agent of the call was a voice-based AI (*i.e.*, "an artificial intelligence assistant") ($M_{\text{human}} = 3.19$, $SD = 2.25$; $M_{\text{AI}} = 6.02$, $SD = 1.77$; $t(119.581) = -7.962$, $p < 0.001$). Additionally, the results indicate that our experiment offers high realism, significantly differing from the scale's midpoint ($M = 4.55$; $SD = 2.31$, $t\ (129) = 5.16$, $p < 0.001$) (Pavone et al., 2023).

Analyses. To investigate the effects of the type of agent on customer orientation and positive service outcomes (H1), we conducted multiple independent samples t-tests. The type of agent (X) served as the independent variable – coded as 0 = human *vs.* 1 = voice-driven AI; perceived customer orientation and positive service outcomes (*i.e.*, customer satisfaction and behavioral loyalty) as the dependent variables (Y). To explore the mediating role of perceived empathy on perceived customer orientation and customer positive service outcomes (H2), we used PROCESS Model 4 with 5,000 bootstrapped samples and a 95% confidence interval (Hayes, 2017). The type of agent served as the independent variable (X); perceived empathy as the mediator (M); perceived customer orientation and positive service outcomes as the dependent variables (Y).

4.2.1. Customer orientation

Main effect. As predicted, there was a significant main effect of agent type on perceived customer orientation. Participants interacting with voice-based AI agents (M = 5.09, SD = 1.67) reported lower perceived customer orientation compared to those interacting with human agents (M = 5.68, SD = 1.24; $t(119.871) = 2.268$, $p = 0.013$). These findings support H1, demonstrating that interactions with voice-based AI agents, compared to human agents, damage the perception of customer orientation (see figure 2).

Mediation. First, we assessed the effect of agent type on perceived empathy (the mediator). The overall model predicting perceived empathy was significant ($R^2 = 0.187$, $F(1,128) = 29.420$, $p < 0.001$). Agent type had a significant effect on perceived empathy (β = -1.63, SE = 0.30, $t(128) = -5.424$, $p < 0.001$), indicating that participants interacting with voice-based AI agents perceived lower empathy compared to those interacting with human agents.

Next, we examined the effect of agent type and perceived empathy on perceived customer orientation (the dependent variable). The overall model was significant ($R^2 = 0.344$, $F(2,127) = 33.228$, $p < 0.001$). Perceived empathy emerged as a significant predictor of perceived customer orientation (β = 0.49, SE = 0.06, $t(127) = 7.684$, $p < 0.001$). In contrast, the direct effect of agent type on perceived customer orientation was not significant when controlling for perceived empathy (β = 0.21, SE = 0.24, $t(127) = 0.868$, $p = 0.387$), suggesting that the effect of agent type on perceived customer orientation is fully mediated by perceived empathy. The analysis of the indirect effect of agent type on perceived customer orientation through perceived empathy revealed a significant indirect effect (*Effect* = -0.79, *BootSE* = 0.20, 95% *CI* [-1.21, -0.45]), confirming the mediating role of perceived empathy (H2).

4.2.3. Customer satisfaction

Main effect. As predicted, there was a significant main effect of agent type on customer satisfaction. Participants interacting with voice-based AI agents ($M = 5.87$, $SD = 1.34$) reported lower customer satisfaction compared to those interacting with human agents ($M = 6.39$, $SD = 0.90$; $t(113.577) = 2.612$, $p = 0.005$). These findings support H1, demonstrating that interactions with voice-based AI agents, compared to human agents, damage customer satisfaction.

Mediation. We examined the effect of agent type and perceived empathy on customer satisfaction (the dependent variable). The overall model was significant ($R^2 = 0.201$, $F(2,127) = 15.933$, $p < 0.001$). Perceived empathy emerged as a significant predictor of customer satisfaction ($\beta = 0.27$, $SE = 0.05$, $t(127) = 4.891$, $p < 0.001$). In contrast, the direct effect of agent type on customer satisfaction was not significant when controlling for perceived empathy ($\beta = -0.09$, $SE = 0.21$, $t(127) = -0.428$, $p = 0.670$), suggesting that the effect of agent type on customer satisfaction is fully mediated by perceived empathy. The analysis of the indirect effect of agent type on customer satisfaction through perceived empathy revealed a significant indirect effect (*Effect* = -0.43, *BootSE* = 0.12, 95% *CI* [-0.69, -0.22]), further supporting the mediating role of perceived empathy (H2).

4.2.3. Behavioral loyalty

Main effect. As predicted, there was a significant main effect of agent type on behavioral loyalty. Participants interacting with voice-based AI agents ($M = 5.42$, $SD = 1.40$) reported lower behavioral loyalty compared to those interacting with human agents ($M = 6.16$, $SD = 0.94$; $t(114.136) = 3.517$, $p < 0.001$). These findings support H1, demonstrating that interactions with voice-based AI agents, compared to human agents, damage behavioral loyalty.

Mediation effect. We examined the effect of agent type and perceived empathy on behavioral loyalty (the dependent variable). The overall model was significant ($R^2 = 0.196$, $F(2,127) = 15.524$, $p < 0.001$). As expected, perceived empathy emerged as a significant predictor of behavioral loyalty ($\beta = 0.24$, $SE = 0.06$, $t(127) = 4.156$, $p < 0.001$). The analysis of the indirect effect of agent type on behavioral loyalty through perceived empathy revealed a significant indirect effect (*Effect* = -0.39, *BootSE* = 0.13, 95% *CI* [-0.66, -0.16]). In contrast, the direct effect of agent type on behavioral loyalty was not significant when controlling for perceived empathy ($\beta = -0.34$, $SE = 0.22$, $t(127) = -1.550$, $p = 0.124$), suggesting that the effect of agent type on behavioral loyalty is fully mediated by perceived empathy (H2).

4.3. Discussion

Study 1 shows that voice-based AI agents reduce perceived empathy, which in turn lowers customer orientation, satisfaction, and behavioral loyalty. Customers interacting with AI agents saw them as less empathetic than human agents, consistent with prior work on AI's emotional limitations (Van Doorn et al., 2017). This lack of empathy how customer-oriented the agents seemed, and how satisfied and loyal participants felt. The results support the idea of parametric reductionism (Valenzuela et al., 2024): when AI simplifies complex human emotions into scripts, the interaction feels less authentic and less relational. Study 2 builds on these findings by offering a conceptual replication and testing whether these effects extend to other outcomes (i.e., repurchase intention and word-of-mouth).

-------------------------Please Add Figure 2 about here-------------------------

5. Study 2

Study 2 extends the previous study by examining whether the type of service recovery agent (human *vs.* voice-driven AI) influences perceived customer orientation and perceived empathy, while also investigating repurchase intentions and positive word-of-mouth as additional behavioral outcomes. Furthermore, one could argue that our results might differ if participants interacted with a voice-based AI assistant, as vocal cues could enhance perceptions of humanness and emotional connection. To address this concern, we enhanced the strength of our manipulation by including a voice stimulus simulating a phone call interaction with the agent to ensure a more immersive and ecologically valid experimental setup.

5.1. Method

Design and Participants. Study 2 uses a single-factor (agent type: voice-driven AI *vs.* human) scenario-based, between-subjects experiment. A total of one hundred and thirty-five U.S. participants were recruited via CloudResearch in exchange for monetary compensation. We excluded participants who failed the attention check (n = 3), leaving the final sample including 132 participants (49.2% male, 47.7% female, 3% other; M_{age} = 37.69; SD_{age} = 10.39).

Procedures and Stimuli. Participants were randomly assigned to the same scenario as Study 1, describing a phone call interaction with either a human or voice-driven AI customer service agent named Mario. Only minor adjustments were made to improve clarity in the English version (*e.g.*, changing "a first-line assistant" to "a human agent").

Additionally, to ensure the realism of our experiment and that the compensation amount was realistic for the U.S. market, we examined online forums (*e.g.*, Reddit) where customers discussed gift card offerings in service recovery situations. Since compensation values were comparable, we maintained the same amount but converted it from euros to dollars (€15 to $15). Furthermore, following the review team's feedback, we strengthened the voice-based AI manipulation by incorporating an audio recording of Mario's scripted dialogue from the scenario. This audio stimulus was presented to participants in both conditions after the text description. To verify participant engagement with the audio element, we implemented a timing feature in Qualtrics to track whether participants listened to the clip. The complete stimulus of the experiment is presented in Appendix A.

Measures. We assessed all variables using a nine-point Likert scale (1 = strongly disagree, 9 = strongly agree). Our primary dependent variable, perceived customer orientation, was measured using the same five-item scale as Study 1, adapted from Homburg et al. (2009) ($\alpha = .87$). To capture service outcomes, we measured customer satisfaction using the same three-item scale as Study 1, adapted from Wieseke et al. (2012) ($\alpha = .95$), repurchase intention using a four-item scale adapted from Ho & Chung (2020) ($\alpha = .89$), and positive word-of-mouth using a three-item scale adapted from Kirk et al. (2023) ($\alpha = .96$). The hypothesized mediating variable, perceived empathy, was assessed using the same four-item scale as Study 1, adapted from Wilder et al. (2014) ($\alpha = .96$). Similar to study 1, to validate our experimental design, we included two manipulation check questions to verify the participants' recognition of the agent type (voice-based AI *vs.* human) and a question regarding the scenario's realism to ensure external validity. Additionally, we included an attention check question instructing participants to indicate a specific response ("To confirm your attention, please select option 3 for this question"; adapted from Pavone et al., (2023). Demographic characteristics were also included. All measures had reliabilities greater than the recommended 0.7 threshold (Juquelier et al., 2025).

5.2. Results

Manipulation checks. Results showed that the manipulation worked as intended. An independent samples t-test revealed that participants in the human condition showed significantly higher agreement with the statement that the agent of the call was a human (M_{human} = 8.27, SD = 1.75; M_{AI} = 1.29, SD = 1.17; $t(115.316) = -26.959$, $p < 0.001$). Similarly, participants in the voice-based AI condition reported significantly higher agreement with the statement that the agent of the call was a voice-based AI (M_{human} = 2.22, SD = 2.47; M_{AI} = 8.86, SD = 0.46; $t(70.772) = 21.582$, $p < 0.001$). The results also indicate that our experiment offered an ecologically valid service encounter, being rated as highly realistic, significantly above the scale's midpoint (M = 7.31; SD = 1.66, $t(131) = 19.42$, $p < .001$) (Pavone et al., 2023).

Analyses. To investigate the effects of the type of agent on customer orientation and positive service outcomes (H1), we conducted multiple independent samples t-tests. Similar to study 1, the type of agent (X) served as the independent variable – coded as 0 = human *vs.* 1 = voice-driven AI; perceived customer orientation and positive service outcomes (*i.e.*, customer satisfaction; repurchase intention, and positive word-of-mouth) as the dependent variables (Y). To explore the mediating role of perceived empathy on perceived customer orientation and customer positive service outcomes (H2), we used PROCESS, model 4, with 5,000 bootstrapped samples, with a 95% confidence interval (Hayes, 2017). The type of agent served as the independent variable (X); perceived empathy as the mediator (M); perceived customer orientation and positive service outcomes as the dependent variables (Y).

5.2.1. Customer orientation

Main effect. As predicted, there was a significant main effect of agent type on perceived customer orientation. Participants interacting with voice-based AI agents (M =

7.08, *SD* = 1.56) reported lower perceived customer orientation compared to those interacting with human agents (*M* = 7.73, *SD* = 1.11; $t(130) = -2.78$, $p = 0.006$). These findings offer additional support for H1, reinforcing that interactions with voice-based AI agents, compared to human agents, damage the perception of customer orientation (see figure 3).

Mediation. First, we assessed the effect of agent type on perceived empathy (the mediator). The overall model predicting perceived empathy was significant ($R^2 = 0.154$, $F(1,130) = 23.583$, $p < 0.001$). Agent type had a significant effect on perceived empathy (β = -1.70, *SE* = 0.35, $t(130) = -4.856$, $p < 0.001$), indicating that participants interacting with voice-based AI agents perceived lower empathy compared to those interacting with human agents.

Next, we examined the effect of agent type and perceived empathy on perceived customer orientation (the dependent variable). The overall model was significant ($R^2 = 0.535$, $F(2,129) = 74.185$, $p < 0.001$). Perceived empathy emerged as a significant predictor of perceived customer orientation (β = 0.48, *SE* = 0.04, $t(129) = 11.523$, $p < 0.001$). In contrast, the direct effect of agent type on perceived customer orientation was not significant when controlling for perceived empathy (β = 0.16, *SE* = 0.18, $t(129) = 0.882$, $p = 0.379$), suggesting that the effect of agent type on perceived customer orientation is fully mediated by perceived empathy. The analysis of the indirect effect of agent type on perceived customer orientation through perceived empathy revealed a significant indirect effect (*Effect* = -0.81, *BootSE* = 0.20, 95% *CI* [-1.24, -0.46]), thus supporting H2.

5.2.2. Customer satisfaction

Main effect. As predicted, there was a significant main effect of agent type on customer satisfaction. Participants interacting with voice-based AI agents (*M* = 7.76, *SD* = 1.49) reported lower satisfaction compared to those interacting with human agents (*M* = 8.30, *SD* = 1.20; *t*(122.68) = -2.27, *p* = 0.025). These findings offer additional support for H1, reinforcing that interactions with voice-based AI agents, compared to human agents, damage customer satisfaction.

Mediation. We examined the effect of agent type and perceived empathy on customer satisfaction (the dependent variable). The overall model was significant ($R^2 = 0.357$, $F(2,129) = 35.765$, $p < .001$). Perceived empathy emerged as a significant predictor of customer satisfaction ($\beta = 0.39$, $SE = 0.05$, $t(129) = 7.991$, $p < .001$). In contrast, the direct effect of agent type on customer satisfaction was not significant when controlling for perceived empathy ($\beta = 0.12$, $SE = 0.21$, $t(129) = 0.582$, $p = 0.562$), suggesting that the effect of agent type on customer satisfaction is fully mediated by perceived empathy. The analysis of the indirect effect of agent type on customer satisfaction through perceived empathy revealed a significant indirect effect (*Effect* = -0.66, *BootSE* = 0.18, 95% *CI* [-1.05, -0.35]), also supporting our proposed mediation.

5.2.3. Repurchase intentions

Main effect. As predicted, there was a significant main effect of agent type on repurchase intentions. Participants interacting with voice-based AI agents ($M = 6.82$, $SD = 1.55$) reported lower repurchase intentions compared to those interacting with human agents ($M = 7.31$, $SD = 1.60$; $t(130) = -1.77$, $p = 0.039$). These findings offer additional support for H1, reinforcing that interactions with voice-based AI agents, compared to human agents, reduce customers' intentions to repurchase from the company.

Mediation. We examined the effect of agent type and perceived empathy on repurchase intentions (the dependent variable). The overall model was significant ($R^2 = 0.372$, $F(2,129) = 38.133$, $p < 0.001$). Perceived empathy emerged as a significant predictor of repurchase intentions ($\beta = 0.47$, $SE = 0.06$, $t(129) = 8.45$, $p < 0.001$). In contrast, the direct effect of agent type on repurchase intentions was not significant when controlling for perceived empathy ($\beta = 0.31$, $SE = 0.24$, $t(129) = 1.285$, $p = 0.201$), suggesting that the effect of agent type on repurchase intentions is fully mediated by perceived empathy. The analysis of the indirect effect of agent type on repurchase intentions through perceived empathy revealed a significant indirect effect (*Effect* = -0.80, *BootSE* = 0.18, 95% CI [-1.18, -0.45]), further confirming H2.

5.2.4. Positive word-of-mouth

Main effect. Results show no significant main effect of agent type on positive word-of-mouth. Participants interacting with voice-based AI agents ($M = 7.35$, $SD = 1.62$) did not significantly differ from those interacting with human agents ($M = 7.66$, $SD = 1.67$; $t(130) = -1.06$, $p = 0.292$). These findings suggest that while interactions with voice-based AI agents negatively impact customer orientation and satisfaction, they do not significantly reduce customers' willingness to recommend the company to others.

5.3. Discussion

The results of Study 2 extend Study 1's findings by showing that voice-driven AI agents, compared to human agents, lead to lower repurchase intentions, with empathy continuing as a mediator of these effects. Interestingly, while these negative effects extended to perceived customer orientation, satisfaction, and behavioral loyalty, positive word-of-mouth intentions were not significantly affected. These findings suggest that although voice-driven AI agents diminish multiple customer perceptions and behavioral outcomes, this negative impact does not significantly influence customers' willingness to recommend the company to others.

To extend these findings, Study 3 tests whether the negative effects of voice-driven AI agents on perceived empathy and customer orientation happen because of a mismatch between the skill required by the recovery task (feeling vs. thinking) and the capabilities demonstrated by the agent.

-------------------------Please Add Figure 3 about here-------------------------

6. Study 3

Study 3 examines whether the effect of agent type (human vs. voice-driven AI) on perceived empathy and customer orientation depends on the fit between the skill required by the service recovery task and the agent's ability to perform it. In this study, we test whether the negative effects observed in prior studies emerge specifically when a voice-driven AI agent is assigned to a feeling-oriented task, for which it lacks the emotional fluency required. In contrast, when the same agent performs a thinking-oriented task, more aligned with its analytical strengths, we expect this mismatch to be less pronounced and customer evaluations to remain unaffected (H3).

6.1. Method

Design and Participants. Study 3 was a 2 (agent type: voice-driven AI *vs.* human) x 2 (type of task: feeling *vs.* thinking) scenario-based, between-subjects experiment. A total of three hundred and twenty-one Portuguese participants were recruited via Prolific in exchange for monetary compensation. We excluded participants who failed two or more attention checks (n = 10) and exhibited unusually rapid survey completion times (n = 3), leaving the final sample of 308 participants (56.5% male, 42.9% female, 0.6% other; M_{age} = 29.04; SD_{age} = 8.84).

Procedures and Stimuli. Participants were randomly assigned to a scenario similar to the previous studies, describing a phone call interaction with either a human or voice-driven

AI customer service agent named Mario. The scenario involved receiving a damaged tablet (with deep scratches) from an online order delivered to their home. Drawing from the literature, we designed descriptive manipulations that reflected the distinct skills required to perform feeling and thinking tasks (Rust et al., 2021). The survey was translated from English to Portuguese using a double translation (Pavone et al., 2023). The complete stimuli of the experiment are presented in Appendix A.

Measures. We assessed all variables using a nine-point Likert scale (1 = strongly disagree, 7 = strongly agree). Our primary dependent variable, perceived customer orientation, was measured using the same five-item scale as the previous studies, adapted from Homburg et al. (2009) ($\alpha = .87$). The hypothesized mediating variable, perceived empathy, was assessed using the same four-item scale as the previous studies, adapted from Wilder et al. (2014) ($\alpha = .84$). To validate our experimental design, we included four manipulation check questions to verify participants' recognition of the agent type (a voice-driven AI or a human) and the agent's skill type (social or technical skills). These questions also served as attention checks, leading to the exclusion of participants who failed to answer two or more questions correctly by selecting extreme responses (1 = strongly disagree or 9 = strongly agree). We also added the same question used in the previous studies to assess the scenario's realism. Finally, demographic characteristics were included. All measures had reliabilities greater than the recommended 0.7 threshold (Juquelier et al., 2025).

6.2. Results

Manipulation checks. Results showed that the manipulation worked as intended. A one-way ANOVA revealed that participants in the human condition showed significantly higher agreement with the statement that the agent of the call was a human ($M_{\text{human}} = 6.51$, $SD = 1.38$; $M_{\text{AI}} = 1.42$, $SD = 1.32$; $F(1, 306) = 1094.305$, $p < 0.001$). Similarly, participants in the voice-based AI condition reported significantly higher agreement with the statement that the agent of the call was a voice-based AI (M_{human} = 1.56, SD = 1.44; M_{AI} = 6.73, SD = 1.02; $F(1, 306) = 2057.195$, $p < 0.001$). In addition, participants in the feeling task condition showed significantly higher agreement with the statement that the agent of the call had social skills (M_{feeling} = 5.99, SD = 1.36; M_{thinking} = 4.27, SD = 2.03; $F(1, 306) = 74.712$, $p < 0.001$). Similarly, participants in the thinking task condition reported significantly higher agreement with the statement that the agent of the call had technical skills (M_{feeling} = 4.54, SD = 1.64; M_{thinking} = 6.38, SD = 1.05; $F(1, 306) = 139.747$, $p < 0.001$). Finally, the results indicated that our experiment offered an ecologically valid service encounter, being rated as significantly higher than the scale midpoint (M = 4.91; SD = 1.55, $t\,(307) = 15.91$, $p < .001$) (Pavone et al., 2023).

Analyses. To investigate the moderation effects of the type of task on customer orientation (H3), we conducted a two-way ANOVA. The perceived customer orientation (Y) served as the dependent variable, and the agent and the type of skill as factors. To examine whether the type of skill shown by the agent (thinking *vs.* feeling) moderates the indirect effect of agent type on perceived customer orientation via perceived empathy, we conducted a moderated mediation analysis using the PROCESS macro (Model 7, Hayes, 2017) with 5000 bootstrap samples and a 95% confidence interval (CI). The agent type (X) (coded as 0 = human, 1 = voice-driven AI) was used as the independent variable, perceived empathy (M) as the mediator, perceived customer orientation (DV) as the dependent variable, and task type (W) (coded as: 0 = thinking, 1 = feeling) as the moderator.

Main effect. Results of the two-way ANOVA revealed a significant main effect of the agent (M_{human} = 6.25, SD = 0.81 *vs.* M_{AI} = 6.05, SD = 0.90, $F(1, 304) = 3.848$, $p = 0.051$). The main effect of the type of skill was not significant (M_{feeling} = 6.17, SD = 0.87 *vs.* M_{thinking} = 6.13, SD = 0.85, $F(1, 304) = 0.190$, $p = 0.663$). Most importantly, there was a significant two-way interaction between agent type and skill type on customer orientation (F (1, 304) = 8.414, p = 0.004).

Planned contrasts revealed that in feeling-oriented tasks, human agents were rated as significantly more customer-oriented (M_{human} = 6.41, SD = 0.70 *vs.* M_{AI} = 5.94, SD = 0.96; F (1, 304) = 12.385, $p < 0.001$). However, in thinking-oriented tasks, there was no significant difference between human and AI agents (M_{human} = 6.08, SD = 0.89 *vs.* M_{AI} = 6.18, SD = 0.82; F (1, 304) = 0.422, p = 0.517) (see figure 4).

Results from a two-way ANOVA with perceived empathy as the dependent variable, and the agent and the type of skill as factors show a significant main effect of the agent (M_{human} = 5.87, SD = 1.19 *vs.* M_{AI} = 4.96, SD = 1.52, F(1, 304) = 34.038, $p < 0.001$). The main effect of the type of skill was not significant (M_{feeling} = 5.31, SD = 1.61 *vs.* M_{thinking} = 5.54, SD = 1.22, F(1, 304) = 2.347, p = 0.127). Most importantly, there was a significant two-way interaction between agent and type of skill on perceived empathy (F (1, 304) = 8.342, p = 0.004).

Planned contrasts revealed that perceived empathy in the feeling task was higher in the human agent condition (M = 5.97, SD = 0.16), as compared to the AI agent condition (M = 4.63, SD = 0.15, $p < 0.001$). In the thinking task, participants also perceived the human agent as more empathetic (M = 5.77, SD = 0.15) than the AI agent (M = 5.31, SD = 0.16, p = 0.042).

Moderated mediation. Results show that the interaction between agent type and task type was significant (β = -0.888, $BootSE$ = 0.308, p = 0.004, 95% CI [-1.493, -0.283]). In thinking tasks the effect of agent type on perceived empathy was significant (β = -0.453, $BootSE$ = 0.222, p = 0.043, 95% CI [-0.890, -0.015]). In feeling tasks the effect was also significant, with voice-driven AI perceived as less empathetic compared to human agents (β = -1.341, $BootSE$ = 0.212, $p < 0.001$, 95% CI [-1.759, -0.923]). When examining the indirect effect of agent type on perceived customer orientation through perceived empathy, we found a significant indirect effect both on thinking (β = -0.175, *BootSE* = 0.075, 95% *CI* [-0.321, -0.027]) and feeling tasks (β = -0.517, *BootSE* = 0.097, 95% *CI* [-0.713, -0.330]). As in the prior studies, the direct effect of agent type on perceived customer orientation was not significant (β = 0.150, *BootSE* = 0.081, p = .066, 95% *CI* [-0.010, 0.310]), which suggests a full mediation.

-------------------------Please Add Figure 4 about here-------------------------

6.3. Discussion

Study 3 provides evidence for our theorized moderation, confirming that the negative effects of voice-driven AI agents on perceived customer orientation depend on a mismatch between the type of skill required and the capabilities of the agent. When the recovery task demanded feeling-based skills, AI agents—constrained by scripted expressions of empathy by definition—were perceived as less empathetic than human agents, reducing perceptions of customer orientation. This aligns with the notion of parametric reductionism (Valenzuela et al., 2024), where AI simplifies complex emotional experiences into predefined parameters. However, when the task involved thinking-based skills, where emotional nuance was less critical, this mismatch was not present, and customer evaluations did not differ significantly across agents. These findings confirm that it is not AI's use per se, but rather the fit between the task's emotional demands and the agent's capabilities, that drives customer responses in service recovery contexts.

7. General discussion

As companies integrate voice-driven AI into service recovery, the question is no longer whether AI can replace human agents but rather how its presence affects consumer perceptions. While AI is often advocated for its efficiency, our findings reveal a critical trade-off: this efficiency comes at the expense of perceived customer orientation, especially in service recovery tasks requiring feeling skills, rather than thinking skills.

Across three experimental studies, we demonstrate that customers perceive service providers as less customer-oriented when service recovery is handled by voice-driven AI rather than human agents, ultimately lowering satisfaction, behavioral loyalty, and repurchase intentions. We further show that this effect is mediated by perceived empathy, as consumers perceive that voice-driven AI struggles to replicate the emotional cues that make human service interactions feel genuine and relational.

Finally, our findings also reveal the conditions under which voice-driven AI agents' limitations are more pronounced. We show that such limitations depend on the type of skill required to perform the service recovery task: when the task demands emotional sensitivity, AI's limitations are more evident. But when the task is analytical or procedural, where emotional expression is less important, customer responses to AI and human agents are comparable. Hence, AI constraints face greater resistance when customers expect feeling skills and human-like emotional engagement.

7.1. Theoretical contributions

This research advances the growing body of literature on voice-driven AI in service encounters in three ways. First, this research provides empirical support for the framework recently proposed by Valenzuela et al. (2024) to explain how AI constrains the human experience in service contexts. According to this framework, AI limits human experience along three dimensions: (1) agency transference, where consumers delegate intentionality and decision-making to AI systems; (2) regulated expression, where users adjust or restrain their emotional responses to better align with the limited interpretative range of AI; and (3) parametric reductionism, whereby AI transforms rich, relational, and context-sensitive human experiences into simplified, pre-structured variables for processing. While some empirical research has begun to examine agency transference in service contexts (e.g., Wagner et al., 2025), the current work is, to our knowledge, the first to empirically test the experiential consequences of parametric reductionism in consumer-facing interactions. Specifically, we show that even when AI agents use empathy scripts, offer apologies, or provide restitution, consumers interpret these actions as emotionally hollow, potentially because they follow pre-programmed rules rather than emerging from genuine human understanding. This suggests that customers are sensitive to what is said by AI, as well as to how and why it is said.

Second, this research contributes to the service recovery literature by introducing parametric reductionism as a theoretical lens to explain why AI-led recovery efforts often fail. While prior studies have documented that AI agents underperform in emotionally charged recovery contexts (e.g., Zhu et al., 2023; Liang et al., 2024), we build on Valenzuela et al.’s (2024) AI Constraints framework to theorize and demonstrate that customers reject the scripted, mechanical quality of AI’s expressions, which are perceived as less empathetic responses rather than context-sensitive emotional attunement. This aligns with research showing that consumers scrutinize the authenticity of service agents' emotional displays (e.g., Grandey et al., 2005). We empirically demonstrate that this mechanistic perception undermines perceived empathy, which in turn explains the negative effect of AI on downstream outcomes such as customer orientation, satisfaction, and behavioral loyalty. By doing so, our work highlights an important gap in the literature: it is not only the presence or absence of recovery behaviors that matters, but the perceived authenticity of the interaction delivering them. Thus, our findings extend empathy research in service recovery (e.g., Mattila & Enz, 2002; DeWitt et al., 2008) by showing that technological mediation—when poorly matched to relational demands—can erode the interpersonal meaning of recovery gestures.

Third, we contribute to the emerging dialogue between the feeling economy framework (Huang et al., 2019; Vorobeva et al., 2022) and AI-based service recovery by identifying a boundary condition that helps explain when and why voice-driven AI harms customer outcomes. We show that the mismatch between the skills required by the recovery task and the expressive capabilities of the AI agent is a determinant of customer evaluations. This mismatch becomes especially salient in feeling-oriented tasks, those that demand emotional nuance and empathy. In these situations, the scripted, reductionist delivery of AI agents fails to meet the emotional expectations of customers, reducing evaluations of customer orientation, satisfaction, loyalty, and repurchase intention. However, when the task instead calls for thinking-oriented skills, such as executing an analytical fix or issuing a procedural solution, AI's limitations in affective expression are less likely to be noticed or penalized by customers.

This is not because AI becomes more emotionally competent, but because emotional expressiveness is less relevant in such contexts. Our findings thus support and refine prior work on human-AI fit in service delivery (e.g., Mende et al., 2019; Belanche et al., 2020) by showing that task-agent alignment is crucial: AI can be effective, but only when assigned to tasks that match its capabilities. In doing so, we also respond to recent calls (e.g., Marinova et al., 2017; Puntoni et al., 2021) to move beyond general assessments of AI performance and instead identify the specific conditions under which AI can support or undermine service relationships. Our research advances this agenda, empirically showing that voice-driven AI is not inherently detrimental to service recovery; rather, its impact depends on how well the nature of the task matches the type of intelligence the AI can convincingly project, analytical precision versus emotional resonance.

7.2. Managerial implications

Our findings also offer guidance for companies and practitioners regarding when and how voice-driven AI should be integrated into service recovery strategies. Despite the proven benefits of voice-driven AI agents in maximizing service efficiency, our findings reveal that they fail to meet consumers' expectations for empathy, thereby undermining customer orientation and positive service outcomes (*i.e.*, satisfaction, repurchase intent, and behavioral loyalty). While artificial agents are often positioned as a cost-effective and scalable solution, our research reveals an unintended consequence: AI's inability to convey genuine empathy can make even an otherwise effective service recovery feel inadequate to customers. This suggests that blindly replacing human agents with AI, even in structured service recovery processes, may backfire in ways that firms do not anticipate.

A prevailing assumption in voice-driven AI design is that making AI more human-like, through conversational language, emotional scripting, or synthetic empathy, can mitigate its limitations (Huang & Rust, 2024). However, our findings suggest that customers recognize the AI's attempt at mimicking human empathy as less customer-oriented, due to parametric reductionism bias (Valenzuela et al., 2024). Instead of focusing on replicating intrinsic human-related feeling skills, companies might benefit from positioning AI's thinking skills as a competent, efficient, and consistent problem-solver in service recoveries.

Rather than treating AI and human agents as interchangeable, firms should consider service recovery models based on skills (thinking *vs*. feeling), where AI efficiently processes and resolves the issue, but human agents step in to provide emotional reassurance when necessary. For example, AI can handle thinking skills-based problem resolution (*e.g.*, refunds, technical troubleshooting**)**, while human agents focus on feeling skills-based relationship and emotional repair (*e.g.*, apologies, personalized reassurance). Structuring service recovery around feeling (*vs*. thinking) skillsets allows companies to take advantage of voice-driven AI's strengths**.**

7.3. Limitations and future research

While our study adds insights to the intersection of voice-driven AI, customer orientation, and service recovery, we acknowledge certain limitations that open avenues for future research. First, a common justification for AI adoption is its ability to provide faster service recovery than human agents. Customers may interpret an instant AI-driven resolution as a superficial, automated response rather than a sign of superior efficiency. While our research has not covered how speed affects perceived customer orientation, future research could consider how introducing controlled delays or personalization strategies that mimic the thoughtful, deliberative nature of human responses could counteract this perception.

In addition, while we identified perceived empathy as a mediator between agent type (voice-driven AI *vs.* human) and perceived customer orientation, we did not examine how familiarity with AI and technology might shape these perceptions. Prior research suggests two competing perspectives: on the one hand, greater familiarity with AI and robots is often associated with higher acceptance and perceived usefulness, as individuals develop an understanding of their capabilities (Belanche et al., 2019; Chi et al., 2021). On the other hand, some studies indicate that high familiarity can also breed skepticism or resistance, as individuals with more experience may be more aware of AI's limitations, biases, or lack of human-like qualities (Belanche et al., 2019; Horowitz et al., 2024). Future research should explore how these contrasting effects unfold in AI-driven service recovery, identifying whether familiarity enhances trust and perceived effectiveness or, conversely, reinforces doubts about AI's ability to handle emotional interactions.

Future research could further investigate how individual differences affect customer preferences for voice AI-driven service recovery. While our findings suggest that customers generally prefer human agents in emotionally charged service encounters, this may not hold across all consumer segments. Individuals with high social anxiety, discomfort with confrontation, or a preference for transactional rather than relational service experiences may actually feel more at ease interacting with AI than with human agents. Future studies could explore how psychological traits, past experiences, and situational factors influence the acceptance of AI in service recovery. In addition, research could examine whether offering customers a choice between AI and human support improves overall satisfaction or inadvertently increases decision fatigue.

A limitation of our study is that we focused on voice AI agents executing service recovery flawlessly, assuming that technical accuracy and problem resolution were held constant. However, service agents are not perfect – mistakes are common in service interactions, and when accompanied by a sincere apology, errors can paradoxically strengthen customer trust (Schniter et al., 2013). In contrast, AI is typically held to higher standards of perfection, and even minor errors can be perceived as incompetence rather than human-like fallibility. Future research could explore whether **s**trategically placed, low-stakes AI mistakes – such as slight misinterpretations, minor corrections, or even AI hallucinations – could enhance AI's perceived authenticity and trustworthiness or if this effect applies only to interactions with human agents. While intuitive thinking suggests that AI should deliver flawless performance, could small imperfections make AI interactions feel more natural and relatable?

Finally, as AI continues to be applied across all industries, it is possible that, in the near future, AI-driven interactions may no longer be perceived as cold or impersonal but instead signal competence and innovation. While our findings suggest that AI-driven service recovery currently faces limitations due to a perceived lack of empathy, future research could explore whether shifting societal norms and increasing AI integration will change consumer expectations over time.

Overall, we call for future research to move beyond traditional comparisons between voice-driven AI and human agents and instead explore how voice-driven AI can be optimized to enhance customer perceptions. Our findings highlight important challenges in voice-driven AI service recovery, but they also open avenues for investigating when and how AI agents can paradoxically increase trust, emotional connection, and customer loyalty. Future studies should examine the evolving expectations consumers have for AI interactions, the long-term effects of AI-driven service recovery on customer-provider relationships, and the potential for hybrid models that balance AI efficiency with human warmth.

References:

Alabed, A., Javornik, A., Gregory-Smith, D., & Casey, R. (2023). More than just a chat: A taxonomy of consumers' relationships with conversational AI agents and their well-being implications. *European Journal of Marketing*, *58*(2), 373–409. https://doi.org/10.1108/EJM-01-2023-0037

Andreassen, T. W. (2000). Antecedents to satisfaction with service recovery. *European Journal of Marketing*, *34*(1/2), 156–175. https://doi.org/10.1108/03090560010306269

Ariffin, A. A. M. (2013). Generic dimensionality of hospitality in the hotel industry: A host–guest relationship perspective. *International Journal of Hospitality Management*, *35*, 171–179. https://doi.org/10.1016/j.ijhm.2013.06.002

Bag, S., Gupta, S., Kumar, A., & Sivarajah, U. (2021). An integrated artificial intelligence framework for knowledge creation and B2B marketing rational decision making for improving firm performance. *Industrial Marketing Management*, *92*, 178–189. https://doi.org/10.1016/j.indmarman.2020.12.001

Bagozzi, R. P., Brady, M. K., & Huang, M.-H. (2022). AI Service and Emotion. *Journal of Service Research*, *25*(4), 499–504. https://doi.org/10.1177/10946705221118579

Basso, K., & Pizzutti, C. (2016). Trust Recovery Following a Double Deviation. *Journal of Service Research*, *19*(2), 209–223. https://doi.org/10.1177/1094670515625455

Belanche, D., Casaló, L. V., & Flavián, C. (2019). Artificial Intelligence in FinTech: Understanding robo-advisors adoption among customers. *Industrial Management and Data Systems*, *119*(7), 1411–1430. https://doi.org/10.1108/IMDS-08-2018-0368

Belanche, D., Casaló, L. V., Flavián, C., & Schepers, J. (2020). Service robot implementation: A theoretical framework and research agenda. *The Service Industries Journal, 40*(3–4), 203–225.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, *3*(21), eaat5954.

Bergner, A. S., Hildebrand, C., & Häubl, G. (2023). Machine Talk: How Verbal Embodiment in Conversational AI Shapes Consumer–Brand Relationships. *Journal of Consumer Research*, *50*(4), 742–764. https://doi.org/10.1093/jcr/ucad014

Blocker, C. P., Flint, D. J., Myers, M. B., & Slater, S. F. (2011). Proactive customer orientation and its role for creating customer value in global markets. *Journal of the Academy of Marketing Science*, *39*(2), 216–233. https://doi.org/10.1007/s11747-010-0202-9

Brady, M. K., & Cronin, J. J. (2001). Customer Orientation: Effects on Customer Service Perceptions and Outcome Behaviors. *Journal of Service Research*, *3*(3), 241–251. https://doi.org/10.1177/109467050133005

Cartabuke, M., Westerman, J. W., Bergman, J. Z., Whitaker, B. G., Westerman, J., & Beekun, R. I. (2019). Empathy as an Antecedent of Social Justice Attitudes and Perceptions. *Journal of Business Ethics*, *157*(3), 605–615. https://doi.org/10.1007/s10551-017-3677-1

Chandra, S., Shirish, A., & Srivastava, S. C. (2022). To Be or Not to Be …Human? Theorizing the Role of Human-Like Competencies in Conversational Artificial Intelligence Agents. *Journal of Management Information Systems*, *39*(4), 969–1005. https://doi.org/10.1080/07421222.2022.2127441

Chen, A., Pan, Y., Li, L., & Yu, Y. (2022). Are you willing to forgive AI? Service recovery from medical AI service failure. *Industrial Management & Data Systems*, *122*(11), 2540–2557. https://doi.org/10.1108/IMDS-12-2021-0801

Chen, C., Liao, M., Walther, J. B., & Sundar, S. S. (2024). When an AI Doctor Gets Personal: The Effects of Social and Medical Individuation in Encounters With Human and AI Doctors. *Communication Research*, *51*(7), 747–781. https://doi.org/10.1177/00936502241263482

Chi, O. H., Jia, S., Li, Y., & Gursoy, D. (2021). Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (AI) social robots in service delivery. *Computers in Human Behavior*, *118*, 106700. https://doi.org/10.1016/j.chb.2021.106700

Chih, W.-H., Wang, K.-Y., & Banda, H. W. (2025). Chatbots at the frontline: Unveiling antecedents of customers' willingness to accept chatbot intervention in service recovery. *Journal of Retailing and Consumer Services*, *84*, 104254. https://doi.org/10.1016/j.jretconser.2025.104254

Chung, M., Ko, E., Joung, H., & Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*, *117*, 587–595. https://doi.org/10.1016/j.jbusres.2018.10.004

Crolic, C., Thomaz, F., Hadi, R., & Stephen, A. T. (2022). Blame the Bot: Anthropomorphism and Anger in Customer–Chatbot Interactions. *Journal of Marketing*, *86*(1), 132–148. https://doi.org/10.1177/00222429211045687

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24–42. https://doi.org/10.1007/s11747-019-00696-0

De Matos, C. A., Henrique, J. L., & Alberto Vargas Rossi, C. (2007). Service Recovery Paradox: A Meta-Analysis. *Journal of Service Research*, *10*(1), 60–77. https://doi.org/10.1177/1094670507303012

DeWitt, T., & Brady, M. K. (2003). Rethinking Service Recovery Strategies: The Effect of Rapport on Consumer Responses to Service Failure. *Journal of Service Research*, *6*(2), 193–207. https://doi.org/10.1177/1094670503257048

DeWitt, T., Nguyen, D. T., & Marshall, R. (2008). Exploring Customer Loyalty Following Service Recovery. *Journal of Service Research*, *10*(3), 269–281. https://doi.org/10.1177/1094670507310767

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Fan, H., Han, B., & Wang, W. (2024). Aligning (In)Congruent Chatbot–Employee Empathic Responses with Service-Recovery Contexts for Customer Retention. *Journal of Travel Research*, *63*(8), 1870–1893. https://doi.org/10.1177/00472875231201505

Gabbott, M., Tsarenko, Y., & Wai Hoe Mok. (2011). Emotional Intelligence as a Moderator of Coping Strategies and Service Outcomes in Circumstances of Service Failure. *Journal of Service Research*, *14*(2), 234–248. https://doi.org/10.1177/1094670510391078

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, *97*, 304–316. https://doi.org/10.1016/j.chb.2019.01.020

Grandey, A. A., Fisk, G. M., Mattila, A. S., Jansen, K. J., & Sideman, L. A. (2005). Is “service with a smile” enough? Authenticity of positive displays during service encounters. *Organizational Behavior and Human Decision Processes, 96*(1), 38–55.

Grewal, D., Roggeveen, A. L., & Nordfält, J. (2017). The future of retailing. *Journal of Retailing, 93*(1), 1-6. https://doi.org/10.1016/j.jretai.2016.12.008

Grewal, D., Roggeveen, A. L., & Tsiros, M. (2008). The effect of compensation on repurchase intentions in service recovery. *Journal of Retailing, 84*(4), 424–434.

Gustafsson, A. (2009). Customer satisfaction with service recovery. *Journal of Business Research, 62*(11), 1220–1222.

Han, E., Yin, D., & Zhang, H. (2023). Bots with Feelings: Should AI Agents Express Positive Emotion in Customer Service? *Information Systems Research*, *34*(3), 1296–1311. https://doi.org/10.1287/isre.2022.1179

Haupt, M., Rozumowski, A., Freidank, J., & Haas, A. (2023). Seeking empathy or suggesting a solution? Effects of chatbot messages on service failure recovery. *Electronic Markets*, *33*(1), 56. https://doi.org/10.1007/s12525-023-00673-0

Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition: A Regression-Based Approach*. Guilford Publications.

Heo, M., & Lee, K. J. (2018). Chatbot as a New Business Communication Tool: The Case of Naver TalkTalk. *Business Communication Research and Practice*, *1*(1), 41–45. https://doi.org/10.22682/bcrp.2018.1.1.41

Hess Jr., R. L., Ganesan, S., & Klein, N. M. (2003). Service Failure and Recovery: The Impact of Relationship Factors on Customer Satisfaction. *Journal of the Academy of Marketing Science*, *31*(2), 127–145. https://doi.org/10.1177/0092070302250898

Ho, M. H.-W., & Chung, H. F. L. (2020). Customer engagement, customer equity and repurchase intention in mobile apps. *Journal of Business Research*, *121*, 13–21. https://doi.org/10.1016/j.jbusres.2020.07.046

Homburg, C., Wieseke, J., & Bornemann, T. (2009). Implementing the Marketing Concept at the Employee-Customer Interface: The Role of Customer Need Knowledge. *Journal of Marketing*, *73*, 64–81.

Horowitz, M. C., Kahn, L., Macdonald, J., & Schneider, J. (2024). Adopting AI: How familiarity breeds both trust and contempt. *AI & SOCIETY*, *39*(4), 1721–1735. https://doi.org/10.1007/s00146-023-01666-5

Hsu, C.-L., & Lin, J. C.-C. (2023). Understanding the user satisfaction and loyalty of customer service chatbots. *Journal of Retailing and Consumer Services*, *71*, 103211. https://doi.org/10.1016/j.jretconser.2022.103211

Huang, D., Markovitch, D. G., & Stough, R. A. (2024). Can chatbot customer service match human service agents on customer satisfaction? An investigation in the role of trust. *Journal of Retailing and Consumer Services*, *76*, 103600. https://doi.org/10.1016/j.jretconser.2023.103600

Huang, M. H., & Rust, R. T. (2024). The Caring Machine: Feeling AI for Customer Care. *Journal of Marketing*, *88*(5), 1–23. https://doi.org/10.1177/00222429231224748

Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172. https://doi.org/10.1177/1094670517752459

Huang, M.-H., & Rust, R. T. (2022). A Framework for Collaborative Artificial Intelligence in Marketing. *Journal of Retailing*, *98*(2), 209–223. https://doi.org/10.1016/j.jretai.2021.03.001

Huang, M.-H., Rust, R., & Maksimovic, V. (2019). The Feeling Economy: Managing in the Next Generation of Artificial Intelligence (AI). *California Management Review*, *61*(4), 43–65. https://doi.org/10.1177/0008125619863436

Jörling, M., Böhm, R., & Paluch, S. (2019). Service Robots: Drivers of Perceived Responsibility for Service Outcomes. *Journal of Service Research*, *22*(4), 404–420. https://doi.org/10.1177/1094670519842334

Juquelier, A., Poncin, I., & Hazée, S. (2025). Empathic chatbots: A double-edged sword in customer experiences. *Journal of Business Research*, *188*, 115074. https://doi.org/10.1016/j.jbusres.2024.115074

Kecht, C., Egger, A., Kratsch, W., & Röglinger, M. (2023). Quantifying chatbots' ability to learn business processes. *Information Systems*, *113*, 102176. https://doi.org/10.1016/j.is.2023.102176

Kemper, T. D., & Lazarus, R. S. (1992). Emotion and Adaptation. *Contemporary Sociology*, *21*(4), 522. https://doi.org/10.2307/2075902

Kennedy, K. N., Goolsby, J. R., & Arnould, E. J. (2003). Implementing a Customer Orientation: Extension of Theory and Application. *Journal of Marketing*, *67*(4), 67–81. https://doi.org/10.1509/jmkg.67.4.67.18682

Khoa, D. T., Gip, H. Q., Guchait, P., & Wang, C.-Y. (2023). Competition or collaboration for human–robot relationship: A critical reflection on future cobotics in hospitality. *International Journal of Contemporary Hospitality Management*, *35*(6), 2202–2215. https://doi.org/10.1108/IJCHM-04-2022-0434

Kim, S. S., Kim, J., Badu-Baiden, F., Giroux, M., & Choi, Y. (2021). Preference for robot service or human service in hotels? Impacts of the COVID-19 pandemic. *International Journal of Hospitality Management*, *93*, 102795. https://doi.org/10.1016/j.ijhm.2020.102795

Kirk, D., Van Eijnatten, E., & Camps, G. (2023). Comparison of Answers between ChatGPT and Human Dieticians to Common Nutrition Questions. *Journal of Nutrition and Metabolism*, *2023*. Scopus. https://doi.org/10.1155/2023/5548684

Kozinets, R. V., & Gretzel, U. (2021). Commentary: Artificial intelligence: The marketer's dilemma. *Journal of Marketing*, *85*(1), 156–159.

Leo, X., & Huh, Y. E. (2020). Who gets the blame for service failures? Attribution of responsibility toward robot versus human service providers and service firms. *Computers in Human Behavior*, *113*, 106520. https://doi.org/10.1016/j.chb.2020.106520

Liang, S., Li, R., Lan, B., Chu, Y., Zhang, M., & Li, L. (2024). Untouchable them: The effect of chatbot gender on angry customers. *Journal of Research in Interactive Marketing*, *18*(6), 1099–1135. https://doi.org/10.1108/JRIM-02-2023-0061

Liu, B., & Sundar, S. S. (2018). Should Machines Express Sympathy and Empathy? Experiments with a Health Advice Chatbot. *Cyberpsychology, Behavior, and Social Networking*, *21*(10), 625–636. https://doi.org/10.1089/cyber.2018.0110

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*, mksc.2019.1192. https://doi.org/10.1287/mksc.2019.1192

Lv, X., Liu, Y., Luo, J., Liu, Y., & Li, C. (2021). Does a cute artificial intelligence assistant soften the blow? The impact of cuteness on customer tolerance of assistant service failure. *Annals of Tourism Research*, *87*, 103114. https://doi.org/10.1016/j.annals.2020.103114

Lv, X., Yang, Y., Qin, D., Cao, X., & Xu, H. (2022). Artificial intelligence service recovery: The role of empathic response in hospitality customers' continuous usage intention. *Computers in Human Behavior*, *126*, 106993. https://doi.org/10.1016/j.chb.2021.106993

Marinova, D., de Ruyter, K., Huang, M. H., Meuter, M. L., & Challagalla, G. (2017). Getting smart: Learning from technology-empowered frontline interactions. *Journal of Service Research, 20*(1), 29–42.

Markovitch, D. G., Stough, R. A., & Huang, D. (2024). Consumer reactions to chatbot versus human service: An investigation in the role of outcome valence and perceived empathy. *Journal of Retailing and Consumer Services*, *79*, 103847. https://doi.org/10.1016/j.jretconser.2024.103847

Mattila, A. S. (2001). The effectiveness of service recovery in a multi-industry setting. *Journal of Services Marketing, 15*(7), 583–596.

Mattila, A. S., & Enz, C. A. (2002). The role of emotions in service encounters. *Journal of Service Research, 4*(4), 268–277.

Maxham III, J. G., & Netemeyer, R. G. (2002). A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts. *Journal of Marketing, 66*(4), 57–71.

Mende, M., Scott, M. L., Van Doorn, J., Grewal, D., & Shanks, I. (2019). Service robots rising: How humanoid robots influence service experiences and elicit compensatory consumer responses. *Journal of Marketing Research, 56*(4), 535–556.

Miller, J. L., Craighead, C. W., & Karwan, K. R. (2000). Service recovery: A framework and empirical investigation. *Journal of Operations Management*, *18*(4), 387–400. https://doi.org/10.1016/S0272-6963(00)00032-2

Moorman, C., Van Heerde, H. J., Moreau, C. P., & Palmatier, R. W. (2024). Marketing in the Health Care Sector: Disrupted Exchanges and New Research Directions. *Journal of Marketing*, *88*(1), 1–14. https://doi.org/10.1177/00222429231213154

Palmatier, R. W., Scheer, L. K., & Steenkamp, J.-B. E. M. (2007). Customer Loyalty to Whom? Managing the Benefits and Risks of Salesperson-Owned Loyalty. *Journal of Marketing Research*, *44*(2), 185–199. https://doi.org/10.1509/jmkr.44.2.185

Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, *42*(6), 1051–1068. https://doi.org/10.1037/0022-3514.42.6.1051

Pavone, G., Meyer-Waarden, L., & Munzel, A. (2023). Rage Against the Machine: Experimental Insights into Customers’ Negative Emotional Responses, Attributions of Responsibility, and Coping Strategies in Artificial Intelligence–Based Service Failures. *Journal of Interactive Marketing*, *58*(1), 52–71. https://doi.org/10.1177/10949968221134492

Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and Artificial Intelligence: An Experiential Perspective. *Journal of Marketing*, *85*(1), 131–151. https://doi.org/10.1177/0022242920953847

Roy, R., & Naidoo, V. (2021). Enhancing chatbot effectiveness: The role of anthropomorphic conversational styles and time orientation. *Journal of Business Research*, *126*, 23–34. https://doi.org/10.1016/j.jbusres.2020.12.051

Rust, R. T., Huang, M.-H., Rust, R. T., & Huang, M.-H. (2021). *The feeling economy*. Springer.

Sands, S., Ferraro, C., Campbell, C., & Tsao, H.-Y. (2021). Managing the human–chatbot divide: How service scripts influence service experience. *Journal of Service Management*, *32*(2), 246–264. https://doi.org/10.1108/JOSM-06-2019-0203

Schniter, E., Sheremeta, R. M., & Sznycer, D. (2013). Building and rebuilding trust with promises and apologies. *Journal of Economic Behavior & Organization*, *94*, 242–256. https://doi.org/10.1016/j.jebo.2012.09.011

Schuetz, S., & Venkatesh, V. (2020). The rise of human machines: How cognitive computing systems challenge assumptions of user-system interaction. *Journal of the Association for Information Systems*, *21*(2), 460–482. https://ssrn.com/abstract=3680306

Sheehan, B., Jin, H. S., & Gottlieb, U. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research*, *115*, 14–24. https://doi.org/10.1016/j.jbusres.2020.04.030

Shuqair, S., Pinto, D. C., & Mattila, A. S. (2019). Benefits of authenticity: Post-failure loyalty in the sharing economy. *Annals of Tourism Research*, *78*, 102741. https://doi.org/10.1016/j.annals.2019.06.008

Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery. *Journal of Marketing Research*, *36*, 356–372.

Song, M., Zhang, H., Xing, X., & Duan, Y. (2023). Appreciation vs. apology: Research on the influence mechanism of chatbot service recovery based on politeness theory. *Journal of Retailing and Consumer Services*, *73*, 103323. https://doi.org/10.1016/j.jretconser.2023.103323

Tax, S. S., Brown, S. W., & Chandrashekaran, M. (1998). Customer evaluations of service complaint experiences: Implications for relationship marketing. *Journal of Marketing, 62*(2), 60–76.

Valenzuela, A., Puntoni, S., Hoffman, D., Castelo, N., De Freitas, J., Dietvorst, B., Hildebrand, C., Huh, Y. E., Meyer, R., Sweeney, M. E., Talaifar, S., Tomaino, G., & Wertenbroch, K. (2024). How Artificial Intelligence Constrains the Human Experience. *Journal of the Association for Consumer Research*, *9*(3), 241–256. https://doi.org/10.1086/730709

Valuates Reports. (2021, September). *Conversational AI Market Statistics 2030*. Valuates Reports. https://reports.valuates.com/reports/ALLI-Manu-0Z68/conversational-ai

Van Der Goot, M. J., Koubayová, N., & Van Reijmersdal, E. A. (2024). Understanding users' responses to disclosed vs. undisclosed customer service chatbots: A mixed methods study. *AI & SOCIETY*, *39*(6), 2947–2960. https://doi.org/10.1007/s00146-023-01818-7 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. https://doi.org/10.1177/1094670516679272

Vorobeva, D., El Fassi, Y., Costa Pinto, D., Hildebrand, D., Herter, M. M., & Mattila, A. S. (2022). Thinking Skills Don't Protect Service Workers from Replacement by Artificial Intelligence. *Journal of Service Research*, *25*(4), 601–613. https://doi.org/10.1177/10946705221104312

Wagner, R., Costa Pinto, D., Hildebrand, D., Pacheco, N. P., Dhillon, G., & Herter, M. (2025). Algorithm delegation: How embedded AI facilitates agency transference in medical services. *Psychology and Marketing*.

Wei, C., Liu, M. W., & Keh, H. T. (2020). The road to consumer forgiveness is paved with money or apology? The roles of empathy and power in service recovery. *Journal of Business Research*, *118*, 321–334. https://doi.org/10.1016/j.jbusres.2020.06.061

Wieseke, J., Geigenmüller, A., & Kraus, F. (2012). On the Role of Empathy in Customer-Employee Interactions. *Journal of Service Research*, *15*(3), 316–331. https://doi.org/10.1177/1094670512439743

Wilder, K. M., Collier, J. E., & Barnes, D. C. (2014). Tailoring to Customers' Needs: Understanding How to Promote an Adaptive Service Experience With Frontline Employees. *Journal of Service Research*, *17*(4), 446–459. https://doi.org/10.1177/1094670514530043

Xu, X., & Liu, J. (2022). Artificial intelligence humor in service recovery. *Annals of Tourism Research*, *95*, 103439. https://doi.org/10.1016/j.annals.2022.103439

Yun, J., & Park, J. (2022). The Effects of Chatbot Service Recovery With Emotion Words on Customer Satisfaction, Repurchase Intention, and Positive Word-Of-Mouth. *Frontiers in Psychology*, *13*, 922503. https://doi.org/10.3389/fpsyg.2022.922503

Zaki, H. S., & Al-Romeedy, B. S. (2024). Chatbot symbolic recovery and customer forgiveness: A moderated mediation model. *Journal of Hospitality and Tourism Technology*, *15*(4), 610–628. https://doi.org/10.1108/JHTT-11-2023-0374

Zhang, J., Lu, J., Wang, X., Liu, L., & Feng, Y. (2024). Emotional expressions of care and concern by customer service chatbots: Improved customer attitudes despite perceived inauthenticity. *Decision Support Systems*, *186*, 114314. https://doi.org/10.1016/j.dss.2024.114314

Zhang, J., Zhu, Y., Wu, J., & Yu-Buck, G. F. (2023). A natural apology is sincere: Understanding chatbots' performance in symbolic recovery. *International Journal of Hospitality Management*, *108*, 103387. https://doi.org/10.1016/j.ijhm.2022.103387

Zhou, C., & Chang, Q. (2024). Informational or emotional? Exploring the relative effects of chatbots' self-recovery strategies on consumer satisfaction. *Journal of Retailing and Consumer Services*, *78*, 103779. https://doi.org/10.1016/j.jretconser.2024.103779

List of figures

Figure 1

Theoretical model of research.

[see separate file]

Figure 2

The impact of agent type on Perceived Customer Orientation and Positive Service Outcomes (Study 1).

Caption: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

[see separate file]

Figure 3

The impact of agent type on perceived customer orientation and positive service outcomes (Study 2).

Caption: * $p < 0.05$; ** $p < 0.01$.

[see separate file]

Figure 4

The impact of agent and task type on perceived customer orientation (Study 3).

Caption: n.s. – non significant; *** $p < 0.001$.

[see separate file]

Table 2
Measurement scales used across studies

Measures	Cronbach's α S1	S2	S3
Perceived Customer Orientation adapted from Homburg et al. (2009) — Mario tried to find out what my needs were. — Mario had my best interests in mind. — Mario took a problem-solving approach with me. — Mario recommended the most suitable solutions to solve the problems. — Mario tried to figure out what kinds of solutions would be most useful to me.	.89	.87	.87
Perceived empathy adapted from Wilder et al. (2014) — Mario tries to empathize with the customers' feelings. — It's easy for Mario to see things from the customer's perspective. — Mario tries to put himself in the "customer's shoes". — Mario tries to understand the customer's point of view.	.92	.96	.94
Customer satisfaction adapted from Wieseke et al. (2012) — All in all, I am very satisfied with my call to this company. — The call to this company met my expectations of an ideal customer support encounter. — The service encounter fulfilled my expectations.	.81	.95	*
Behavioral loyalty adapted from Shuqair et al. (2019) — I consider this company as my first choice. — I will continue to purchase from this company. — I will not buy products from this company in the future. (R)	.78	*	*
Repurchase intentions adapted from Ho & Chung (2020) — I would intend to continue purchasing this company's products in the future. — I would like to recommend this company's products to others, even if they are existing customers. — I would look forward to the new product launches by this company and its associated suppliers.	*	.89	*

— I would like to have first-hand information about this company's new products.			
Positive Word-of-Mouth adapted from Kirk & Givi (2025)			
— Say positive things about this company. — Recommend this company to others. — Recommend this company to someone else who seeks my advice.	*	.96	*
*Online Shopping Frequency *** How often do you shop online?			
— Daily — Weekly — Monthly — Yearly — I never shop online		n.a.	
*Recent service failures with online shopping *** In the past 12 months, have you experienced any problems with at least one online purchase?			
— Yes — No — Not sure		n.a.	
*Previous knowledge of AI *** Which of the following best describes your knowledge about AI (Artificial Intelligence)?			
— Never heard of AI — Heard of AI — Basic knowledge of AI — Intermediate knowledge of AI — Advanced knowledge of AI — Active AI research/development		n.a.	

** not included in the study; ** self-created for the study*